Research & Development

2011

Mechanical Engineering Letters, Szent István University
Annual Technical-Scientific Journal of the Mechanical Engineering Faculty,
Szent István University, Gödöllő, Hungary

HU ISSN 2060-3789

Páter K. u. 1., Gödöllő, H-2103 Hungary
dekan@gek.szie.hu, www.gek.szie.hu,

Volume 6 (2011)

SELECTED COLLECTION FROM THE RESEARCH RESULTS OF YEAR 2011

CONTENTS

1. INSTITUTE FOR MATHEMATICS AND INFORMATICS

PROFESSOR DR. SÁNDOR MOLNÁR
DIRECTOR OF INSTITUTE

Dear Reader,

Our Institute consists of three departments: **Department of Mathematics** (head: Prof. Dr. Zoltán Varga) and **Department of Informatics** (head: Prof. Dr. Sándor Molnár). The **External Department of Applied Informatics** (head: Prof. Dr. Andras Edelmayer) is located at the **Computer and Automation Research Institute** of the Hungarian Academy of Sciences (MTA SZTAKI). Our research activities in 2011 were as follows.

Department of Mathematics
In 2011the first year of the a three-year research project on "Dynamic modelling of stage-structured multi-species population systems", funded by the Hungarian Scientific Research Fund (OTKA), has been successfully closed, with 19 publications, 16 of them prepared in international cooperation, with a cumulated impact factor 12.405. Further research fields are mathematical demography, patch detection in forest use, real-time observer design for solar heating systems, marine fishery with reserve area, game-theoretical modelling of economic behaviour, etc. Two PhD theses entitled "Absolute convergence of double trigonometric Fourier series and Walsh-Fourier series" and "On the stability of second order differential equations with step function coefficients" have been successfully defended.

Department of Informatics:
The Department's research activities focused on the following three major fields:

- Application of enviromental informatics and modeling methodologies. A VLS database and a geostatistical model was developed for the Paks NPP in the area of groundwater level forecasting with consideration of river-level trends.
- Analysis of qualitative parameters of dynamical systems. The research was financed by OTKA and provided contributions to time-dependent problems with differential algebraic methods, conditional and phase-limited control problems. The research resulted in 20 articles in international journals and 16 articles in international conference proceedings were published with a cumulated impact factor of 13.469.

– Laguerre representation-based pole-reconstruction method. This method forms a promising approach to the identification of frequency domain models of various systems while allowing good practical applicability.

External Department of Applied Informatics

The Department focuses on system and control theory application, fault detection, distributed networking, advanced filtering methods. Another field of research is intelligent metering and transportation vehicle fleet guidance under traffic conditions.

ESTABLISHMENT OF OPTIMAL REALIZATION-INDEPENDENT COST FUNCTIONS

Sándor MOLNÁR[1], Moira MIRANDA[2], Márk MOLNÁR[3], Alexandros SOUMELIDIS[4]

[1]Department of Informatics, Institute for Mathematics and Informatics
[2]Laboratory of Unit Operations, Department of Chemical Engineering
Faculty of Engineering, University of Los Andes, Mérida, Venezuela
[3]Department of Macroeconomics, Institute of Economics, Szent Istvan University
[4]System and Controls Lab, Computer and Automation Research Institute
Hungarian Academy of Sciences

Abstract

In practical applications systems are usually given in input-output forms. The cost of operating a system depends on its state, input and output variables. Since the same system can be described by different realizations, the cost functions must not depend on the realization considered. In this paper realization-independent cost functions will be introduced and examined, and methods for their optimization will be given.

Keywords

Input-output systems, realizations, verticuum-type systems

1. Introduction

Consider a finite dimensional realization of an input-output system

$$\begin{aligned} &\dot{x} = Ax + Bu, x(0) = 0, \\ &y = Cx \end{aligned} \tag{1}$$

The input-output function can be given as

$$f(u)(t) = C\int_0^t \exp A(t-s)Bu(s)ds \tag{2}$$

Let $D \in \mathbf{R}^{k\times k}, E \in \mathbf{R}^{m\times k}, F \in \mathbf{R}^{m\times m}, d \in \mathbf{R}^{k}, e \in \mathbf{R}^{m}$. The cost function can be defined as follows. For a pair of *(u,y)* of input and output define

$$I(u,y)=\int_0^{T_u}(\langle Dy,y\rangle+\langle Ey,u\rangle+\langle Fu,u\rangle+\langle d,y\rangle+\langle e,u\rangle+1) \tag{3}$$

This cost function is obviously realization-independent. Consider next a realization *S=(A,B,C)* of our input-output system. Obviously *y=Cx,* and therefore the cost function can be rewritten using the state variables.

$$\begin{aligned} &J_S(u,x)=I(u,Cx)= \\ &=\int_0^{T_u}(\langle C^*DCx,x\rangle+\langle ECx,u\rangle+\langle Fu,u\rangle+\langle C^*d,x\rangle+ \\ &+\langle e,u\rangle+1)= \\ &=\int_0^{T_u}(\langle D_SX,x\rangle+E_S\langle x,u\rangle+\langle Fu,u\rangle+\langle d_S,x\rangle+\langle e,u\rangle+1) \end{aligned} \tag{4}$$

LEMMA. *Assume that for realization S=(A,B,C), k≤ n and the rank of C is k, then the realization independent cost function I can be rewritten by using cost function J_S.*

Proof. By assumption D_s, E_s, d_s are known. We shall show how matrices *D*, *E* and vector *d* can be determined. They can be obtained by solving the following equations

$$\begin{aligned} &1.\quad C^*DC=D_S \\ &2.\quad EC=E_S \\ &3.\quad C^*d=d_s \end{aligned} \tag{5}$$

By premultiplying eq. 5/1 by *C* and postmultiplying eq. 5/2 by C^*, and premultiplying eq. 5/3 by *C* we have

$$\begin{aligned} &CC^*D(CC^*)=CD_SC^*, \\ &E(CC^*)=E_SC^*, \\ &CC^*d=Cd_S. \end{aligned} \tag{6}$$

The rank of $CC^*\in\mathbf{R}^{k\times k}$ is *k*, so this matrix is invertible. Therefore

$$\begin{aligned} &D=(CC^*)^{-1}CD_Sc^*(CC^*)^{-1}, \\ &E=E_SC^*(CC^*)^{-1} \\ &d=(CC^*)^{-1}Cd_S. \end{aligned} \tag{7}$$

$$I(u,y)=\int_0^{T_u}(\langle (CC^*)^{-1}CD_Sc^*(CC^*)^{-1}y,y\rangle+\langle E_SC^*(CC^*)^{-1}y,u\rangle+ \\ +\langle Fu,u\rangle+\langle (CC^*)^{-1}Cd_S,y\rangle+\langle e,u\rangle+1)dt. \tag{8}$$

From which the assertion follows.

Next, we shall discuss the problem how cost function J_S changes when realization $S= (A,B,C)$ is transformed to realization $S'=(A',B',C')$.
LEMMA. *Assume that realizations S and S' are connected by transformations*

$$A_{S'}=P^{-1}A_SP, B_{S'}=P^{-1}B_S, C_{S'}=C_SP$$

Then

$$J_{S'}(u,x)=\int_0^{T_u}(\langle D_{S'}x,x\rangle+\langle E_{S'}x,x\rangle+\langle Fu,u\rangle+\langle d_{S'},x\rangle+\langle e,u\rangle+1)=$$

$$=\int_0^{T_u}\left(\langle P^*D_SPx,x\rangle+\langle E_SPx,u\rangle+\langle Fu,u\rangle+\langle P^*d_s,x\rangle+\langle e,u\rangle+1\right)$$

Proof. By using equations 5.1..5.3 we have

$$D_{S'}=C_{S'}{}^*DC_{S'}=(C_SP)^*D(C_SP)=P^*(C_S{}^*DC_S)P=P^*D_SP,$$
$$E_{S'}=EC_{S'}=E(C_SP)=(EC_S)P=E_SP,$$
$$d_{S'}=C_{S'}{}^*d=(C_SP)d=P^*(C_Sd)=P^*d_S,$$

from which the assertion follows.

2. Optimization of input-ouput systems

First an optimum problem concerning input-output systems will be formulated. Let $U \in \mathbf{R}^m$ be a convex set such that $0 \in U$ and $\eta \in \mathbf{R}^k$. A function $u:[0,T_u]\rightarrow U$ is called η-feasible, or simply a feasible control, if

1. U is piecewise continous
2. $f(u)(T_u)= \eta$.

Let the set of η-feasible controls be denoted by M_η. We shall optimize functions $u \rightarrow I(u, f(u))$ on set M_η. Formally the following problem is defined

Minimize (u,y) $\rightarrow$ I(u,y) (7)

Subject to
$y=f(u)$
$y(0)=y(T_u)=\eta$.

A realization-depedent necessary condition is first verified.
THEOREM 1. *Assume that $u_0 \in M_\eta$ is an optimal solution of (7). Let S=(A,B,C) be an arbitrary realization of the input-output system f. Then there exist a vector $v \in \mathbf{R}^k$ and solutions x_0 and $\varphi_0 \neq 0$ of the differential equations*

$$\begin{aligned} &\dot{x} = Ax + Bu_0 \\ &\dot{\varphi} = -A\varphi - 2C^*DCx - C^*E^*u_0 - C^*d \\ &x(0) = 0, Cx(T_u) = \eta, \varphi(T_u) = C^*v \end{aligned}$$

such that in every (one sided) continuity point t of u_0,the minimum principle

$$\min_{u\in U}\{\langle \varphi_0(t), Bu\rangle + \langle ECx_0(t), u\rangle + \langle Fu, u\rangle + \langle e, u\rangle\} =$$

$$= \langle \varphi_0(t), Bu_0(t)\rangle + \langle ECx_0(t), u_0(t)\rangle + \langle e, u_0(t)\rangle. \quad (8)$$

Proof. *Consider the following optimal control problem:*

$$\begin{aligned} &\dot{x} = Ax + Bu \\ &x(0) = 0, Cx(T_u) = \eta \\ &J_S(u,x) = \int_0^{T_u} (\langle D_S x, x\rangle + \langle E_S x, x\rangle + \langle Fu, u\rangle + \langle d_S, x\rangle + \langle e, u\rangle + 1). \end{aligned} \quad (9)$$

It is obvious, that *(u,x)*is feasible if and only if $u \in M_\eta$.Then $y=Cx$ and $g(T_u)=Cx(T_u)=\eta$. Furthermore $J_S(u,v)=I(u,Cx)=I(u,y)$, and therefore (u_0,x_0) is an optimal solution for (9) if and only if $u_0 \in M_\eta$.is an optimal solution for (7). If only if $u_0 \in M_\eta$.is an optimal solution for (7), then (u_0, x_0) satisfies the minimum-principle of Pontryagin associated to problem (9) which is equivalent to relations (8) □
For time optimum problems sufficient conditions can easily be derived. Assume that the functional has the form

$$I(u,x) = \int_0^{T_U} 1 = T_U, \quad (10)$$

then problem (7) is called time-optimum problem. This is a special case, when $0 = D \in r^{k\times k}, 0 = E \in \Re^{m\times k}, 0 = F \in \Re^{m\times m}, 0 = d \in \Re^{k}, 0 = e \in \Re^{m}$. We shall demonstrate that under certain additional conditions the minimum principle is sufficient. This additional condition is that $0 \in U$ is an interior point of U. In this case the following theorem holds.

THEOREM 1. *Let $u_0 \in M_\eta$. Assume that S=(A,B,C) is a controllable realization of input-output system f.If there exist vector $\nu \in \boldsymbol{R}^k$ and solutions x_0, $\phi_0 \neq 0$ of differential equations*

$$\begin{aligned} &\dot{x} = Ax + Bu_0, \\ &\dot{\varphi} = -A^*\varphi, \\ &x(0) = 0, Cx(T_u) = \eta, \varphi(T_u) = C^*\nu \end{aligned} \tag{11}$$

such that the minimum principle

$$\langle \varphi_0(t), Bu_0(t) \rangle = \min_{u\in U} \langle \varphi_0(t), Bu \rangle \tag{12}$$

Proof. Consider the time-optimum problem

$$\begin{aligned} &\dot{x} = Ax + Bu \\ &x(0) = 0, Cx(T_u) = \eta \\ &T_u \to \min! \end{aligned} \tag{13}$$

Since S is controllable and $0 \in U$ is an interior point, we are able to apply the general formulation of the minimum principle of Pontryagin for optima of general positions, that, the process (u_0,x_0) which satisfies the minimum principle is time-optimal, and so, u_0 is optimal for problem (7).

3. Optimization of verticum-type systems

In our earlier papers (Molnár (1987a,b,c), Molnár(1994)) and in recent publications (Gámez et al, 2010) we have seen that under certain conditions input-output systems can be represented by verticuum-type realizations. Assume that realization *S=(A,B,C)* has the special block structure

$$
\begin{aligned}
&\dot{x}_0 = Ax_0 + B_0u_0\\
&\dot{x}_1 = A_1x_1 + A_{10}x_0 + B_1u_1\\
&\vdots\\
&\dot{x}_l = A_lx_l + A_{ll-1}x_{ll-1} + B_lu_l\\
&x(0) = 0, x_1(0) = 0, ..., x_l(0) = 0,\\
&y_0 = C_0x_0,\\
&y_1 = C_1x_1,\\
&\vdots\\
&y_l = C_lx_l
\end{aligned}
\tag{14}
$$

where

$$
\begin{aligned}
&A_i \in \mathbf{R}^{n_i \times n_i}, A_{ii-1} \in \mathbf{R}^{n_i \times n_{i-1}}, B \in \mathbf{R}^{n_i \times m_i},\\
&C_i \in \mathbf{R}^{k_i \times n_i}, \sum_{i=0}^{l} n_i = n, \sum_{i=0}^{l} m_i = m, \sum_{i=0}^{l} k_i = k.
\end{aligned}
\tag{15}
$$

Introduce matrices

$$
\begin{aligned}
D &= \begin{pmatrix} D_{00} & D_{01} & \cdots & D_{0l} \\ D_{10} & & & \\ \vdots & & & \\ D_{l0} & \cdots & & D_{ll} \end{pmatrix},\\
F &= \begin{pmatrix} F_{00} & F_{01} & \cdots & F_{0l} \\ F_{10} & & & \\ \vdots & & & \\ F_{l0} & \cdots & & F_{ll} \end{pmatrix},\\
d &= \begin{pmatrix} d_0 \\ d_1 \\ \\ d_l \end{pmatrix}, E = (E_0, E_1, \ldots, E_l), e = (e_0, e_1, \ldots, e_l)
\end{aligned}
\tag{16}
$$

If conditions

$$
\begin{aligned}
C_i^* D_{ij} C_j &= 0 (i \neq j)\\
F_{ij} &= 0 (i \neq j)
\end{aligned}
$$

hold, then the above matrices are said to be verticum-compatible. Cost function J_s can be then rewritten as

$$J_s(u,x) = \int_0^{T_u} (\langle C^* DCx, x\rangle + \langle ECx, u\rangle + \langle Fu, u\rangle + \langle C^* d, x\rangle + \\ + \langle e, u\rangle + 1) = \sum_{i=0}^{l} \int_0^{T_u} (\frac{1}{l+1} + \langle C_i^* D_{ii} C_i x_i, x_i\rangle + \langle E_i C_i x_i, u_i\rangle + \\ \langle F_{ii} u_i, u_i\rangle + \langle C_i^* d_i, x_i\rangle + \langle e_i, u_i\rangle) = \sum_{i=0}^{l} J_{S_i}(u_i, x_i), \qquad (17)$$

where J_S is the cost function of subsystem i, where the i^{th} control system is defined as

$$\begin{aligned} \dot{x} &= A_i x_i + A_{ii-1} z + B_i u_i \\ x_i(0) &= 0, \\ y_i &= C_i x_i. \end{aligned} \qquad (18)$$

Obviously, cost function J_{Si} depends also on the fixed function z. For any fixed z_i define verticum V_i as follows:

$$\begin{aligned} \dot{x}_i &= A_i x_i + A_{i,i-1} z + B_i u_i \\ \dot{x}_{i+1} &= A_{i+1} x_{i+1} + A_{i+1,i} x_i + B_{i+1} u_{i+1}, \\ &\vdots \\ \dot{x}_l &= A_l x_l + A_{l,l-1} x_{l-1} + B_l u_l \end{aligned} \qquad (19)$$

And define cost function J_{Vi} as

$$J_{V_i}(u_i, u_{i+1}, \ldots, u_l, x_i, x_{i+1}, \ldots, x_l) = \sum_{j=1}^{l} J_{s_j}(u_j, x_j).$$

Assume that $U \subset \mathbf{R^m}$ is a convex set and can be described as the Descartean product $U = \prod_{i=0}^{l} U_i$ of convex sets $U_i \subset \mathbf{R^m_i}$. Let $u : [0, T_u] \to \prod_{i=0}^{l} U_i$ be piecewise continous. A control u is called ($\underline{\eta_i}$,z) feasible with $\underline{\eta_i} = (\eta_i, \eta_{i+1}, \ldots, \eta_l)$, if the solution $x(0)=0, x_{i+1}(0)=0, \ldots, x_l(0)=0$ satisfies $C^* \underline{x_i}(T_u) = \underline{\eta_i}$. Let $M_{(\underline{\eta i},z)}$ denote the ($\underline{\eta_i}$,z) feasible controls. Consider now the following optimization problem:

$$
\begin{aligned}
&\dot{x}_i = A_i x_i + A_{i,i-1} z + B_i u_i \\
&\dot{x}_{i+1} = A_{i+1} x_{i+1} + A_{i+1,i} z + B_{i+1} u_{i+1} \\
&\vdots \\
&\dot{x}_l = A_l x_l + A_{l,l-1} x_{l-1} + B_l u_l, \\
&x_i(0) = 0, \ldots, x_l(0) = 0, C_i x_i(T_u) = \eta_i, \ldots, C_l x_l(T_u) = \eta_l \\
&J_{V_i}(u_i, u_{i+1}, \ldots u_l, x_i, x_{i+1}, \ldots, x_l) \to \min!
\end{aligned}
\tag{20}
$$

Assume that with process $(u_{i,0}, u_{i+1,0}, \ldots, u_{l,0}, x_{i,0}, \ldots, x_{l,0})$ this problem has a minimum. Let the minimum be denoted by

$$
J_{M_i}(z) = \min_{M(\eta_i, z)} J_{V_i}(u_i, u_{i+1}, \ldots, u_l, x_i, x_{i+1}, \ldots, x_l). \tag{21}
$$

By using the principle of dynamic programming, the following is obtained. If $(u_{0,0}, u_{1,0}, \ldots, u_{l,0}, x_{0,0}, \ldots, x_{l,0})$ is optimal, then $(0_{ix_{i-1,0}})$ and $(u_{0,0}, u_{1,0}, \ldots, u_{i-1,0}, x_{0,0}, \ldots, x_{i-1,0})$ is optimal for problem

$$
\begin{aligned}
&\dot{x} = A_0 x_0 + B_0 u_0 \\
&\dot{x}_1 = A_1 x_1 + A_{10} x_0 + B_1 u_1 \\
&\vdots \\
&\dot{x}_{i-1} = A_{i-1} x_{i-1} + A_{i-1,i-2} x_{i-2} + B_{i-1} u_{i-1} \\
&x(0) = 0, \ldots, x_{i-1}(0) = 0, C_0 x_0(T_u) = \eta_0, \ldots \\
&\ldots, C_{i-1} x_{i-1}(T_u) = \eta_{i-1}, \\
&J_{M_1}(x_{i-1}) + \sum_{j=0}^{i-1} J_{S_j}(u_j, x_j) \to \min!
\end{aligned}
\tag{22}
$$

and furthermore

$$
\min \sum_{j=0}^{l} J_{S_j}(u_j, x_j) = (J_{M_i}(x_{i-1})) + \sum_{j=0}^{i-1} J_{S_j}(u_j, x_j) \tag{23}
$$

for all $i=1,2...,$. The corresponding optimum-principle is given next as

THEOREM 3. Assume that $u_0 \in M$ is an optimal for problem (7), and S=(A,B,C) is a verticum-type realization of system f. It is also assumed that the blocks of cost function I are verticum-compatible. Then for all i=1,2...,l there exist vectors $v_j \in \mathbf{R}^{kj}(j \geq i)$ and solutions $(x_{i,0}, \ldots, x_{l,0}, \varphi_{i,0}, \ldots, \varphi_{l,0})$ of differential equations

$$\dot{\phi}_i = -(A_i^* \phi_i + A_{i+1}^* \phi_{i+1} + 2C_i D_{ii} C_i x_i + C_i^* E_i u_{i0} + C_i^* d_i),$$
$$\vdots$$
$$\dot{\phi}_{l-1} = -(A_{l-1}{}^* \phi_{l-1} + A_{l,l-1}^* \phi_l + 2C_{l-1} D_{l-1,l-1} C_{l-1} x_{l-1} + C_{l-1}^* E_{l-1} u_{l-1,0} + C_{l-1}^* d_{l-1})$$
$$\dot{\phi}_l = -(A_l^* \phi_l + 2C_l D_{ll} C_l x_l + C_l^* E_l u_{l,0} + C_l^* d_l)$$
$$x_i(0) = 0, \ldots, x_l(0) = 0, C_i x_i(T_u) = \eta_i, \ldots, C_l x_l(T_u) = \eta_l,$$
$$\phi_i(T_u) = C_i^* v_i, \ldots, \phi_l(T_u) = C^*{}_l v_l$$

such that $0 \neq (\phi_{i,0}, \phi_{i+1,0}, \ldots, \phi_{l,0})$ *and equations*

$$\min_{u_j \in U_j} \sum_{j=1}^{l} (\langle \phi_{j,0}(t), B_j u_j \rangle + \langle E_j C_j x_{j,0}(t), u_j \rangle + \langle F_j u_j, u_j \rangle + \langle e_j, u_j \rangle) =$$
$$= \sum_{j=1}^{l} (\langle \phi_{j,0}(t), B_j u_{j,0}(t) \rangle + \langle E_j C_j x_{j,0}(t), u_{j,0}(t) \rangle + \langle F_j u_{j,0}(t), u_{j,0}(t) \rangle + \langle e_j, u_{j,0}(t) \rangle)$$

are satisfied at all continuity points t of u_0.

Proof. The principle of dynamic programming implies that if u_0 is optimal for (7)then $(u_{i,0}, u_{i+1,0}, \ldots, u_{l,0}, x_{i,0}, \ldots, x_{l,0})$is optimal for $(0_{ix_{i-1,0}})$ of the verticum type realization of $S=(A,B,C)$. The minimum-principle of this optimality implies the assertion.

Remarks. 1. The differential equations of the theorem contain the undefined trajectories x_0. However they are known, since from $u_{0,0}, u_{1,0}, \ldots, u_{l,0}$ they can be determined by solving the corresponding control differential equations. The equations for $x_i, x_{i+1}, \ldots, x_l$ are simply the repetitions of those equations, and the uniqueness of the solutions implies that $x_{i,0}, \ldots, x_{l,0}$ are really the solutions.

2. Assume that the optimum $J_{M_i}(z)$ can be represented in the integral form

$$J_{M_i}(z) = \int_0^{T_u} K_i(t, z(t)) dt$$

then the minimum principle can be also formulated for problem 0_i. If u_0 is optimal for (7), then for all $i=1,2,\ldots l$ there exist vectors $v_j \in \mathbf{R}^{kj}$, $(j \leq i-1)$ and solutions $x_{0,0}, \ldots, x_{i-1,0}$, $\varphi_{0,0}, \ldots, \varphi_{i-1,0}$ of differential equations

$$\dot{x} = A_0 x_0 + B_0 u_{00}$$
$$\dot{x}_1 = A_1 x_1 + A_{10} x_0 + B_1 u_{10}$$
$$\vdots$$
$$\dot{x}_{i-1} = A_{i-1} x_{i-1} + A_{i-1,i-2} x_{i-1} + B_{i-1} u_{i-1,0}$$

$$\dot{\phi}_0 = -A(A_0^* \phi_0 + A_{10}^* \phi_1 + 2C_0^* D_{00} C_0 x_0 + C_0^* E_0 u_{00} + C_0^* d_0)$$
$$\vdots$$
$$\dot{\phi}_{i-2} = (A_{i-2}^* \phi_{i-2} + A_{i-1,i-2}^* \phi_{i-1} + 2C_{i-2}^* D_{i-1,i-2} C_{i-2} x_{i-2} +$$
$$+C_{i-2}^* E_{i-2} u_{i-2,0} + C_{i-2}^* d_{i-2})$$

$$\dot{\phi}_{i-1} = -(A_{i-1}^* \phi_{i-1} + \partial_{x_{i-1}} K_i(., x_{i-1}) + 2C_{i-1}^* D_{i,i-1} C_{i-1} x_{i-1} +$$
$$+C_{i-1}^* E_{i-1} u_{i-1,0} + C_{i-1}^* d_{i-1})$$
$$x_0(0) = 0, \ldots, x_{i-1}(0) = 0, Cx_0(T_u) = \eta_0, \ldots, Cx_{i-1}(0) = \eta_{i-1},$$
$$\phi_0(T_u) = C_0^* v_0, \ldots, \phi_{i-1}(T_u) = C_{i-1}^* v_{i-1}$$

such that $(\phi_{00}, \phi_{10}, \ldots, \phi_{i-1,0}) \neq 0$ and

$$\min_{u_j \in U_j} \sum_{j=0}^{i-1} \left(\left\langle \phi_{j,0}(t), B_j u_j \right\rangle + \left\langle E_j C_j x_{j0}(t), u_j \right\rangle + \left\langle F_j u_j, u_j \right\rangle + \left\langle e_j, u_j \right\rangle \right)$$
$$= \sum_{j=0}^{i-1} (\left\langle \phi_{j,0}(t), B_j u_{j0}(t) \right\rangle + \left\langle E_j C_j x_{j0}(t), u_{j0}(t) \right\rangle + \left\langle F_j u_{j0}(t), u_{j0}(t) \right\rangle +$$
$$+ \left\langle e_j, u_{j0}(t) \right\rangle)$$

at all continuity points of u_0. The condition that the optimum has an integral representation can not be verified easily since even in special cases the relation between the optimum and inital conditions is rather complicated.
Conditions for time optimum can be also formulated for verticum-type realizations. Assume that $0 \in U$ is an interior point of U. If $(0,...,0) \in \prod_i (U_i) = U$ is an interior point if and only if $0 \in U_u$ is interior in each U_i. Under these additional assumptions the following theorem can be verified:

THEOREM 4. Let $u_0 \in M_\eta$ and assume that S=(A,B,C) is a controllable realization of input-output system f. If there exist vectors $v_j \in \mathbf{R}^{kj}$ (j=0...l) and solutions $(x_{0,0}, \ldots, x_{l,0}, \varphi_{0,0}, \ldots, \varphi_{l,0})$ on differential equations

$$
\begin{aligned}
&\dot{x} = A_0 x_0 + B_0 u_{00} \\
&\dot{x}_1 = A_1 x_1 + A_{10} x_0 + B_1 u_{10} \\
&\vdots \\
&\dot{x}_l = A_l x_l + A_{l,l-1} x_{l-1} + B_l u_{l,0} \\
&\dot{\phi}_0 = -A_0^* \phi_0 - A_{1,0}^* \phi_1, \\
&\vdots \\
&\dot{\phi}_{l-1} = -A_{l-1}^* \phi_{l-1} - A_{l,l-1}^* \phi_l \\
&\dot{\phi}_l = -A^* \phi_l
\end{aligned}
$$

such that $(\varphi_{0,0},\ldots,\varphi_{l,0})\neq 0$ and equation

$$\min_{u_j \in U_j} \sum_{j=0}^{l} \left\langle \phi_{j,0}(t), B_j u_j \right\rangle = \sum_{j=i}^{l} \left\langle \phi_{j,0}(t), B_j u_{j,0}(t) \right\rangle$$

holds for all continuity points of u_0,then u_0 is a time-optimal control.

Proof. The assertion follows from Theorem 2.□

Remark. Conditions for controllability of realization S have been given in our earlier papers (Molnar, 1987a,b,c).

References

Molnár, S. (1987a,b,c). Observability and Controllability of Decomposed Systems. Parts I, II, III., Mathematical Analysis and Systems Theory, Vol 5., pp 57.-80.

Molnár S., Szigeti F (1994): On "Verticum"-Type Linear Systems with Time-Dependent Linkage", Applied Mathematics and Computation, Vol. 60., pp. 89-102.,

Gámez M., López I., Szabó I., Varga Z., (2010): Verticum-type systems applied to ecological monitoring, Applied Mathematics and Computation 215 (2010) pp. 3230–3238

IDENTIFYING HARMONICS IN MECHANICAL SYSTEMS BY USING HYPERBOLIC WAVELET CONSTRUCTS

Alexandros SOUMELIDIS[1], Sándor MOLNAR[2], Ferenc SCHIPP[3]

[1]Systems and Control Lab, Computer and Automation Research Institute
Hungarian Academy of Sciences
[2] Department of Informatics, Institute for Mathematics and Informatics
[3]Department of Numerical Analysis, Faculty of Informatics, Eötvös Loránd University

Abstract

Identification and detection of periodic – or almost periodic – phenomena in the behavior of mechanical systems plays fundamental role in system design, quality assessment, and maintenance. An adequate mode to identify periodicities is the derivation of the *poles* of the spectral functions – Fourier-transforms or transfer functions. The classical methods of frequency domain identification – spreading from those using FFT-based spectral estimation to the representations in rational orthogonal bases – cannot ultimately solve this problem. A new method for identifying the poles of functions belonging to the Hardy space H^2 on the unit disc – corresponding to the spectral representation of discrete-time signals and systems – is introduced. The method is based on the hyperbolic wavelet constructions generated on the Blaschke group. An algorithm is outlined that on the basis of frequency domain measurement data results in efficient estimates of the poles. A numerical simulation example is also presented to illustrate the efficiency of the method.

Keywords

System identification, Mechanical systems, Harmonic vibrations, Discrete-time signal representations, Group representations, Wavelets, Hyperbolic geometry

1. Introduction

Identification and detection of periodic – or almost periodic – phenomena in the behavior of mechanical systems plays fundamental role in system design, quality assessment, and maintenance. The elastic mechanical components with inertia in combination with losses arising from friction and drag result in harmonic vibrations that can either be normal component of the system operation or artifact. External forces and moments can influence the system behavior in directions that can enhance the periodic effects – by arising resonances – resulting in harmful phenomena; let us refer to the collapse of the Tacoma Narrows Bridge (Washington, US, 1940) by the cause of a resonance

catastrophe. Applying active components (pneumatic, hydraulic or electrical actuators, motors, drives), i.e. components that inject additional energy into the system, involve further effects that can result in periodic processes.

Identifying periodicities – implying both real periodic processes and damped vibration phenomena – can be performed on the basis of applying the rules of physics (i.e. the Newton axioms, the conservation rules, or the elements of the Lagrangian or Hamiltonian mechanics complementing them with the rules of elasticity, friction, etc.), which can be considered as the use of the paradigms of the deduction. Modal analysis – a classical method used in mechanical engineering – is mainly based on this paradigm, analyses or simulations can be performed on deductively constructed models of the systems; structural or finite element models (FEM) are typically used.

An alternative way of identifying periodicities in mechanical systems arises from using the paradigm of induction, i.e. by acquiring knowledge on the basis of observations, practically concluding upon measurements. Signal processing combined with the principles of systems theory forms the conceptual basis of doing this: analyzing frequency-domain or spectral representations of signals, and the transfer-functions associated with input-output models of systems gives adequate tools for detecting periodicities, since periodic components in signals appear as peaks of different height and width in the spectral functions. The classical tool to detect periodicities is Fourier-transform. Fourier-transform maps the time-domain signals onto spectral functions, hence plays fundamental role in the theory, furthermore, due to its computationally efficient discrete realizations – the algorithm of Fast Fourier-transform (FFT) – it has also become a powerful and popular means of practical spectral analysis methods. One can be state that digital signal processing – especially in the context of spectral analysis – begins with the discovery of the FFT; for the details and the main issues of the classical spectral analysis see e.g. (A.V. Oppenheim at al., 1999).

However spectral analysis based on FFT can efficiently be performed, the exact spectral behavior of the systems cannot be determined in this way. Spectral peaks cannot be separated in many cases, as well as analyzing the height and width of the peaks does not result in unique solutions for determining all the periodicities that are present in the signals.

A deeper view in the spectral behavior of signals and systems can be obtained if the notion of *poles* is introduced. The poles are associated with the transfer function of linear time-invariant systems; the transfer function has the form of rational functions that can be described by sets of poles and zeros. Since digital processing methods are aimed to elaborate, let us focus on discrete-time systems, i.e. to systems with inputs and outputs being sequences. A commonly used class of discrete-time signals are those belonging to the sequence class ℓ^2. The

spectral representation of these signals can be obtained by the so-called *z*-transform that maps stable signals onto the unit disc on the complex plane with poles residing within the unit circle. An efficient mathematical description of the spectral representations of discrete-time signals belonging to stable systems is the theory of Hardy-space H^2 on the unit disc, that can be considered as the orthogonal complement of the space spanned by the *z*-transforms. Since there exist a unique one-to-one mapping between H^2 and its orthogonal complement $H_\perp^2$, the H^2 nomenclature will be used hereinafter in the paper. An excellent introduction to Hardy spaces including also H^2 in the context of signal processing and systems theory can be found e.g. in (C. K. Chui and G. Chen, 1992)

The frequency-domain or spectral functions in the H^2 theory correspond to the boundary functions of the complex functions $F(z) \in H^2$, i.e. to $F(e^{\mathbf{i}\omega})$ where ω denotes the circular frequency related to the Nyquist-frequency belonging to the time-domain sampling rule. It is assumed that the time-domain signals are uniformly sampled with frequency obeying to the Shannon rule of sampling, hence the spectral functions are periodic with period ω_s, where ω_s denotes the circular sample frequency. Any discrete measurement in the frequency domain results data on the unit circle, i.e. for normalized circular frequency ω_k a complex value $F(e^{\mathbf{i}\omega_k})$ that can be interpreted as an amplitude and a phase associated to the given frequency. Deriving the periodic components belonging to a signal or system associated with $F(z)$ function in H^2 on the basis of physically realizable measurements means deriving the the poles of $F(z)$ on the basis of boundary function points $F(e^{\mathbf{i}\omega_k})$. This is an approximation problem in H^2.

In association with vibrating mechanical systems it should be mentioned that the notion of poles is in strict connection with linear time-invariant systems, hence the methods described here can be used in these types of systems. This is the reason that harmonically vibrating systems are referred in many parts of the present paper. Harmonically vibrating mechanical structures – at least as abstraction – occur very frequently everywhere in the industry, transportation, buildings, etc. Using these methods – at least in local sense – in systems that do not obey to linearity or time-invariance is subject of present and future research.

It is obvious that classical FFT-based spectral methods cannot solve the approximation problem set up above, hence cannot derive or estimate the poles belonging to signals and systems. In the 80's the linear model-based spectral methods based on autoregressive (AR) and autoregressive-movingaverage (ARMA) models and their variants resulted in new opportunities. The system

identification methods created on this basis solved many problems and became standard methods built in software packages of scientific computations and became parts of popular textbooks, e.g. (L. Ljung, 1999). Among others the pole estimation problem in system models of predefined structure, i.e. pole estimation has been solved as a parameter-estimation problem. Structure-estimation remained essentially unsolved, some attempts to solve the problem, e.g. the information criteria, resulted only in particular solutions. An explicit method among the parametric identification methods that aims to estimate periodicities in noisy environment is the Prony-method (attributed to baron Gaspard de Prony from the 18-19. century) and a variant of it, the Pisarenko-method (S. L. Marple, 1987).

In the 90's a more general theory of H^2 and H^∞ approximations has been used in the field of system identification that resulted in the principles of rational orthogonal bases (P. S. C. Heuberger at all, 2005). The H^2 approximation problem has been solved as infinite sequence of nested subspaces generated by finite sets of poles. Exact knowledge on the exact pole locations results in finite representations involving unique rational function representation. Hence the efficient use of the rational orthogonal bases in system identification needs a priori information on the pole locations – at least approximately. Special attention paid on the problems of pole selection and validation (J. Bokor at all, 1999, T. O. e Silva, 2005). There exist methods to refine the pole locations starting from an approximate placement of poles (A. Soumelidis, 2002), however the general pole-identification problem has not been solved so far.

Starting from the results of the theory of rational orthogonal bases a new approach of identifying poles of functions belonging to the H^2 space on the unit disc D – denoted by $H^2(\mathrm{D})$ in the forthcoming – will be introduced in the paper. The method is based on the hyperbolic wavelet constructions generated by the group formed upon the Blaschke functions – a key notion in the theory of rational orthogonal bases – and the Laguerre representations. Based upon the hyperbolic metric generated by the Blaschke group a method is set up that can find the poles of the function under consideration, and an algorithm is outlined that – by using frequency domain measurement data – results in efficient estimates of the poles. A numerical simulation example will also be presented to illustrate the efficiency of the method.

2. Laguerre system representations, Blaschke-functions, hyperbolic geometry, wavelets

A key concept in the $H^2(\mathrm{D})$ system and signal representations is the Blaschke function based upon a parameter $b \in \mathrm{D}$, which can be considered as an *inverse pole* ($b = 1/\overline{p}$) of the function. The Blaschke function is defined as

$$B_{\mathfrak{b}}(z) := \varepsilon \frac{z-b}{1-\bar{b}z} \quad (z \in \mathrm{C},\ \mathfrak{b} = (b,\varepsilon) \in \mathrm{B} := \mathrm{D}\times\mathrm{T}), \tag{2.1}$$

where D and T denotes the open unit disc and the unit circle, respectively. If $\mathfrak{b} \in \mathrm{B}$, then $B_{\mathfrak{b}}$ is an $1-1$ map on T and D, respectively. The restrictions of the Blaschke functions on the set D or T with the operation $(B_{\mathfrak{b}_1} \circ B_{\mathfrak{b}_2})(z) := B_{\mathfrak{b}_1}(B_{\mathfrak{b}_2}(z))$ form a group. In the set of the parameters $\mathrm{B} := \mathrm{D}\times\mathrm{T}$ let us define the operation induced by the function composition in the following way $B_{\mathfrak{b}_1} \circ B_{\mathfrak{b}_2} = B_{\mathfrak{b}_1 \circ \mathfrak{b}_2}$. The group $(\mathrm{B}, \circ)$ will be isomorphic with the group $(\{B_{\mathfrak{b}}, \mathfrak{b} \in \mathrm{B}\}, \circ)$. The neutral element of the group $(\mathrm{B}, \circ)$ is $\mathfrak{e} := (0,1) \in \mathrm{B}$ and the inverse element of $\mathfrak{b} = (b,\varepsilon) \in \mathrm{B}$ is $\mathfrak{b}^{-1} = (-b\varepsilon, \bar{\varepsilon})$.
It can be proved that the map

$$\rho(z_1, z_2) := \frac{|z_1 - z_2|}{|1 - \bar{z}_1 z_2|} = |B_{z_1}(z_2)| \qquad (B_{z_1} := B_{(z_1,1)}, z_1, z_2 \in \mathrm{D}) \tag{2.2}$$

is a metric on D, called pseudohyperbolic metric (see e.g. (L. Ahlfors, 1973)). Moreover the Blaschke functions B_b ($b \in \mathrm{D}$) are isometries with respect to this metric, i.e.

$$\rho(B_b(z_1), B_b(z_2)) = \rho(z_1, z_2) \quad (b \in \mathrm{D}, z_1, z_2 \in \mathrm{D}). \tag{2.3}$$

The lines in this model are the sets

$$\mathrm{L}_{\mathfrak{b}} := \{B_{\mathfrak{b}}(r) : -1 < r < 1\} \quad (\mathfrak{b} \in \mathrm{B}),$$

i.e. circles crossing perpendicularly the unit circle. This model is known in the hyperbolic geometry as the unit disc Poincaré model.

The discrete Laguerre-functions are defined by

$$L_n^b(z) := \frac{\sqrt{1-|b|^2}}{1-\bar{b}z}\left(\frac{z-b}{1-\bar{b}z}\right)^n \quad (z \in \overline{\mathrm{D}}, b \in \mathrm{D}, n \in \mathrm{N}). \tag{2.4}$$

Using the function

$$F_{\mathfrak{b}}(z) := \frac{\sqrt{\varepsilon(1-|b|^2)}}{1-\bar{b}z} \quad (\mathfrak{b} := (b,\varepsilon) \in \mathrm{B}, b \in \mathrm{D}, \varepsilon \in T, z \in \overline{\mathrm{D}}), \tag{2.5}$$

and the Blaschke maps according to (2.1), the discrete Laguerre-functions can be expressed in the form

$$L_n^b = F_{\mathfrak{b}} B_{\mathfrak{b}}^n \quad (\mathfrak{b} = (b,1) \in \mathrm{B}, n \in \mathrm{N}). \tag{2.6}$$

The Hardy space H^2 forms also a Hilbert space with the scalar product

$$\langle f, g \rangle := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{it})\overline{g}(e^{it})\,dt \quad (f, g \in H^2(\mathrm{T})). \tag{2.7}$$

We introduce the collection of operators $U_{\mathfrak{b}} : H^2(\mathrm{T}) \to H^2(\mathrm{B})$ $(\mathfrak{b} \in \mathrm{B})$ defined by

$$U_{\mathfrak{b}} f := F_{\mathfrak{b}^{-1}}\, f \circ B_{\mathfrak{b}^{-1}} \quad (f \in H^2(\mathrm{T}), \mathfrak{b} \in \mathrm{B}). \tag{2.8}$$

It is known that $U_{\mathfrak{b}}$, $(\mathfrak{b} \in \mathrm{B})$ is a unitary representation of the Blaschke group B (M. Pap and F. Schipp, 2006), i.e.

(i) $U_{\mathfrak{b}_1}(U_{\mathfrak{b}_2} f)) = U_{\mathfrak{b}_1 \circ \mathfrak{b}_2} f$,

(ii) $\langle U_{\mathfrak{b}} f, U_{\mathfrak{b}} g \rangle = \langle f, g \rangle$ $(f, g \in H^2(\mathrm{T}))$.

The discrete Laguerre-functions L_n^b can be introduced as image of the power function $h_n(z) := z^n$ by the representation $U_{\mathfrak{b}}$:

$$L_n^b := U_{\mathfrak{b}}^{-1} h_n \quad (n \in \mathrm{N}, \mathfrak{b} = (b,1) \in \mathrm{T}). \tag{2.9}$$

Since $U_{\mathfrak{b}}$ is unitary, $U_{\mathfrak{b}}^* = U_{\mathfrak{b}}^{-1} = U_{\mathfrak{b}^{-1}}$ and consequently for any $m, n \in \mathrm{N}$

(i) $\langle L_n^b, L_m^b \rangle = \langle U_{\mathfrak{b}^{-1}} h_n, U_{\mathfrak{b}^{-1}} h_m \rangle = \langle h_n, h_m \rangle = \delta_{mn}$

(ii) $\langle f, L_n^b \rangle = \langle f, U_{\mathfrak{b}^{-1}} h_n \rangle = \langle U_{\mathfrak{b}} f, h_n \rangle$.

Thus the discrete Laguerre-Fourier coefficients of f are equal to the trigonometric Fourier coefficients of the function $U_b f$. This relation can be used to compute the discrete Laguerre-Fourier coefficients.

The representation of any function $f \in H^2(\mathbb{D})$ in the Laguerre-system can be expressed as

$$f(z) = \sum_{n=0}^{\infty} c_{n,f} L_n^b(z), \tag{2.10}$$

where $c_{n,f}$ coefficients – i.e. the so called Laguerre-Fourier coefficients belonging to function f – can be computed by the scalar products $\langle f, L_n^b \rangle$ ($n \in \mathbb{N}$).

The unitary representation of the Blaschke group allow us to introduce the concept of the wavelets in the space $H^2(\mathbb{D})$.

The wavelet transform – first formulated in the field of geophysics (P. Goupillaud, 1984), than improved by (Y. Meyer, 1990), (S. G. Mallat, 1989), and (I. Daubechies, 1988) – gained numerous applications in several areas of science. Its success in applying it in signals and systems theory raises - among others - from its localization capabilities both in the time and frequency domain of signals. The conventional wavelets has been constructed to handle signals belonging to the $L^2(\mathbb{R})$ space.

The continuous wavelet transform on a function $f \in L^2(R)$ is formed by taking translation and dilation of a function ψ named the *mother wavelet;* the integral operator with the kernel

$$\psi^{pq}(x) := \frac{\psi((x-q)/p)}{\sqrt{p}} \quad (x \in \mathbb{R}, p \in (0,\infty), q \in \mathbb{R}) \tag{2.11}$$

is called *wavelet transform:*

$$(W_\psi f)(p,q) := \frac{1}{\sqrt{p}} \int_{\mathbb{R}} f(x) \overline{\psi}\left(\frac{x-q}{p}\right) dx = \langle f, \psi^{pq} \rangle \quad (f \in L^2(\mathbb{R})), \tag{2.12}$$

where $\langle \cdot, \cdot \rangle$ means the inner product of the Hilbert-space $L^2(R)$. A reconstruction formula also exists, under some conditions concerning ψ, the function $f \in L^2(R)$ can be reconstructed from its wavelet transform $W_\psi f$. To ensure the existence of the wavelet transform and its inverse the mother-wavelet ψ should satisfy some *admissibility* conditions, expressing it qualitatively: ψ should be localized in respect of its independent variable. See for more detailed introduction on the continuous wavelet transform in (C. K. Chui, 1992) and (I. Daubechies, 1992).

In the continuous wavelet transform (2.12) the mother-wavelet ψ is dilated and translated by the parameters (p,q) as it can be seen in (2.11). In the wavelet transform ψ is submitted to an *affine transform,* which can be formulated as follows:

$$\ell_a(x) := px + q \quad (x \in \mathrm{R}, a := (p,q) \in \mathrm{A} := (0,\infty) \times (-\infty,\infty)).$$

The set of affine maps $(\ell_a, a \in \mathrm{A})$ form a *group* with respect to function composition. The identity element of this group is the map $\ell_e(x) := x$, while the inverse element of ℓ_a is the inverse function of this map, i.e. $\ell_a^{-1} = \ell_{a^{-1}}$, where $a^{-1} := (1/p, -q/p)$ if $a = (p,q) \in \mathrm{A}$. Introducing a group operation on A by $a_1 \circ a_2 := (p_1 p_2, q_1 + p_1 q_2)$ $(a_j := (p_j, q_j) \in \mathrm{A}, j = 1,2)$ we get a group $(\mathrm{A}, \circ)$ isomorphic to the group of affine maps, i.e.

$$\ell_{a_1} \circ \ell_{a_2} = \ell_{a_1 \circ a_2}.$$

The continuous wavelet transform can be defined on the basis of the operators

$$U_a f := \frac{1}{\sqrt{p}} f \circ \ell_{a^{-1}} \quad (a = (p,q) \in \mathrm{A}). \tag{2.13}$$

It is easy to see that the collection $(U_a, a \in \mathrm{A})$ is a *unitary representation* of the group $(\mathrm{A}, \circ)$ on the Hilbert-space $H := L^2(\mathrm{R})$, see for proofs in (C. E. Heil and D. F. Walnut, 1989). The wavelet transform can be expressed by this representation in the form

$$(\mathrm{W}_\psi f)(a) = \langle f, U_a \psi \rangle \quad (f, \psi \in H, a \in \mathrm{A}), \tag{2.14}$$

where $\langle \cdot, \cdot \rangle$ is the scalar product on H. Hence the conventional continuous wavelet transform can be referred as *affine wavelet transform.*

A similar transform can be introduced by using the unitary representation on $H^2(\mathrm{D})$ $U_\mathfrak{b}$ $(\mathfrak{b} \in \mathrm{B})$ defined in (2.8):

$$(W_\phi f)(\mathfrak{b}) = \langle f, U_\mathfrak{b}\varphi \rangle \quad (f, \varphi \in H^2(\mathrm{D}), \mathfrak{b} \in \mathrm{B}), \tag{2.15}$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in the Hardy space $H^2(\mathrm{D})$. This construction forms a transform related to the Blaschke group analogous to the wavelet transform on the affine group. The function ϕ can be considered as the mother wavelet. The operations realized by the one-parameter subgroups $(r,1) \in \mathrm{B}_1$ ($r \in [0,1)$) and $(0,\varepsilon) \in \mathrm{B}_2$ can be associated with the translation and dilation operations respectively. Translation in the unit disc corresponds to a rotation, hence $(0,\varepsilon) \in \mathrm{B}_2$ really corresponds to the translation parameter. The correspondence between $(r,1) \in \mathrm{B}_1$ and the dilation parameter belonging to the affine wavelet is not so evident. A dilation effect can undoubtedly be observed on the Blaschke group elements depending on the distance of their zero *b* from the unit circle, however it cannot be separated from the rotation. Figure 1 presents four instances of the unitary representation belonging to the function $\varphi(z) = z$ as an example with different group elements, i.e. Blaschke function associated with $\mathfrak{b} = (r, \delta)$ equal to $(0.8,1)$, $(0.9,1)$, $(0.95,1)$, $(0.99,1)$, respectively. As it can be observed that as parameter r increases toward the unit, the peak formed by the function gets sharper.

The construction (2.15) – as it is related to the hyperbolic Blaschke group – can be referred as *Blaschke* or *hyperbolic wavelet*.

According to (2.9) the Laguerre system generates a sequence of hyperbolic wavelets on the basis of mother wavelet $\varphi(z) := z^n$. The elements of the wavelet-transform sequence can be computed as the Laguerre-Fourier coefficients belonging to the Laguerre representation of the function to be transformed. There exist efficient algorithms to compute Laguerre-Fourier coefficients belonging to the representation of any function in $H^2(\mathrm{D})$ (A. Soumelidis, F. Schipp, and J. Bokor, 2002).

3. A method of identifying poles

Let $\mathfrak{R}$ denote the set of rational functions with poles falling outside the closed unit disc. It is obvious that functions

$$r_{n,a}(z) := \frac{z^n}{(1-\bar{a}z)^{n+1}} \quad (a \in \mathrm{D}, z \in \bar{\mathrm{D}}, n \in \mathrm{N}) \tag{3.1}$$

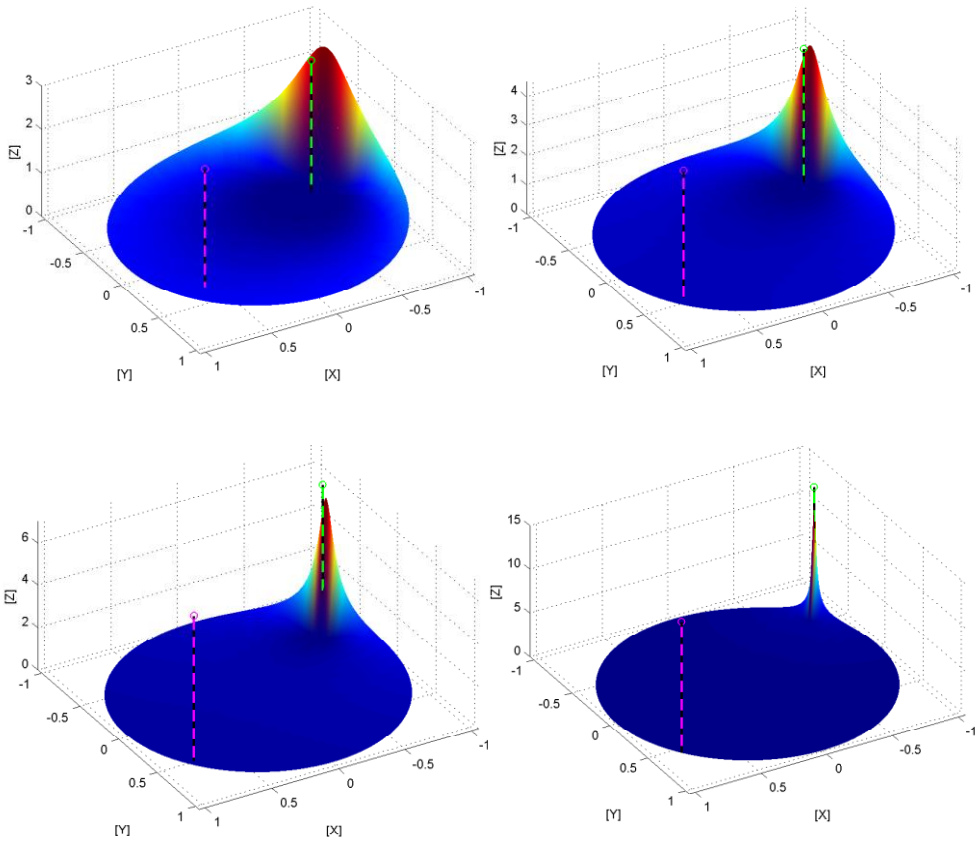


Fig. 1. Unitary group representation on function $\varphi(z) = z$ for different $b \in \mathrm{B}_1$.

belong to $\Re$. $a^* := 1/\bar{a}$ is a pole of multiplicity $(n+1)$ of the function $r_{n,a}$, which is the inverse map of a with respect to the unit circle. a will be referred as the "inverse pole" of the function $r_{n,a}$ in the subsequent part of the paper. It is well known that the functions of form (3.1) generate the function class $\Re$, i.e. any $f \in \Re$ can be expressed in the form

$$f := \sum_{k=1}^{P} \Lambda_{a_k} \qquad \Lambda_{a_k} := \sum_{i=0}^{m_k-1} \lambda_{ki} r_{i,a_k}, \tag{3.2}$$

where $a_k \in \mathrm{D}$ $(k = 1, \ldots, P)$ denote the inverse poles of the function f with their multiplicity m_k.

The following lemma will be used in computing the Laguerre-Fourier coefficients of *f*.

Lemma 3.1 *For every function* $g \in \mathfrak{R}$

$$\langle g, r_{n,a} \rangle = \frac{g^{(n)}(a)}{n!} \quad (n \in \mathrm{N}, a \in \mathrm{D}). \tag{3.3}$$

Proof. By definition

$$\langle g, r_{n,a} \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{g(e^{it}) e^{-\mathrm{int}}}{(1 - ae^{-it})^{n+1}} dt =$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{g(e^{it}) e^{it}}{(e^{it} - a)^{n+1}} dt = \frac{1}{2\pi i} \int_{|\zeta|=1} \frac{g(\zeta)}{(\zeta - a)^{n+1}} d\zeta.$$

Hence by Cauchy's integral formula we get (3.3).

In the case if $m_k = 1$ the associated term will be r_{0,a_k}, and the conjugate of the Laguerre-Fourier coefficients belonging to it are directly given by (3.3) as

$$\langle L_n^b, r_{0,a_k} \rangle = L_n^b(a_k), \tag{3.4}$$

that is equal to coefficient $c_{n,f}$. To indicate that $c_{n,f}$ coefficients belong to the parameter b used in the Laguerre representation, let us denote them as $c_{n,f}(b)$.

Suppose that the system under consideration contains only a single pole of multiplicity 1, in this case the conjugated Laguerre-Fourier coefficients are given as $c_{n,f}(b) = L_n^b(a)$, and the quotients

$$q_n(b) = \frac{c_{n+1,f}(b)}{c_{n,f}(b)} = B_b(a) \quad (n \in \mathrm{N}), \tag{3.5}$$

form a constant sequence and its elements are equal to a Blaschke function applied to a. This fact can be used to identify the position of inverse pole a,

$$a = B_{b^{-1}}\left(\frac{c_{n+1,\,f}(b)}{c_{n,\,f}(b)}\right), \tag{3.6}$$

where $B_{b^{-1}}$ is the inverse of B_b, i.e. a is given by applying a hyperbolic transform corresponding to the inverse group element belonging to b.

This concept can be extended to multiple poles, it will be shown that in the case of multiple poles there exist a region of D where the sequence of the quotients generated by the conjugated Laguerre-Fourier coefficients converge.

Let the inverse poles $a_1, a_2, \ldots, a_P \in \mathrm{D}$ of function f be fixed. Applying the hyperbolic distance as defined by (2.2) let us introduce subsets of D as follows:

$$\begin{aligned} D_i &:= \{b \in \mathrm{D} \,:\, \rho(b, a_i) > \max_{1 \le j \le P, i \ne j} \rho(b, a_j)\}, \\ D &:= \bigcup_{j=1}^{P} D_i \quad (i = 1, 2, \ldots, P). \end{aligned} \tag{3.7}$$

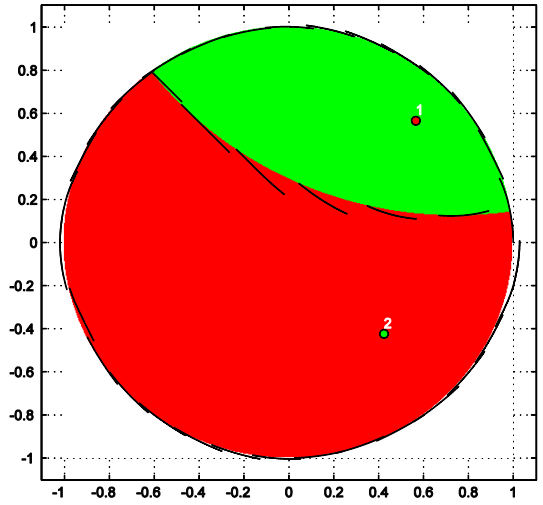


Fig. 2. Domains D_i belonging to a pair of poles.

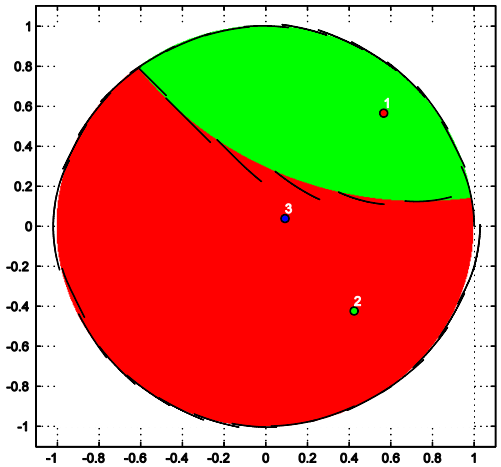


Fig. 3. Domains D_i belonging to three poles, one of them isolated.

Concerning these sets a rather informative interpretation can be given: the set

$$\mathbf{L}_{ij} := \{b \in \mathrm{D} : \rho(a_i, b) = \rho(a_j, b)\} \tag{3.8}$$

can be considered as the hyperbolic perpendicular bisector of the points a_i, a_j that divides D in two hyperbolic half-planes. Let the following notations be introduced:

$$\begin{aligned} D_{ij} &:= \{b \in \mathrm{D} : \rho(a_i, b) > \rho(a_j, b)\} \\ D_{ji} &:= \{b \in \mathrm{D} : \rho(a_i, b) < \rho(a_j, b)\}. \end{aligned} \tag{3.9}$$

The sets D_i can be generated as an intersection of the hyperbolic half-planes, i.e. according to the definitions in (3.7) and (3.9)

$$D_i = \bigcap_{k=1, k\neq i}^{P} D_{ik} \quad (i = 1, 2, \ldots, P). \tag{3.10}$$

As a consequence the sets D_i are hyperbolically convex regions, i.e. any hyperbolic line segment connecting two points belonging to any D_i is located as a whole in the same region.

An example illustrating the placement of regions D_i belonging to 2 poles can be seen in Figure 2. The limit between the two regions is a hyperbolic line satisfying the condition (3.8). In Figure 3 an additional pole has been taken

(pole no. 3), however it cannot generate a nonempty region, i.e. region D_3 cannot be seen.

It will be shown that in any point of D the limit

$$(\mathfrak{Q}f) := \lim_{n\to\infty} \frac{c_{n+1,f}(b)}{c_{n,f}(b)} \quad (f \in \mathfrak{R}) \tag{3.11}$$

does exist, and it can be used to reconstruct the poles of function f. It should be mentioned that operator $\mathfrak{Q}$ defined on domain $\mathfrak{R}$ is nonlinear.

Theorem 3.1 *For any rational function* f *of the form (3.2) in any point* b *of* D *the limit* (3.11) *exists, and*

$$(\mathfrak{Q}f)(b) = B_b(a_i), \quad b \in D_i \quad (i = 1, 2, \ldots, P). \tag{3.12}$$

In the case of poles of multiplicity 1 for the speed of convergence the estimation

$$\left| \frac{c_{n+1,f}(b)}{c_{n,f}(b)} - B_b(a_i) \right| = O(q_i^n) \; (n \in \mathrm{N}, b \in D_i, q_i < 1) \tag{3.13}$$

can be given.

The proof can be found in (F. Schipp and A. Soumelidis, 2011).

According to Theorem 3.1

$$B_b^{-1}((\mathfrak{Q}f)(b)) = a_i \quad (b \in D_i, i = 1, 2, \ldots, P) \tag{3.14}$$

which can be used to reconstruct all the poles with region $D_i \neq 0$ belonging to them.

4. Algorithm of identifying poles

On the basis of Theorem 3.1 and its corollary (3.14) a practically realizable method can be constructed for the reconstruction of system poles by using frequency-domain signal measurements. Two problems has to be solved within the procedure:

1. Estimation of the Laguerre-Fourier coefficients belonging Laguerre parameter b on the basis of frequency domain measurements.
2. Reconstruction the poles as a limit of quotients of consecutive Laguerre-Fourier coefficients.

The Laguerre-Fourier coefficients belonging to function f and parameter b can be computed as trigonometric Fourier coefficients of the unitary representation $U_b f$. An efficient estimation of them is obtained by computing discrete Fourier transform of function $U_b f$ discretized in a specific nonuniform scale obtained as the inverse of the so-called argument function belonging to the Blaschke function B_b – see for details in (A. Soumelidis at all, 2002) and (A. Soumelidis at all, 2011).

In the reconstruction of poles the critical point is finding the limit from finite number of quotients' values. In the most cases the quotient values are stabilized and represent an approximately constant value above a given index – this value can be considered as the limit. If the quotients are not stabilized within the number of the computed coefficients, a new selection of b is advantageous. Finding multiple poles, or all the detectable poles can be done by selecting a sequence of parameters b arranged randomly or in arbitrary order. Poles with empty regions D_i according to (3.7) remain hidden, i.e. no any selection of b exists that leads to a reconstruction of them.

5. A numerical example

The identification of the poles of a simulated function possessing 3 poles, one real pole with inverse pole position $a_1 = 0.8$ and residue $\lambda_1 = 1.5$, as well as a conjugated complex pair of poles in position $a_{2,3} = 0.8 * e^{\pm i\frac{\pi}{4}}$, with associated residues $\lambda_{2,3} = 1$ is presented.

Figure 4 and 6 presents a visualization of the iteration processes for finding specific poles. The poles that belong to a given a_i and the Laguerre parameter used are drawn by color and white circles, respectively. The elements of the quotient sequence generated by (3.5) are transformed by the hyperbolic transform $B_{b^{-1}}$ to locate them at the same region where the poles are located, and these are drawn by white points in the figures. Furthermore, the D_i regions, that are given in (3.7), and belong to several poles a_i are visualized in the figures with colors identical with those of the poles.

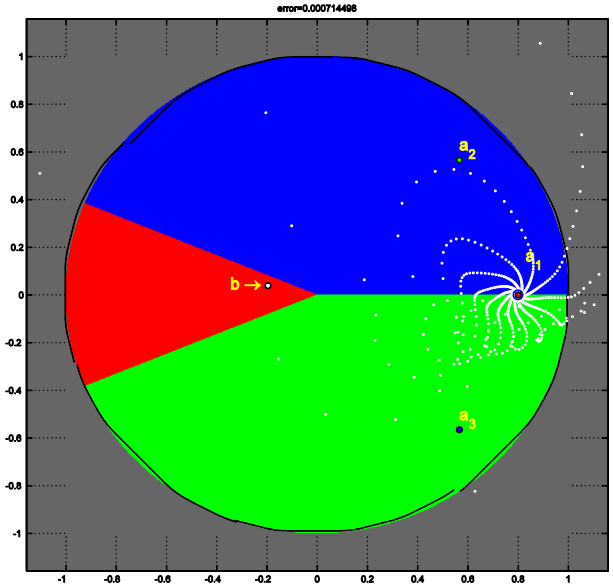


Fig. 4. Example 1: – Finding pole a_1

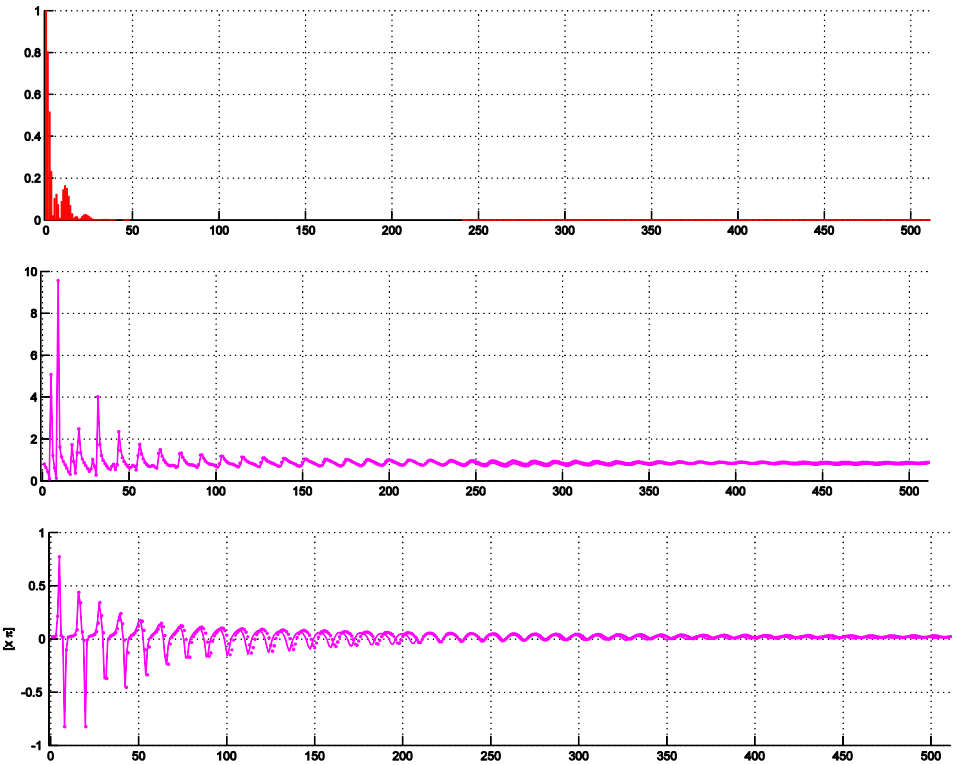


Fig. 5. Example 1 – Abs of L-F coefficients, Abs–Phase of sequence $q_{n.}$

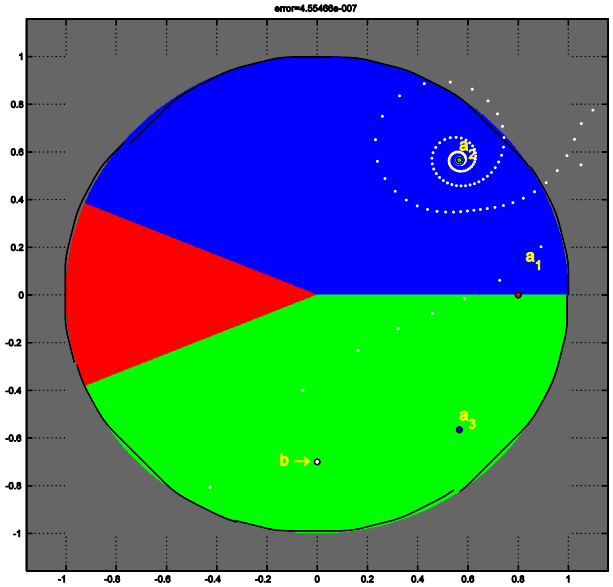


Fig. 6. Example 2 – Finding pole a_2.

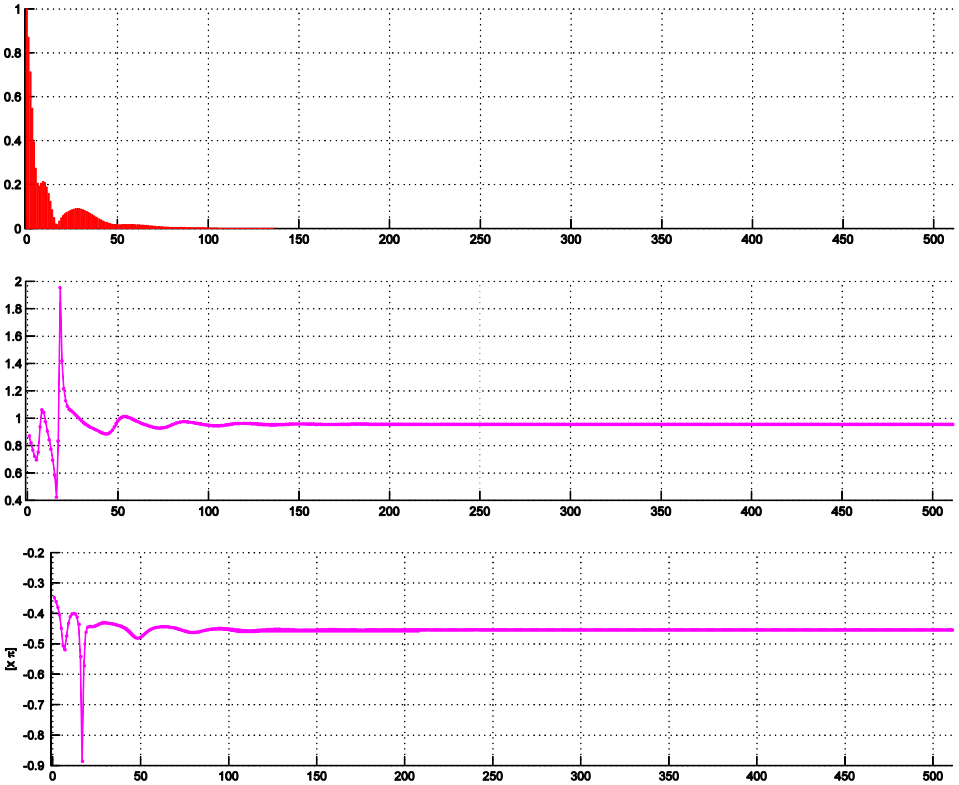


Fig. 7. Example 2 – Abs of L-F coefficients, Abs–Phase of sequence q_n.

The pole to be identified in Example 1 is a_1, hence the parameter b is selected to lie in the region D_2, i.e. $b = 0.2e^{i\frac{15\pi}{16}}$. Similarly, in Example 2 $b = 0.7e^{-i\frac{\pi}{2}}$ has been selected in the region D_2 with the purpose to find a_2.

It can be observed that the transformed sequences of the quotients converge to the specific poles in both examples. The convergence can be checked on the lower two diagrams in Figures 5 and 7 for both examples respectively. The absolute value and the phase of the quotient sequences has been plotted against the indices. The upper diagram in these figures depicts the absolute value of the Laguerre-Fourier coefficients belonging to the specific representation. The reconstruction error – defined as a root-mean-square difference – is rather small in both examples, typically it falls in the magnitude $10^{-5} \dots 10^{-7}$.

The third pole a_3 can be identified analogously to a_2, due to the symmetry of the current pole locations. The identified poles represent the function f with high accuracy, f can be reconstructed by them with root-means-quare error in the magnitude less than 10^{-5}.

Conclusion

A new method of identifying poles of functions belonging to the H^2 space on the unit disc D has been presented in the paper. The method is based on the hyperbolic wavelet constructions generated by the group formed upon the Blaschke functions and the Laguerre representations. Based upon the hyperbolic metric generated by the Blaschke group a method has been set up that can find the poles of the function under consideration, and an algorithm has been outlined that – by using frequency domain measurement data – results in efficient estimates of the poles. The pole identification method presented give an ultimate solution in identifying periodicities in harmonically vibrating mechanical structures on the basis of measurement data.

References

L. Ahlfors. *Conformal invariants*. McGraw-Hill, New York, 1973.

J. Bokor, B. Ninness, T. O. e Silva, P. M. J. Van den Hof, and B. Wahlberg. Selection of poles in GOBF models. In *Modelling and Identification with Orthogonal Basis Functions, Workshop notes of a pre-conference tutorial presented at the 14th IFAC World Congress*, Beijing, China, 1999. Chapter 5.

C. K. Chui. *An introduction to wavelets*. Academic Press, Inc., Boston, San Diego, New York, 1992.

C.K. Chui and G. Chen. *Signal Processing and Systems Theory*. Springer-Verlag, Berlin, Heidelberg, New York, 1992.

I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math*., 41:909-996, 1988.

I. Daubechies. *Ten lectures on wavelets*. CBMS/NSF Series in Applied Mathematics; 61. SIAM, Philadelphia, Pennsylvania, 1992.

T. O. e Silva. Pole selection in GOBF models. In P. S. C. Heuberger, P. M. J. Van den Hof, and B. Wahlberg, editors, *Modelling and Identification with Rational Orthogonal Basis Functions*, chapter 11, pages 297-336. Springer-Verlag, 2005.

P. Goupillaud, A. Grossmann, and J. Morlet. Cycle-octave and related transforms in seismic signal analysis. *Geoexploration*, 25:85-102, 1984.

C. E. Heil and D. F. Walnut. Continuous and discrete wavelet transforms. *SIAM REVIEW*, 31(4):628–666, 1989.

P. S. C. Heuberger, P. M. J. Van den Hof, and B. Wahlberg. *Modelling and Identification with Rational Orthogonal Basis Functions*. Springer-Verlag, New York, 2005.

L. Ljung. *System Identification Theory for the User*. Prentice Hall, Upper Saddle River, New Yersey 07458, 2nd edition, 1999.

S. G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. 11(7):674–693, 1989.

S. L. Marple. *Digital Spectral Analysis with Applications*. Prentice Hall, Englewood Cliffs, 1987.

Y. Meyer. *Ondolettes et Operateurs*. Hermann, New York, 1990.

A.V. Oppenheim, R.W. Schafer, and J.R. Buck. *Discrete-Time Signal Processing*. Prentice Hall, Upper Saddle River, New Yersey 07458, 2nd edition, 1999.

M. Pap and F. Schipp. The voice transform on the Blaschke group I. *Pure Mathematics and Applications*, 17(3-4):387–395, 2006.

F. Schipp and A. Soumelidis. On the Fourier coefficients with respect to the discrete Laguerre system. *Annales Univ. Sci. Budapest, Sect. Comput*., 34:223-233, 2011.

A. Soumelidis, M. Pap, F. Schipp, and J. Bokor. Frequency domain identification of partial fraction models. In *Proceedings of the15th IFAC World Congress*, page on CD, Barcelona, Spain, June 2002.

A. Soumelidis, F. Schipp, and J. Bokor. Frequency domain representation of signals in rational orthogonal bases. In *Proc. of the 10th Mediterranean Conference on Control and Automation*, page on CD, Lisboa, Portugal, 2002. MED'2002.

A. Soumelidis, F. Schipp, and J. Bokor. On hyperbolic wavelets. In S. Bittandi, Cenedese, and S. Zampieri, editors, *Preprints of the 18th IFAC World Congress*, pages 2309–2314, Milano, Italy, August 28 – September 2 2011.

SOME RECENT RESULTS ON STABILITY PROPERTIES OF SECOND ORDER DIFFERENTIAL EQUATIONS WITH STEP FUNCTION COEFFICIENTS

László HATVANI[1], László SZÉKELY[2]

[1]Bolyai Institute, University of Szeged, Hungary
[2]Department of Mathematics, Institute for Mathematics and Informatics

Abstract

In this paper we examine two problems related to differential equations with step function coefficients that have been discussed in our recent papers and in the unpublished thesis of the second author. First, we consider second order linear differential equations, where both elasticity and damping coefficients are step functions. For such equations, we give sufficient condition for the existence of a small solution, i.e., the existence of a non-trivial solution tending to 0 with respect to x as t tends to infinity. For the proof of the theorem, as a tool, we recall our earlier result guaranteeing the existence of a small solution of linear systems of difference equations of arbitrarily finite dimension. Here, we also generalize this result to some nonlinear systems of difference equations. As a second problem, we consider the Armellini-Tonelli-Sansone theorem for second order linear differential equations with varying elasticity coefficient. This theorem gives a sufficient condition guaranteeing that all non-trivial solutions of such equations are small. We present the extension of this theorem to the so-called half-linear differential equations in the case when the coefficient is a step function. Half-linear differential equations have many important applications, e.g., describing the mechanics of certain non-Newtonian fluids. For the extension of the Armellini-Tonelli-Sansone theorem to the half-linear case, we needed to prove a new theorem on the asymptotic stability of two dimensional systems of linear difference equations. The proof is based on a geometric method which also applies for the nonlinear case.

Keywords

small solution, asymptotic stability, Armellini-Tonelli-Sansone theorem, linear ODE, damping, step-function coefficients, nonlinear difference equation, half-linear differential equation

1. Introduction

Consider the second order differential equation

$$x'' + a(t)x = 0 \tag{1}$$

describing the motion of a linear oscillator with varying elasticity coefficient. The Pólya-Sonin theorem (see e.g. (Szegő, 1959)) says, that if $a:[0,\infty)\to[0,\infty)$ is a monotone non-decreasing function, then all non-trivial solutions of equation (1) are oscillatory, the local maxima of $|x|$, i.e. the size of the amplitudes is non-increasing, and the distance between consecutive maxima of $|x|$, that is the distances between consecutive local extrema of x are non-increasing.

Definition 1 A non-trivial solution x_0 of equation (1) is called *small* if

$$\lim_{t\to\infty} x_0(t) = 0.$$

Milloux (Milloux, 1934), Prodi (Prodi, 1950) and Trevisan (Trevisan, 1954) proved that if $a:[0,\infty)\to[0,\infty)$ is differentiable and non-decreasing, then equation (1) has at least one small solution if and only if $\lim_{t\to\infty} a(t) = \infty$ holds. Milloux also constructed an example with a step function coefficient a, where not all non-trivial solutions of equation (1) were small. Hartman (Hartman, 1948) investigated the linear system of differential equations

$$x' = A(t)x \tag{2}$$

where x is an m dimensional vector and A is an $m\times m$ matrix having real continuous entries with domain $[0,\infty)$. He proved the following:

Theorem 2 *Suppose that for all solutions x of equation (2)* $\lim_{t\to\infty}\|x(t)\| < \infty$ *holds. Then equation (2) has at least one small solution if and only if*

$$\int_0^t trA(s)ds \to -\infty \quad (t\to\infty). \tag{3}$$

According to Liouville's theorem, condition (3) means that the phase volume of system (2) tends to zero as t tends to infinity. Based on this result, Hartman (Hartman, 1948) extended the theorem of Milloux, Prodi and Trevisan to systems of equations, furthermore, he proved that instead of differentiability, it is sufficient to assume the continuity of a.

The Armellini-Tonelli-Sansone (see e.g. Macki, 1982) theorem was the first to give a sufficient condition guaranteeing that *all* non-trivial solutions of equation (1) are small with the following concept. A nondecreasing function $f:[0,\infty)\to(0,\infty)$ with $\lim_{t\to\infty} f(t) = \infty$ is called to grow *intermittently* if for

every $\varepsilon > 0$ there is a sequence $\{(a_i, b_i)\}_{i=0}^{\infty}$ of disjoint intervals such that $a_i \to \infty$ as $i \to \infty$, and

$$\limsup_{i\to\infty} \sum_{k=1}^{i} \frac{b_k - a_k}{b_i} \le \varepsilon, \sum_{i=1}^{\infty} (f(a_{i+1}) - f(b_i)) < \infty$$

are satisfied. Roughly speaking, this condition means that the growth of *f* cannot be located to a set of small measure. If such a sequence does not exist, then *f* is called to grow *regularly*.

Theorem 3 *If a is continuously differentiable and it grows to infinity regularly as* $t \to \infty$*, then all non-trivial solutions of equation (1) are small.*

It is important to note that this stability property is weaker than the asymptotic stability of the trivial solution of equation (1).
The simplest case of intermittent growth is a monotonously increasing step function. Such equations have an important role for example in the field of control theory thanks to the so-called Bang-Bang principle. Differential equations with step function coefficients can be rewritten as systems of difference equations, thus the proof of theorems on such equations can be deduced to the proof of statements on difference equations.

2. On small solutions of second order linear differential equations with step function coefficients

We consider the equation

$$x'' + c(t)x' + a^2(t)x = 0 \tag{4}$$

describing the motion of an oscillator where both elasticity coefficient *a* and damping coefficient *c* are step functions. Namely, $\{t_n\}_{n=1}^{\infty}$, $\{a_n\}_{n=1}^{\infty}$ and $\{c_n\}_{n=1}^{\infty}$ are real sequences with the following properties:

$$0 = t_0 < t_1 < \ldots < t_{n-1} < t_n < \ldots; \quad \lim_{n\to\infty} t_n = \infty,$$

$$a_n > 0, \quad c_n \ge 0 \quad (n = 1,2,\ldots),$$

furthermore, $a(t) = a_n$ and $c(t) = c_n$ on the interval $[t_{n-1}, t_n)$. In the case when $\lim_{n\to\infty} a_n = \infty$ and damping doesn't act, i.e. $c_n = 0$ $(n = 1,2,\ldots)$, the first

author (Hatvani, 1998) gave the following sufficient condition on the existence of at least one small solution of equation (4).

Theorem 4 (Hatvani, 1998) *Assume that the above conditions on sequences* $\{t_n\}_{n=1}^{\infty}$, $\{a_n\}_{n=1}^{\infty}$ *are satisfied and suppose that* $\lim_{n\to\infty} a_n = \infty$ *and* $c_n = 0$ $(n = 1,2,\ldots)$ *holds. If*

$$\sum_{n=1}^{\infty} \max\left\{\frac{a_n}{a_{n+1}} - 1; 0\right\} < \infty$$

then equation (4) has at least one small solution.

It is natural to guess that damping helps weaken this condition and even the condition $\lim_{n\to\infty} a_n = \infty$. In fact, as the main theorem of this section we prove the following, which is, in addition, an improvement of an earlier result of the authors (Hatvani, Székely, 2006) as well. Actually, the following theorem corrects conditions (i) and (ii) in Theorem 3.2 of (Hatvani, Székely, 2006), which were not accurate.

Theorem 5 *Assume that the above conditions on sequences* $\{a_n\}_{n=1}^{\infty}$, $\{c_n\}_{n=1}^{\infty}$ *and* $\{t_n\}_{n=1}^{\infty}$, *are satisfied, and let us introduce the notation*

$$\gamma_n := \frac{c_n}{2a_n + c_n}[(2a_n - c_n)(t_n - t_{n-1}) - 2].$$

Suppose, in addition, that

(i) $$a_n > c_n / 2 \quad (n = 1,2,\ldots),$$

(ii) $$\sum_{k=1}^{\infty}\left(-\gamma_k + \ln\frac{a_k}{a_{k+1}}\right) = -\infty,$$

(iii) *there is a number K such that for every* $p, q \in N$, $(1 \le p \le q)$

$$\sum_{k=p}^{q}\left(-\frac{\gamma_k}{2} + \ln\max\left\{\frac{a_k}{a_{k+1}}; 1\right\}\right) < K$$

holds. Then equation (4) has at least one small solution.

Equation (4) is equivalent with a two dimensional system of difference equations, therefore for the proof of our theorem we need a sufficient condition guaranteeing the existence of a small solution of such system. To this end, we recall a theorem of the authors which gives a necessary and sufficient condition for the existence of such solutions of arbitrarily finite dimensional systems.
Consider the following nonautonomous system of difference equations

$$x_{n+1} = M_n x_n, \quad n = 0,1,2,\ldots \tag{5}$$

where $m \in N$, $x_n \in R^m$ is a column vector and $M_n \in R^{m\times m}$ is an $m \times m$ matrix having real entries. For vectors $x = (x^1,\ldots x^m)^{\mathrm{T}}$, $y = (y^1,\ldots y^m)^{\mathrm{T}} \in R^m$ and matrix M denote by $x \cdot y$, $\| x \|$, $\| M \|$, respectively, the scalar product, the Euclidian norm and the matrix norm, i.e.

$$x \cdot y := \sum_{i=1}^{m} x^i y^i, \qquad \| x \| := \sqrt{x \cdot x}, \qquad \| M \| := \max_{\|x\|=1} \| Mx \|.$$

A non-trivial solution $\{x_n\}_{n=0}^{\infty}$ of equation (5) is called *small* if $\lim_{n\to\infty} x_n = 0$.
Our following theorem is an extension and even a sharpened version of Hartman's theorem on linear systems of differential equations to linear systems of difference equations. Namely, we showed that the condition that the phase volume of system (5) tends to zero as *n* tends to infinity, i.e. $\prod_{n=0}^{\infty} |\det M_n| = 0$ is necessary and sufficient to have at least one small solution, if we require in some sense only the boundedness of the solutions.

Theorem 6 (Hatvani, Székely, 2006) *Suppose that there is a* $K \in R$ *such that for every* $p, q \in N$, $(0 \le p \le q)$ *we have*

$$\left\| \prod_{n=p}^{q} M_n \right\| < K.$$

Then there exists at least one small solution of (5) if and only if

$$\prod_{n=0}^{\infty} |\det M_n| = 0.$$

Now, we are ready to prove Theorem 5.

Proof. Introducing the variable $y = x'/a_n$ $(n = 1,2,\ldots)$, equation (4) can be rewritten into the first order system

$$x' = a_n y, \quad y' = -a_n x - c_n y \quad (t_{n-1} \le t < t_n, \ n = 1,2,\ldots). \tag{6}$$

Since the solutions of equation (4) are continuously differentiable on the interval $[0,\infty)$, therefore the function $x'=a_n y$ has to be continuous on the same interval. Accordingly, if $t\mapsto(x(t),y(t))$ is a solution of equation (6) on the interval $[0,\infty)$, then we require that the function $t\mapsto y(t)$ is continuous from the right for all $t\geq 0$ and satisfies $a_n y(t_n-0)=a_{n+1}y(t_n)$ for $n=1,2,\ldots,$ where $y(t_n-0)$ denotes the left-hand side limit of y at t_n. This means that equation (4) for x is equivalent with the following system of first order differential equations with impulses

$$\begin{aligned} x'&=a_n y, \qquad y'=-a_n x-c_n y \qquad (t_{n-1}\leq t<t_n)\\ y(t_n)&=\frac{a_n}{a_{n+1}}y(t_n-0), \qquad n=1,2,\ldots \end{aligned} \tag{7}$$

Introducing the polar coordinates (r,φ) by the equations $x=r\cos\varphi$, $y=r\sin\varphi$, we can rewrite system (6) into the form

$$\begin{aligned} r'&=-c_n r\sin^2\varphi,\\ \varphi'&=-a_n-\frac{c_n}{2}\sin 2\varphi \quad (t_{n-1}\leq t<t_n, n=1,2,\ldots). \end{aligned} \tag{8}$$

Note, that condition (i) implies that $\varphi'(t)\leq 0$ holds for all $t\in[t_{n-1},t_n)$ $(n=1,2,\ldots)$. Using the identity $\sin^2\varphi=(1-\cos 2\varphi)/2$ and the Newton-Leibniz theorem, we get the estimate

$$\ln\frac{r(t_n-0)}{r(t_{n-1})}=-c_n\int_{t_{n-1}}^{t_n}\sin^2\varphi(t)\,dt=-c_n\int_{t_{n-1}}^{t_n}\frac{\sin^2\varphi(t)\varphi'(t)}{\varphi'(t)}\,dt$$

$$=-c_n\int_{t_{n-1}}^{t_n}\frac{\sin^2\varphi(t)\varphi'(t)}{-a_n-\frac{c_n}{2}\sin 2\varphi(t)}\,dt=-c_n\int_{\varphi(t_n-0)}^{\varphi(t_{n-1})}\frac{\sin^2 u}{a_n+\frac{c_n}{2}\sin 2u}\,du \tag{9}$$

$$\leq-\frac{c_n}{a_n+\frac{c_n}{2}}\int_{\varphi(t_n-0)}^{\varphi(t_{n-1})}\sin^2 u\,du=-\frac{c_n}{a_n+\frac{c_n}{2}}\left\{\left[\frac{1}{2}u-\frac{\sin 2u}{4}\right]_{\varphi(t_n-0)}^{\varphi(t_{n-1})}\right\}$$

$$= \frac{c_n}{2a_n + c_n}\left[(\varphi(t_n - 0) - \varphi(t_{n-1})) - \frac{1}{2}(\sin 2\varphi(t_n - 0) - \sin 2\varphi(t_{n-1})) \right].$$

From the second equation of (8) we obtain the following estimation for φ

$$\varphi' = -a_n - \frac{c_n}{2}\sin 2\varphi \le -a_n + \frac{c_n}{2} \quad (t_{n-1} \le t < t_n,\ n = 1,2,\ldots).$$

By integration we get the inequality

$$\varphi(t_{n-1}) - \varphi(t_n - 0) = -\int_{t_{n-1}}^{t_n} \varphi'(t)\ dt \ge \left(a_n - \frac{c_n}{2} \right)(t_n - t_{n-1}).$$

Now, we can continue estimate (9):

$$\ln \frac{r(t_n - 0)}{r(t_{n-1})} \le -\frac{c_n}{2(2a_n + c_n)}\left[(2a_n - c_n)(t_n - t_{n-1}) - 2 \right] = -\frac{\gamma_n}{2}. \qquad (10)$$

Let $M_{n-1}^{(1)}$ denote the fundamental matrix for equation (8), i.e.

$$\begin{pmatrix} x(t_n - 0) \\ y(t_n - 0) \end{pmatrix} = M_{n-1}^{(1)} \begin{pmatrix} x(t_{n-1}) \\ y(t_{n-1}) \end{pmatrix}.$$

The vector $(x(t_n), y(t_n))^{\mathrm{T}}$ is defined by equation (7):

$$\begin{pmatrix} x(t_n) \\ y(t_n) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{a_n}{a_{n+1}} \end{pmatrix} \begin{pmatrix} x(t_n - 0) \\ y(t_n - 0) \end{pmatrix} = M_{n-1} \begin{pmatrix} x(t_{n-1}) \\ y(t_{n-1}) \end{pmatrix},$$

where

$$M_{n-1} := \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{a_n}{a_{n+1}} \end{pmatrix} M_{n-1}^{(1)}.$$

Obviously, any stability property of equation (7) is equivalent with the same stability property of the difference equations

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = M_{n-1} \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} (n = 1, 2, \ldots). \tag{11}$$

Using estimation (10) we can estimate the norms of matrices $M_{n-1}^{(1)}$ $(n = 1,2,\ldots)$:

$$\left\| M_{n-1}^{(1)} \right\| = \sup_{0<r(t_{n-1})} \frac{r(t_n - 0)}{r(t_{n-1})} \le \exp\left[-\frac{\gamma_n}{2} \right], \tag{12}$$

where the supremum is taken in the case of each solutions where *r* is positive at time t_{n-1}. Thus, we have

$$\left\| M_{n-1} \right\| = \left\| M_{n-1}^{(1)} \right\| \max\left\{ \frac{a_n}{a_{n+1}}; 1 \right\}$$

$$\le \exp\left[-\frac{\gamma_n}{2} \right] \max\left\{ \frac{a_n}{a_{n+1}}; 1 \right\} = \exp\left[-\frac{\gamma_n}{2} + \ln \max\left\{ \frac{a_n}{a_{n+1}}; 1 \right\} \right].$$

It can easily be seen that due to condition (iii)

$$\prod_{k=p}^{q} \left\| M_{k-1} \right\| \le \prod_{k=p}^{q} \exp\left[-\frac{\gamma_k}{2} + \ln \max\left\{ \frac{a_k}{a_{k+1}}; 1 \right\} \right]$$

$$= \exp\left[\sum_{k=p}^{q} \left(-\frac{\gamma_k}{2} + \ln \max\left\{ \frac{a_k}{a_{k+1}}; 1 \right\} \right) \right] < K$$

holds for arbitrary *p, q* $(1 \le p \le q)$. According to Theorem 6 it is enough to show that for system (11) $\prod_{n=0}^{\infty} |\det M_n| = 0$ is satisfied as well. To see this, we recall the inequality

$$\left| \det M \right| \le \left\| M \right\|^m,$$

which holds for any $m\times m$ $(m\in N)$ matrix having real entries. Based on this inequality and estimate (12) we have

$$\prod_{k=1}^{n}\left|\det M_{k-1}\right|=\prod_{k=1}^{n}\frac{a_k}{a_{k+1}}\left|\det M_{k-1}^{(1)}\right|\le\prod_{k=1}^{n}\left\{\frac{a_k}{a_{k+1}}\left(\sup_{0<r(t_{n-1})}\frac{r(t_n-0)}{r(t_{n-1})}\right)^2\right\}$$

$$\le\prod_{k=1}^{n}\exp\left[-\gamma_k+\ln\frac{a_k}{a_{k+1}}\right]=\exp\left[\sum_{k=1}^{n}\left(-\gamma_k+\ln\frac{a_k}{a_{k+1}}\right)\right].$$

From condition (ii) it follows that $\prod_{n=0}^{\infty}|\det M_n|=0$, thus with an application of Theorem 6 the statement is proved.

Remark 7 *Conditions (ii) and (iii) in Theorem 5 are independent of each other.*

By the choice $c_n=0$, $t_n=n$ and $a_n=a>0$ $(n=1,2,\ldots)$ it can be seen that (iii) does not imply (ii), because in both expressions all elements are 0. To show the opposite direction let $c_n=0$, $t_n=n$ $(n=1,2,\ldots)$ and define a_n in the following way:

$$a_1:=2,\qquad a_{2k}:=2k,\qquad a_{2k+1}=2(2k+1),\qquad k=1,2,\ldots$$

Since $\lim_{n\to\infty}a_n=\infty,$ therefore (ii) holds, but according to the estimation below (iii) doesn't:

$$\sum_{k=p}^{q}\ln\max\left\{\frac{a_k}{a_{k+1}};1\right\}=\sum_{k=\left\lceil\frac{p-1}{2}\right\rceil}^{\left\lfloor\frac{q-1}{2}\right\rfloor}\ln\frac{2(2k+1)}{2(k+1)}$$

$$\ge\sum_{k=\left\lceil\frac{p-1}{2}\right\rceil}^{\left\lfloor\frac{q-1}{2}\right\rfloor}\ln\frac{4}{3}=\ln\frac{4}{3}\left(\left\lfloor\frac{q-1}{2}\right\rfloor-\left\lceil\frac{p-1}{2}\right\rceil+1\right),$$

which cannot be uniformly bounded for all $(1\le p<q)$.

Remark 8 *With the aid of the inequality* $x-1\geq \ln x$ *one can easily see that in the case* $c_n = 0$ $(n=1,2,\ldots)$ *Theorem 5 follows from the theorem of Hatvani (Theorem 4).*

3. On small solutions of nonlinear difference equations

In this section we examine the possible extensions of Theorem 6 to nonlinear systems of difference equations in a similar way to the discussion in (Székely, 2011). Consider the difference equation

$$x_{n+1} = f(n, x_n), \quad n = 0,1,2,\ldots, \tag{13}$$

where $m \in N$, $x_n \in R^m$ is a column vector, and functions $f(n,\cdot)$ have the following properties for all $n \in N_0$:

$$f(n,\cdot): D_n \subset R^m \to R^m, \text{ ran } f(n,\cdot) \subset D_{n+1},$$

$$f(n,0) = 0, \ f(n,\cdot) \in C^1(D_n),$$

where D_n is a convex domain $(n = 0,1,\ldots)$. Define

$$F(q, p;\cdot) := f(q,\cdot) \circ \ldots \circ f(p,\cdot) \qquad (0 \leq p \leq q, p, q \in N_0),$$

furthermore, let $F^j(q, p;\cdot): D_p \to R$ $(j = 1,\ldots,m)$ be the *j*th component of function $F(q, p;\cdot)$, i.e.

$$F(q, p; x) = \begin{pmatrix} F^1(q, p; x) \\ \vdots \\ F^m(q, p; x) \end{pmatrix}.$$

Since $F(n,0;\cdot)$ is the flow of equation (13), then for a solution $\{x_n\}_{n=0}^{\infty}$ of (13) $F(n,0;x_0) = x_n$ holds. For a function $g: R^m \to R$ denote by grad $g(x)$ the gradient vector of g, i.e., grad $g(x) = \left(\partial g(x)/\partial x_1, \ldots, \partial g(x)/\partial x_m\right)$. Let $H_0 \subset D_0$ be the closure of a bounded and connected open set. Then the *n*th image of H_0

under the flow of equation (13) is $H_n = F(n,0;H_0)$, and the phase volume of H_n is

$$\mu(H_n) = \int_{H_0} \left|\det F'(n,0;x)\right| dx,$$

where μ is the Lebesgue measure. Furthermore, let $\overline{H}$, ∂H and $\operatorname{int} H = \overline{H} \setminus \partial H$ be the closure, the boundary and the interior of a set $H \in R^m$, respectively.

With the aid of a Lyapunov function, Karsai, Graef and Li (Graef, Karsai, Li, 2000) gave a sufficient condition for such equations to have at least one small solution. Currently, with the application of our topological method of proof used in the proof of Theorem 6 (Theorem 2.2 in (Hatvani, Székely, 2006)), we could only conclude such result which is a consequence of their theorem. Nevertheless, since the conditions in our theorem are based only on the right hand side of equation (13), we present this result as well (see also in (Székely, 2011)).

Theorem 9 *Suppose that there exists a closed ball* H_0 *around the origin and a number* $K > 0$, *such that for all* $p,q \in N_0$ $(0 \le p \le q)$, $j = 1,\dots,m$ *and* $x \in H_0$

$$\| \operatorname{grad} F^j(q,p;x) \| \le K \tag{14}$$

holds, furthermore

$$\lim_{n\to\infty} \int_{H_0} \left|\det F'(n,0;x)\right| dx = 0\,. \tag{15}$$

Then equation (13) has at least one small solution. In particular, if

$$\lim_{n\to\infty} \det F'(n,0;x) = 0 \qquad (x \in H_0)$$

then the assertion holds.

Proof. Denote by s_n any point of H_n closest to the origin. It is easy to see that condition (15) implies $\lim_{n\to\infty} s_n = 0$. Let r_n be one of the points $x \in \partial H_0$ for which $F(n,0;x) = s_n$ holds. Since H_0 is compact, there is a convergent subsequence $\{r_{n_l}\}_{l=0}^{\infty}$ of the sequence $\{r_n\}_{n=0}^{\infty}$, such that $\lim_{l\to\infty} r_{n_l} =: r \in H_0$.

We will show that the solution with initial point r is a small solution, i.e., $\lim_{n\to\infty} F^j(n,0;r)=0$ $(j=1,\ldots,m)$. Let us fix an index j $(1\le j\le m)$. Since $F(n_l,0;r_{n_l})\to 0$ and $\| r-r_{n_l}\|\to 0$ as $l\to\infty$, for every $\varepsilon>0$ there is an $\tilde{l}\in N$ such that if $l>\tilde{l}$, then $|F^j(n,0;r)|<\varepsilon/(2K)$ and $\| r-r_{n_l}\|<\varepsilon/(2K)$ are both satisfied. Now, choose $l>\tilde{l}$. Then, since the origin is a fixed point, for $n_l>n_{\tilde{l}}$ we have

$$|F^j(n,0;r)|\le |F^j(n,0;r)-F^j(n,0;r_{n_l})|+|F^j(n,0;r_{n_l})|$$

$$=|F^j(n,0;r)-F^j(n,0;r_{n_l})|$$

$$+|F^j(n,n_l+1;F(n_l,0;r_{n_l}))-F^j(n,n_l+1;0)|.$$

Applying Lagrange's theorem for vector-scalar functions to both expressions, we get

$$|F^j(n,0;r)|\le|\operatorname{grad}F^j(n,0;\xi_j)\cdot(r-r_{n_l})|+|\operatorname{grad}F^j(n,n_l+1;\eta_j)\cdot F(n_l,0;r_{n_l})|$$

$$\le\|\operatorname{grad}F^j(n,0;\xi_j)\|\cdot\|(r-r_{n_l})\|+\|\operatorname{grad}F^j(n,n_l+1;\eta_j)\|\cdot\|F(n_l,0;r_{n_l})\|$$

$$<K\frac{\varepsilon}{2K}+K\frac{\varepsilon}{2K}=\varepsilon,$$

where $\xi_j=r+(1-\lambda_j)r_{n_l}$ and $\eta_j=\nu_j F(n_l,0;r_{n_l})$ for some $\lambda_j,\nu_j\in[0,1]$. Since ε and j were arbitrarily chosen the theorem is proved.

■

Remark 10 *From the proof one can easily see that condition (14) implies that all solutions of equation (13) are bounded.*

Remark 11 *In case of the linear equation (5)* $F(q,p;x)=\left(\prod_{n=p}^{q}M_n\right)x$, *therefore the sufficiency of Theorem 6 follows from Theorem 9.*

4. On stability of second order half-linear differential equations with step function coefficients

In the final section we consider the half-linear second order differential equation

$$x''|x'|^{n-1}+q(t)|x|^{n-1}x=0, \qquad n\in R^{+}, \tag{16}$$

which is an important generalization of the second order differential equation (1) and was introduced by Imre Bihari (Bihari, 1957) and Árpád Elbert (Elbert, 1981). They called it half-linear because its solution set is homogeneous, but it is not additive. By the appropriate choice of coefficients $(n=p-1)$, equation (16) can be transformed into the well-known equation

$$(r(t)\Phi(x'))'+c(t)\Phi(x)=0, \qquad \Phi(x)=|x|^{p-1}\operatorname{sgn}x, \quad p>1, \tag{17}$$

which can be considered as a generalization of the so-called Sturm-Liouville differential equation

$$(r(t)x')'+c(t)x=0.$$

A motivation of the qualitative analysis of half-linear differential equations is for example that many physical phenomena can be described by such partial differential equations which include the so-called *p*-Laplacian operator. In the one dimensional case, under certain conditions, these equations can be transformed to the form of (16), or equivalently of (17). The *p*-Laplacian operator is the following:

$$\Delta_p u(x) := \operatorname{div}(\|\nabla u(x)\|^{p-2}\nabla u(x)),$$

where $x=(x_1,\ldots,x_m)\in R^m$, $p>1$, ∇ is the gradient, furthermore, $\operatorname{div}=\sum_{k=1}^{m}\partial/\partial x_k$ is the divergence operator. A specific example for the application of half-linear differential equations is the mechanics of certain non-Newtonian fluids, namely, for which the connection between shear stress (τ_{xy}) and shear velocity (du/dy) is

$$\tau_{xy}=\eta_0\left|\frac{du}{dy}\right|^{n-1}\frac{du}{dy},$$

where η_0 and n are constants characterizing the fluid (in case of $n=1$ the fluid is Newtonian and η_0 is the viscosity). Some examples for such fluids are milk, paints, melt of certain polymers, and slurry with mineral dust. Besides applications (see e.g. (Bognár, 2010), (Bognár, 2011)), the qualitative theory of half-linear differential equations has attracted much attention (see, e.g., the monographs (Agarwal, 2002), (Došlý, 2004), (Došlý, Řehák, 2005) and the references therein).

For equation (16) Bihari (Bihari, 1984) proved an Armellini-Tonelli-Sanone-type theorem, namely, he proved that the trivial solution of (16) is asymptotically stable with respect to x if coefficient q is continuously differentiable and tends regularly to infinity as $t \to \infty$. Such result for this equation with irregularly (or intermittently) growing coefficients was unknown until the appearance of our paper (Hatvani, Székely, 2011). In this section we give a sufficient condition for the asymptotic stability of the trivial solution with respect to x in the case when coefficient q is the most typically intermittently growing, that is when q is a step function. In the proof of our theorem we could successfully replace the method used for the linear case ($n=1$ in (16)) to a geometric technique which does not require linearity. What is more, this new method of proof allows us to sharpen the known results for the linear case. Therefore, our results not just include, but even sharpen the Armellini-Tonelli-Sanone-type theorems of Elbert ((Elbert, 1996/97), (Elbert, 2001)) for linear differential equations with step function coefficients, thus we first introduce this method to linear systems of difference equations.

To this end, we investigate the asymptotic stability of the trivial solution of the linear system of difference equations (5) in the case when it is two dimensional. It is well-known that if $\prod_{n=0}^{\infty} \| M_n \| = 0$, then all solutions of equation (5) tend to zero as $n \to \infty$ (see e.g., (Agarwal, 1992)). Elbert (Elbert, 1997) gave a sufficient condition for the asymptotic stability under the assumptions: (i) $\prod_{n=0}^{\infty} \max\{\| M_n \|;1\} < \infty$, (ii) $0 < \prod_{n=0}^{\infty} \| M_n \|$, (iii) $\prod_{n=0}^{\infty} \max\{|\det M_n|;1\} < \infty$. His proof was based on estimation of the norm of some special matrices and a "tricky" decomposition of matrices M_n. To investigate equation (5), we define a difference equation on the plane which has the same stability properties as equation (5). The construction of this equation is based on the polar factorization theorem (see e.g., (Jacobson, 1953)). Let

$$x_{n+1} = \| M_n \| \begin{pmatrix} 1 & 0 \\ 0 & d_n \end{pmatrix} \begin{pmatrix} \cos\omega_n & -\sin\omega_n \\ \sin\omega_n & \cos\omega_n \end{pmatrix} x_n, \quad 0 \le d_n \le 1, \; n = 0,1,\ldots, \qquad (18)$$

where d_n and ω_n $(n=0,1,\ldots)$ can be calculated from matrices $M_0,\ldots M_n$. Using a geometric method of proof, in (Hatvani, Székely, 2011) we showed that conditions (i)-(iii) of Elbert's theorem can be weakened.

Theorem 12 (Hatvani, Székely, 2011) *Suppose that*

$$\limsup_{n\to\infty}\prod_{k=0}^{n}\| M_k \| < \infty .$$

If

$$\sum_{n=0}^{\infty}\min\{1-d_n,1-d_{n+1}\}\sin^2\omega_{n+1}=\infty, \tag{19}$$

then the zero solution of difference equation (18) is asymptotically stable.

From a geometric point of view, the dynamics of (18) consists of consecutive rotations and contractions along the *y* axis. Due to the contraction each point's distance from the origin decreases except the points which are transformed onto the *x* axis by the rotation. Condition (19) guarantees that there is no point which stays "close" to the *x* axis after the rotations, that is the angles of rotations aren't close to $k\pi$ $(k\in Z)$, and, in addition, that the contractions aren't too small. Based on Theorem 12, as the main result of our paper (Hatvani, Székely, 2011) we could successfully extend the Armellini-Tonelli-Sansone theorem to second order half-linear differential equations with step function coefficients to the case $n>1$.

Theorem 13 (Hatvani, Székely, 2011) *Let* $n>1$ *and*

$$0=t_0<t_1<\ldots<t_k<t_{k+1}<\ldots, \quad \lim_{k\to\infty}t_k=\infty,$$

$$0<q_0\le q_1\le\ldots\le q_k\le q_{k+1}\le\ldots, \lim_{k\to\infty}q_k=\infty.$$

Then all non-trivial solutions of equation

$$x''|x'|^{n-1}+q_k|x|^{n-1}x=0 \qquad (t_k\le t<t_{k+1},\ k=0,1,\ldots) \tag{20}$$

are small, if

$$\sum_{k=0}^{\infty}\min\left\{1-\frac{q_k}{q_{k+1}},1-\frac{q_{k+1}}{q_{k+2}}\right\}\left|S\left(q_{k+1}^{\frac{1}{n+1}}(t_{k+2}-t_{k+1})\right)\right|^{n+1}=\infty.$$

The function S appearing in the theorem is the so-called generalized sine function, that is, the solution of the initial value problem

$$\begin{cases} S''|S'|^{n-1} + |S|^{n-1} S = 0, \\ S(0) = 0, \quad S'(0) = 1. \end{cases}$$

Note, that S satisfies the identity $|S(\Phi)|^{n+1} + |S'(\Phi)|^{n+1} \equiv 1$. The proof is similar to the one of Theorem 12, but due to the appearance of the generalized trigonometric functions we have to modify our estimations. The main difficulty is that exact addition formulae for these functions are unknown. Unfortunately, our method of proof cannot be applied directly to the case $0 < n < 1$. Although, our conjecture is that the theorem remains valid in that case as well. Earlier results on the existence of a small solutions of equation (16) and (20) (Atkinson, Elbert, 2000, Hatvani, 2000), and the extension of the Armellini-Tonelli-Sansone theorem to equation (16) in the continuous case (see Bihari, 1984) are valid for arbitrary $n > 0$, in addition, the proofs of these theorems did not require to distinguish cases with respect to *n*. We ran computer simulations which also support our conjecture. We hope we can prove it by modifying our technique of proof.

References

Agarwal, R. P. (1992): Difference equations and inequalities, Monographs and Textbooks in Pure and Applied Mathematics, 115, Marcel Dekker, Inc. New York.

Agarwal, R. P. (2002): Oscillation theory for second order linear, half-linear, superlinear and sublinear dynamic equations, Kluwer Academic Publishers Dordrecht-Boston-London.

Atkinson, F. V., Elbert, Á. (2000): An extension of Milloux's theorem to half-linear differential equations, Proceedings of the 6th Colloquium on the Qualitative Theory of Differential Equations (Szeged, 1999), No. 8, (electronic), Proc. Colloq. Qual. Theory Differ. Equ., E. J. Qualitative Theory of Diff. Equ., Szeged.

Bihari, I. (1957): Ausdehnung der Sturmschen Oszillations und Vergleichssätze auf die Lösungen gewisser nichtlinearen Differentialgleichungen zweiter Ordnung, Publ. Math. Inst. Hungar. Acad. Sci., 2, 154-165.

Bihari, I. (1984): Asymptotic result concerning equation $x''|x'|^{n-1} + a(t)x^{n} = 0$. Extension of a theorem by Armellini-Tonelli-Sansone, Studia Sci. Math. Hungar. 19, no. 1, 151-157.

Bognár, G. (2010): Similarity solution of a boundary layer flow for non-Newtonian fluids, Int. J. Nonlinear Sci. Numer. Simul., 10, 1555-1566.

Bognár, G. (2011): On similarity solutions to boundary layer problems with upstream moving wall in non-Newtonian power-law fluids, IMA J. Appl. Math., 1-17.

Došlý, O. (2004): Half-linear differential equations, Handbook of differential equations, 161-357, Elsevier/North-Holland Amsterdam.

Došlý, O., Řehák, P. (2005): Half-linear differential equations, Elsevier/North-Holland Amsterdam.

Elbert, Á. (1981): A half-linear second order differential equation. Qualitative theory of differential equations, Vol. I, II (Szeged, 1979), 153-180, Colloq. Math. Soc. János Bolyai, 30, North-Holland Amsterdam-New York.

Elbert, Á. (1996/97): On asymptotic stability of some Sturm-Liouville differential equations, General Seminar in Mathematics, University of Patras, 22-23, 57-66.

Elbert, Á. (1997): Stability of some difference equations, Advances in Difference Equations: Proceedings of the Second International Conference on Difference Equations and Applications, Veszprém, Hungary, 7-11 August 1995, Gordon and Breach Science Publishers Amsterdam, 155-178.

Elbert, Á (2001): On damping of linear oscillators, Studia Sci. Math. Hungar. 38, 191-208.

Hartman, P. (1948): On a theorem of Milloux, Amer. J. Math., 70, 395-399.

Hatvani, L. (1998): On the existence of a small solution to linear second order differential equations with step function coefficients, Dyn. Contin. Discrete Impuls. Syst., 4, 321-330.

Hatvani, L. (2000): On stability properties of solutions of second order differential equations, Proceedings of the 6th Colloquium on the Qualitative Theory of Differential Equations (Szeged, 1999), No. 11, (electronic), Proc. Colloq. Qual. Theory Differ. Equ., E. J. Qualitative Theory of Diff. Equ., Szeged.

Hatvani, L., Székely, L. (2006): On the existence of small solutions of linear difference equations with varying coefficients, J. Difference Equ. Appl., 12, No. 8, 837-845.

Hatvani, L., Székely, L. (2011): Asymptotic stability of two dimensional systems of linear difference equations and of second order half-linear differential equations with step function coefficients, E. J. Qualitative Theory of Diff. Equ., 38, 1-17.

Jacobson, N. (1953): Lectures in abstract algebra, II. Linear algebra, Springer Verlag New York-Heidelberg-Berlin.

Karsai, J., Graef, J. R., Li, M. Y. (2000): On the phase volume method for nonlinear difference equations, Int. J. Differ. Equ. Appl., 1, 17-35.

Macki, J. W. (1983): Regular growth and zero tending solutions, Ordinary differential equations and operators (Dundee, 1982), Springer Berlin, 358-374.

Milloux, H. (1934): Sur l'equation differentielle $x''+A(t)x=0$, Prace Mat.-Fiz., 41, 39-54.

Prodi, G. (1950): Un'osservazione sugl'integrali dell' equazione $y''+A(t)y=0$ nel caso $A(x)\to\infty$ per $x\to\infty$ Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat., 8, 462-464.

Szegő, G. (1959): Orthogonal polynomials, American Mathematical Society New York.

Székely, L. (2011): On Stability Properties of Second Order Differential Equations with Step Function Coefficients (in Hungarian), PhD thesis, Doctoral School in Mathematics and Computer Science, University of Szeged (unpublished).

Trevisan, G. (1954): Sull'equazione differenziale $y''+A(t)y=0$, Rend. Sem. Mat. Univ. Padova, 23, 340-342

2. INSTITUTE FOR PROCESS ENGINEERING

PROFESSOR DR. JÁNOS BEKE
DIRECTOR OF THE INSTITUTE

Dear Reader,

The Institute for Process Engineering is a dominant education and research unit at the School of Mechanical Engineering that was founded as integrations of two earlier independent departments (Department of Automotive and Thermal Technologies and Department of Energetics and Food Industry). The current institute consists of three professional working areas (Automotive Technologies, Energetics and Measurement Technologies)

During our educational and scientific activities we paid special respects to the following topics:

- Basic technical knowledge forming the appropriate engineering approach,
- Engineering thermodynamics, electrical engineering and electronics,
- Basic and applied knowledge of energy conversion, energy utilization and energy economy,
- Special technical knowledge for development of environmental industry and utilization of alternative energy sources,
- Basic and applied knowledge of automotive and off-road techniques and technologies
- Disciplines of measurement technology and process engineering needed to control, regulate and automate different engineering processes,
- Environmental- and market-conscious development of production and processing technologies.

Following our tradition, the Institute for Process Engineering is eager to present the most significant scientific activities from the year of 2011.
In last year we paid special attention on the following scientific topics

- Modelling of terrain and towed vehicle interactions. In frame of this topic we designed and tested mathematical models to determine the stresses on vehicles moving on different types of terrain.
- Analysing the microwave assisted drying Processes of fruits, vegetables and grains. During this research work – by using the results of laboratory experiments – we determined the typical energy efficiency factors for microwave drying of some sorts of agricultural products. Furthermore, we detected the non-thermal effect of the microwave energy.

- Investigation of different heat pump systems. The objective of this project was to find the most suitable solution for heating and cooling conference rooms and other office buildings.
- Utilization possibilities of renewable energy sources. This research topic was about the probable thermal utilization of wind, solar and geothermal energy for heating and air conditioning.
- Modelling different engineering processes from energetic point of view. Inside this activity we tested different mathematical models (theoretical, empirical and semi-empirical) that are frequently used to simulate agricultural energetic processes.

More details and other information can be found: www.fomi.gek.szie.hu

ANALYSING THE MICROWAVE DRYING PROCESS OF POTATO, APPLE AND ONION SAMPLES FROM ENERGETIC POINT OF VIEW

János BEKE, Zoltán KURJÁK, Kornél BESSENYEI

Department of Energetics, Institute for Process Engineering

Abstract

From the practical point of view the efficiency of the energy transfer during drying must be considered as dominant parameter. Although the energy consumption of dehydration under microwave conditions is theoretically lower even by an order of magnitude than that of convective drying, the real efficiency of energy transfer in the microwave field is influenced by numerous factors. In this study, the most determinative relations between the energy transfer efficiency and the drying parameters were investigated. During the experiments several different drying parameters were set up to study the effects of domain variables in the apple, potato and onion drying process.

Based on laboratory experiment, we determined the typical energy efficiency factors for microwave drying of the investigated samples. Furthermore, we detected the non-thermal effect of the microwave energy.

Keywords

Microwave drying, vegetables, non-thermal effect

1. Introduction

Drying is a compound and energy consuming process. Moreover, drying in agriculture is a much more complex procedure than in industry, because in parallel with reducing the water content of the living material, physiological and morphological changes occur. So, the parameters of the drying process should be selected on a scientific base in order to prevent structural damage to the product and to optimize the drying process energetically (Mujumdar, Beke 2002).

Microwave energy transfer is based on various form polarization of the drying material behaving as dielectric substance. Part of the electric energy is transformed into heat, which is suitable to eliminate the redundant moisture content.

Agricultural products are not homogenous dielectric materials, thus their special characteristics largely influence their behaviour in microwave fields, and make processes more difficult to acquaint (Ludányi and Beke 2002).

2. Theoretical background

By puting moist material into a microwave cavity, it causes space perturbation and changes the resonance frequency of the dissipation area as follows (Ludányi and Beke 2000)

$$\frac{\Delta\omega}{\omega_0} = -\frac{1}{W}\int_V \left(\frac{d\varepsilon}{\varepsilon} w_e + \frac{d\mu}{\mu} w_m\right)\frac{dV}{V} \tag{1}$$

Since for biological materials $\mu=1$ and $d\mu=0$ using the practical perturbation equation for ε' and ε'' the next two equations can be derived

$$\varepsilon' = \frac{Kf_0}{Kf_0 - 2\Delta f} \tag{2}$$

$$\varepsilon'' = \frac{P_d}{55{,}6\cdot 10^{-14}\cdot E^2\cdot f} \tag{3}$$

In the case of electromagnetic radiation emitted upon a metal surface the Lebedev-type jet-pressure (p_{jp}) is well known, which is based on the law of energy and momentum conservation (Ludányi, Beke 2002)

$$p_{jp} = \frac{C_B}{c_\lambda} S_P \tag{4}$$

By introducing the reflection coefficient ($\Gamma=e_r/e_i$) the density of the absorbed energy comes from the following formula

$$e_m = \frac{c_i \cos\Psi_i}{c_m \cos\Psi_m}(1-\Gamma)\cdot e_i \tag{5}$$

The microwave energy of a magnetron is forced through a funnel-shaped antenna, and to prevent the bypass radiation wave traps are applied.

The electromagnetic energy is reflected towards the sample surface with the help of special plates, which are adjusted to the following calculated Brewster-angle (Ψ). In this case $\Gamma = 0$.

$$\Psi = \text{arctg}\sqrt{\frac{\rho_m^2 - \varepsilon_m \rho_i^2}{\rho_i^2(\varepsilon_m - 1)}} \tag{6}$$

In cases of usual measuring systems the instruments show detected performance (P_{det}), and the real relevant performance can be calculated by using the so-called coupling coefficient (C_{dB}) as follows

$$P = 10^{\frac{C_{dB}}{10}} \cdot P_{det} \quad [W] \tag{7}$$

It is an essential problem that normally in biological materials there is no homogenous internal moisture and temperature distribution, because these types of materials are living tissues with complex composition and their biochemical processes continue during the significant part of the treatment (Obermayer et al. 2009). Based on previous scientific findings, it can be assumed, that besides the usual microwave field parameters and the physical properties of the material placed in microwave zone – depending on the apparatus construction – the applied microwave performance and the residence time have also significant influence on drying process.

3. Materials and methods

Measuring instrumentation

For measurements, a special microwave device was used, whose power could be continuously adjusted using a toroid transformer. During drying, the transmitted and reflected energy of the magnetron were measured, and the dissipated performance was calculated. The temperature and moisture content of the entering and outgoing air of the drying chamber were also measured. The set-up of the measuring instrumentation can be seen in Fig. 1.

Change in the sample's mass was measured every minute by a scale under the sample holder. Samples were turned 90 degrees after each measurement to compensate the effect of the non-homogeneous microwave field. To prevent precipitation of the moisture evaporating from the samples, the drying chamber was ventilated with a fan.

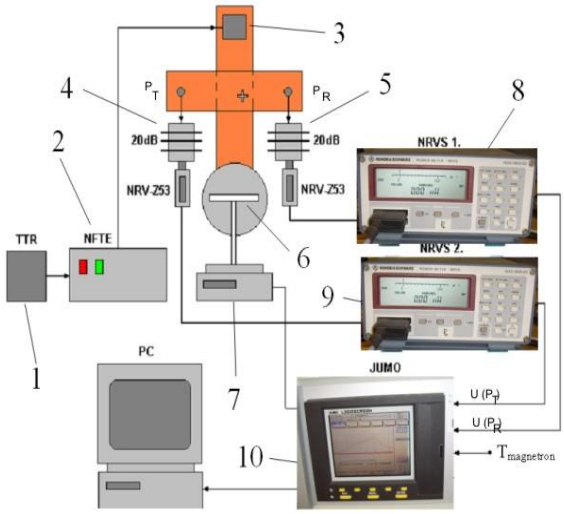


Figure 1. The measuring setup
1. toroid transformer; 2. transformer and switch unit; 3. magnetron;4. sensor for transmitted energy (P_T); 5.sensor for reflected energy (P_R); 6. sample holder; 7. balance; 8 and 9. power meters 10. data recorder

Sample preparation

Drying experiments were performed using three materials, potato, apple and onion, which have different characteristics. To ensure the same starting parameters, sample preparation for the three materials was done in the same way. They were cut into 10 mm cubes, of which 80-100 grams were placed in the drying chambers (see Fig.2.)

The microwave drying continued until the materials reached 5-15% moisture content, depending on the starting produce. After that, the samples were placed in a drying oven for 24 hours to determine the entire moisture content in accordance with the ASAE standards.

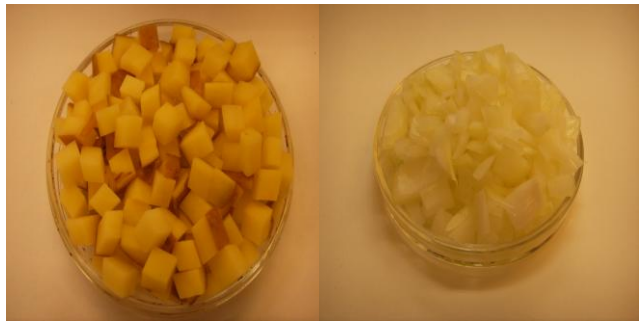


Figure 2. Prepared samples (potato and onion)

Measurements for each product were performed at 140 and 160 transformer voltage accompanied by 456 and 523 Watt power.

4. Results and discussion

Potato drying

Average input and output moisture data of potato drying obtained from the eight series of experiments can be seen in Table 1.

Table 1. Moisture content of potato drying

U_T [V]	w_1 [%]	w_2 [%]
140	80.27	18.31
140	79.39	19.75
140	79.37	19.25
140	80.56	19.08
Average^{140}	79.90	19.10
Δm_w=60,10 g/sample		
160	80.61	18.58
160	80.49	18.71
160	81.50	19.45
160	81.01	19.59
Average^{160}	80.90	19.08
Δm_w=61.14 g/sample		

The obtained kinetic curves are shown by Figure 3. Then the drying rate curves were calculated (Figure 4). Microwave performance can be seen in Figure 5.

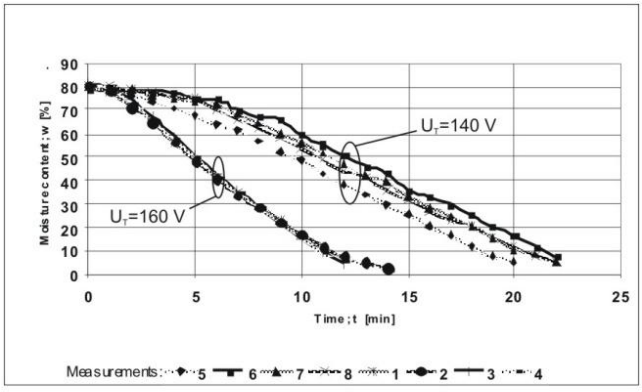


Figure 3. Kinetic curves of potato drying

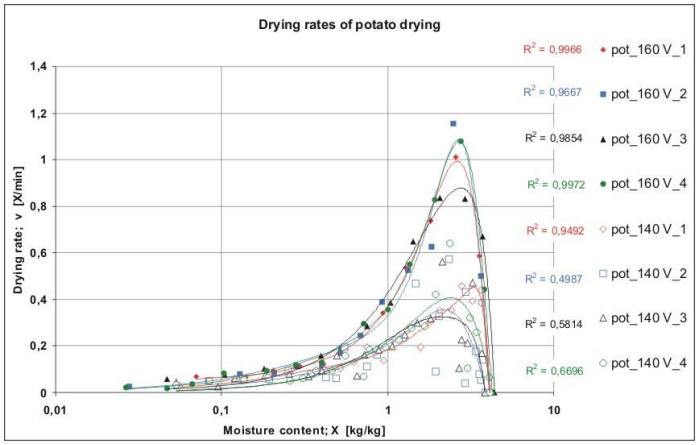


Figure 4. Drying rates of potato drying

By looking at the obtained diagrams, it can be seen at first glance that there are significant differences between the results coming from the 140 V and 160 V series. We can observe much large deviations between the measured values coming from the 140 V series than those of the 160 V series.

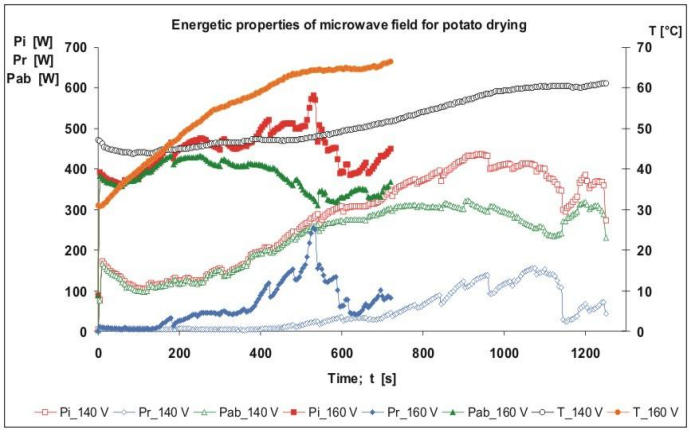


Figure 5. Energetic properties of microwave field for potato drying

These phenomena can be partly caused by the apparatus construction and partly to the bigger and nondescript convective and surface heat loss taking place during the experiments. However, the rate of increase in the drying intensity between the 140 and 160 series cannot be explained by these reasons entirely. It is probable that other factors such as energy absorption or rate of dissipation have serious influences on the dewatering process. To clarify these problems more precisely, additional experiments and discussions are needed.

In spite of some obscurity in data of the 140 series, it is evident that by applying magnetron performance of 160 V, the energy utilization becomes more efficient. As Figure 5 shows, coming to the end of drying process the temperature of the magnetron reaches practically the same high value in both cases, but the 140 V series has longer interval of high temperature and consequently, in this case, the energy loss is higher.

Apple drying

The apple cubes to be dried had 6-8 % more moisture than the potato samples and there were some differences in the final moisture content as well (Table 2.), but these differences had no significant influences on the drying parameters. By processing the measured data from apple, similar kinetic and drying rate diagrams to potato were obtained (Figure 6 and 7). As far as the field parameters are concerned the magnetron temperature reaches 80 ^{o}C at the end of the dewatering process in both cases and – similarly to the case of potato drying – the longer period of high temperature belongs to the 140 V series (Figure 8).

Table 2. Moisture content of apple drying

U_T [V]	w_1 [%]	w_2 [%]
140	86.43	17.54
140	87.06	20.46
140	86.82	17.95
140	87.77	16.88
Average^{140}	87.02	18.21
Δm_w=67.16 g/sample		
160	89.28	16.65
160	89.72	14.77
160	88.97	24.01
160	91.24	20.61
Average^{160}	89.80	19.01
Δm_w=69.92 g/sample		

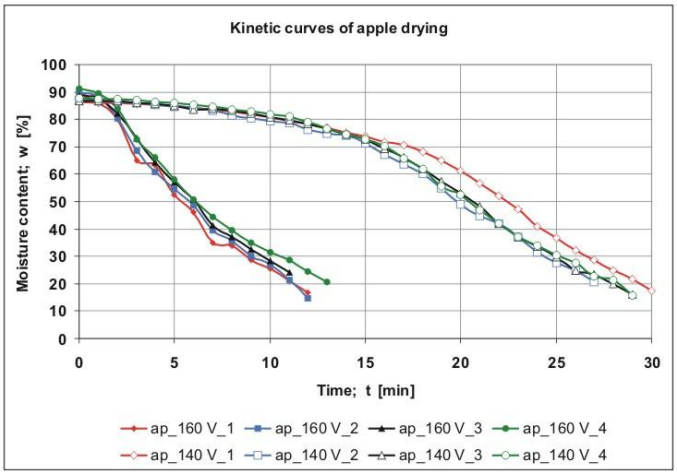


Figure 6. Kinetic curves of apple drying

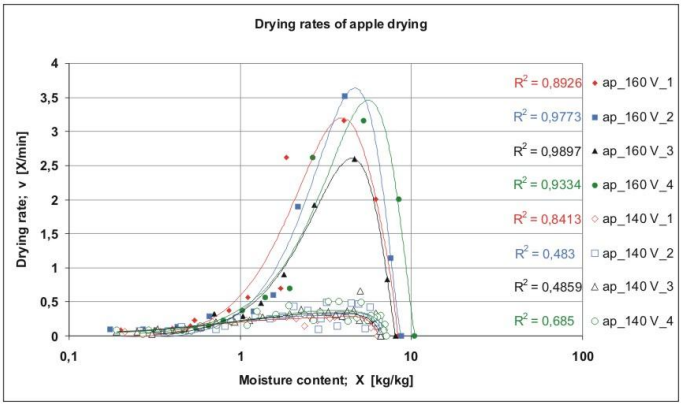


Figure 7. Drying rates of apple drying

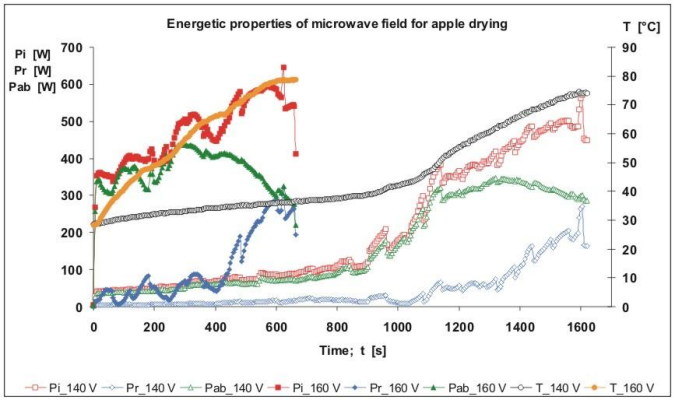


Figure 8. Energetic properties of microwave field for apple drying

Onion drying

The moisture content of onion samples are listed in Table 3 and the process parameters are in Figure 9, 10 and 11. As the data show, onion has the slowest drying process amongst the investigated samples. The most probable reason for this phenomenon is the special leafage structure of onion. Under influence of the energy and mass transport that takes place during drying, the onion leaves move away from each other while a wax layer forms on their surfaces. This wax coat reduces the moisture diffusion to the surface and the surface component transfer coefficient as well. Toward the centre of the onion sample, this process is enhanced and the distances between the leaves is becoming less and less. As a result of this process, the mass transport inside the material continues while the outer leaves become dried and their surface completely closed by the wax film, thus the moisture movement from the inner parts of the samples is almost blocked.

Table 3. Moisture content of onion drying

U_T [V]	w_1 [%]	w_2 [%]
140	92.68	19.47
140	92.83	18.95
140	93.27	19.68
140	94.63	19.74
Average^{140}	93.35	19.48
Δm_w=73.39 g/sample		
160	93.62	18.08
160	93.54	18.14
160	93.13	20.08
160	93.94	21.13
Average^{140}	93.36	19.36
Δm_w=73.40 g/sample		

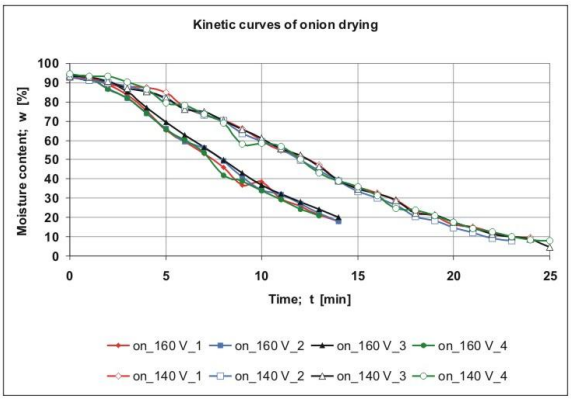


Figure 9. Kinetic curves of onion drying

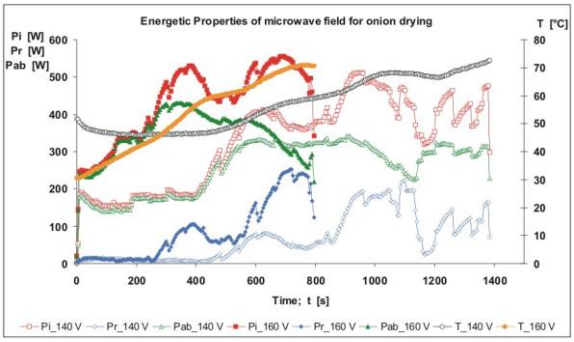


Figure 10. Drying rates of onion drying

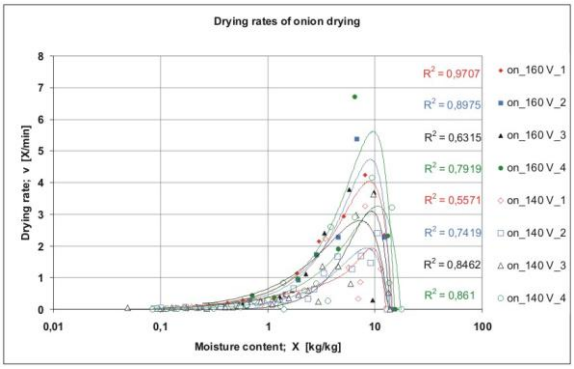


Figure 11. Energetic properties of microwave field for onion drying

Energetic comparison

The typical energetic parameters are listed in Table 4, 5 and 6 for potato, apple and onion respectively.

Table 4. Energetic parameters of potato drying

	U_T [V]	W_i [J]	W_d [J]	W_{ia} [J]	W_{da} [J]	q_{ia} [J/g]	q_{da} [J/g]
1	160V	234081	201106	229506	196070	3753.8	3206.9
2		222741	185215				
3		225914	200289				
4		235287	197670				
5	140V	276834	234308	262531	218445	4368.2	3634.7
6		239157	201954				
7		259259	212302				
8		274874	225219				

The tables display the relevant transformer voltage (U_T), the used energy amounts of different measurement series (W_i, W_d), the mean values for transformer voltage of 140 V and 160 V and finally the average specific energy consumptions (q_i, q_d in J/g water) calculated on the incident and the dissipated microwave energy. Potato has a rather intensive dewatering process. Although the morphological structure of apple somehow is similar to that of potato the mass transfer coefficient – especially in the end of the apple drying – is definitely lower. The sugar content of these samples may explain this. It is possible that partly caramelized sugar is hindering the moisture movement through the cell wall. Because of the low diffusion coefficient in onion drying the difference in energy consumption between the 140 and 160 series is significantly higher than those of potato and apple drying.

Table 5. Energetic parameters of apple drying

	U_T [V]	W_i [J]	W_d [J]	W_{ia} [J]	W_{da} [J]	q_{ia} [J/g]	q_{da} [J/g]
1	140V	264180	225758	329219	258914	4890.4	3846.0
2		348969	276686				
3		353260	266835				
4		350466	266378				
5	160V	308180	240952	314209	243957	4493.8	3489.1
6		312510	243155				
7		310709	249409				
8		325437	242315				

Table 6. Energetic parameters of onion drying

	U_T [V]	W_i [J]	W_d [J]	W_{ia} [J]	W_{da} [J]	q_{ia} [J/g]	q_{da} [J/g]
1	140V	370555	290474	364033	288114	4960.3	4850.4
2		359351	287259				
3		335317	280147				
4		369108	294579				
5	160V	350388	279046	343202	274540	4675.8	3740.3
6		344502	278936				
7		335714	269962				
8		342205	270215				

Using the specific dissipated energy consumption of potato drying (q_{da}) as base-unit (R_d=1) we can create a relative parameter that gives us a chance to compare the drying process of different samples from energetic point of view (Figure 12).

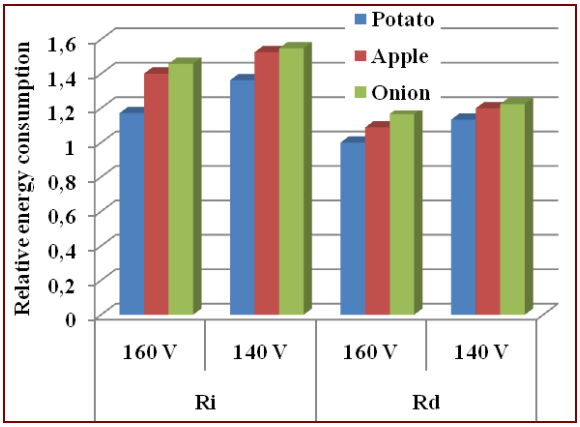


Figure 12. Comparison of relative energy consumption of different materials

By carrying out experiments a serious anomaly could be experienced, because – especially in the most intensive dewatering interval – a part of the exiting moisture was liquid and the material temperature was higher than those that could be derived from the classical heat balance theory. In this way, it is possible that the moisture movement is partly forced by the non-thermal effect of microwave energy. This wave pressure causes a pressure gradient pointing to the surface. To clarify this question theoretically some indirect evidence can be found for its existence using the classical jet pressure equation

$$p_{jp} = e_i \cos^2 \Psi_i + e_r \cos^2 \Psi_i - e_m \cos^2 \Psi_m \tag{8}$$

By substituting the ($\Gamma = e_m/e_i$) reflection coefficient and taking into consideration the possible simplification possibilities the following expression is obtained

$$p_{jp} = e_i \cos^2 \Psi_i (1+\Gamma) - e_i \cos \Psi_i \sqrt{\varepsilon_m - \sin^2 \Psi_i}\,(1-\Gamma) \tag{9}$$

When the radiation is established under the relevant Brewster angle, the reflection will cease ($\Gamma = 0$). In this case $\varepsilon_m=1$, which is impossible under the usual conditions of drying biological materials.
If we visualize the most intensive interval of the dewatering process, the pressure gradient caused by the wave pressure of microwave radiation can be shown directly. In the Figure 13 notable amounts of liquid water can be observed among particulates that came from the drying substance without phase change, thus the moisture movement cannot be caused by the thermal effect of microwave power entirely.

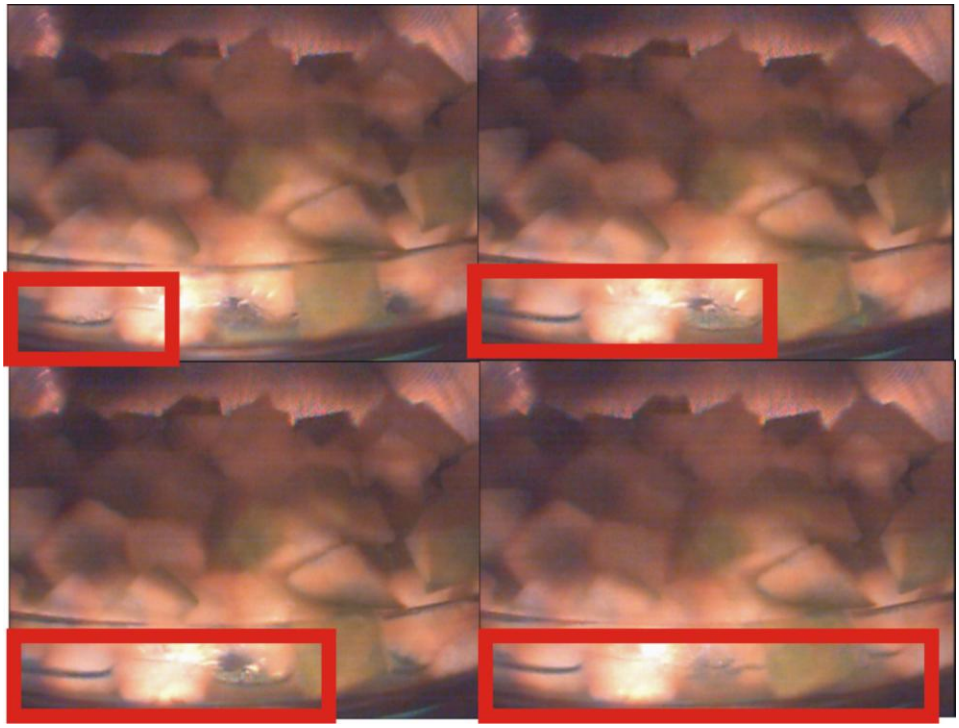


Figure 13. Liquid water among drying apple particulates in electromagnetic field (U_T = 160V)

5. Conclusion

It is evident that drying intensity mainly depends on the morphological structure of the treated material. However, besides the energy and mass transport, the microwave energy can induce physicochemical changes that have significant influences on the drying intensity. In order to optimize energy consumption, the applied microwave performance and residence time should be harmonized with the features of the material to be dried.

It is most probable that in the constant drying rate period of substances with high moisture content the non-thermal effect of microwave energy plays some role in the drying of biological materials.

Symboles

D	diffusion coefficient
C_{dB}	coupling coefficient
E	electric field strength
K	ratio between the absorbent and the resonator

P	performance
R	ratio of specific energy consumptions
S_p	Poynting vector
T	temperature
U	voltage
V	volume
W	energy
X	moisture content (dry basis)
c	specific heat
c_λ	wave velocity
e	wave density
f	frequency
m	mass
p_{jp}	jet pressure
q	specific energy
t	time
v	drying rate
w	moisture content (wet basis)
w_e	electric energy density
w_m	magnetic energy density

Greek letters

Γ	reflection coefficient
Ψ	angle wave related to surface normal
ε	dielectric constant
ε”	imaginary component of complex dielectric constant
ε’	real component of complex dielectric constant
λ	wave length
μ	magnetic permeability
ρ	density
ω	angle frequency
Δω	frequency drift

Subscripts

M	magnetron
T	transformer
a	average
ab	absorbed
i	incident
r	reflected
det	detected
d	dissipated
m	material

w water
0 empty chamber
1 beginning
2 final

References

Bihercz, G. and Kurjak, Z (2003) Comparison of Convective and Microwave Drying Process of Tomato and Carrot Samples. Procedings of EUDrying03. Heraklion-Crete Greece 268-277.

Ludanyi, L. – Beke, J. (2000) Study of Intermittent and Continouos Microwave Drying Method. Hungarian Agricultural Engineering. Budapest 13. 41-43. p.

Ludanyi, L – Beke, J. (2002) Energetical Problems of Dissipation heat transfer by Microwaves. Hungarian Agricultural Engineering. Budapest, 15. 50-52. p

Obermayer, D., Gutmann, B. and Kappe, O.C. (2009) Microwave Chemistry in Silicon Carbide Reaction Vials Separating Thermal from Nonthermal Effects Angew. Chem. Int. Ed. 48. 8321 –8324. p.

Mujumdar, A.S., Beke, J. (2002) Industrial Drying (in Hungarian) Szakturdás Kiadó Ház. Bp., 247 p.

EXPERIMENTS WITH GROUND SOURCE HEAT PUMP SYSTEM

László TÓTH[1], Béla ÁDÁM[2], Károly PETRÓCZKI[3], Zoltán GERGELY[3]

[1]Department of Energetics, Institute for Process Engineering
[2]Hydro Geodrilling Kft.
[3]Department of Measurement Technology, Institute for Process Engineering

Abstract

The objective of the project was to find the most suitable solution for heating and cooling a conference room in an office building at Gödöllő, Hungary. Ground source heat pump system assisted by four solar photovoltaic cells for electric energy generation was installed in 2009. For a collector system two 100 m deep U-type-tubes were placed in separate boreholes. The system works in heating and cooling mode. The heat pump is not used for cooling. The excess heat is delivered to the soil by a heat exchanger situated in a depth of 15 m.
This research project are supported by the European Union and co-financed by the European Social Fund (TÁMOP-4.2.1.B-11/2/KMR-2011-0003).

Keywords

Ground source heating, heat pump, solar energy, heating, cooling

1. Introduction

The price of fossil fuel and the need for an independent and pollution free energy source motivates businesses and households looking for alternative energy sources in Hungary especially for heating and cooling (Komlós, 1997). Heat pumps are suitable for the weather in Hungary as it can provide heating in winter and cooling in summer and soil conditions in this geographical area are suitable for collecting heat from the relatively high temperature ground (Tóth, 2008). The disadvantage of heat pumps is that they require external power and as a solution it is possible to combine it with a solar system for electricity generation. A solar assisted ground source heat pump system is installed in one of the office buildings at Gödöllő. The purpose of this project was to find a cheaper, cleaner and more independent heat energy source.

2. Design and installation

A solar assisted ground source heat pump system is installed for heating and cooling a 100 m^2 conference room of an office building at Gödöllő in 2009. The heat pump requires electricity and to provide cheaper and more environmental friendly energy solar collectors are used for power generation. The system consists of a heat source - ground collector system, water-to-water heat pump

which raises the temperature of the collected heat and transfers it in form of the water to the heat reservoir (HR) and further to a heat distribution system – water is transferred from the heat reservoir to fan-coil units (Ochsner, 2007). The collector system consists of two 100 m deep boreholes. A location of boreholes and a floor plan of conference room are shown in Figure 1.

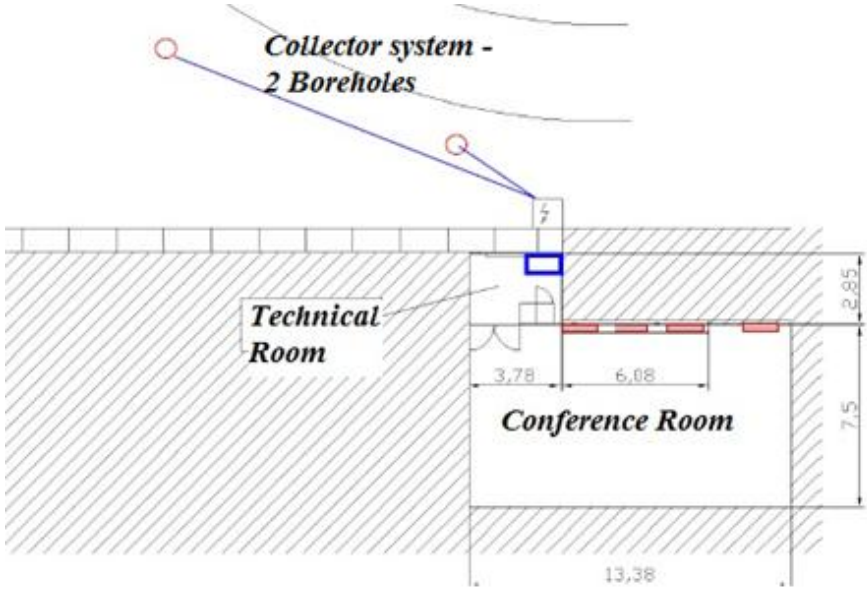


Figure. 1. Plan of conference room and boreholes

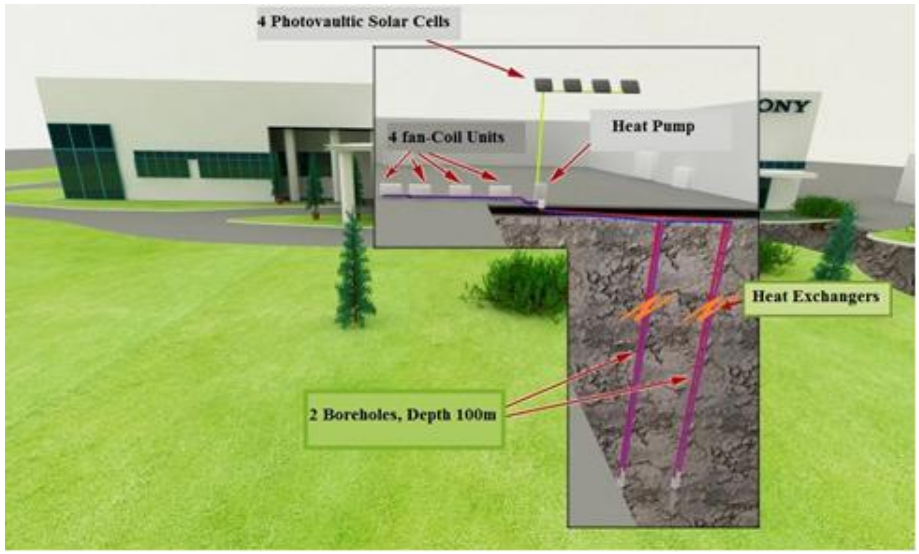


Figure 2. Solar assisted heat pump system

There are two different types of ground source heat pump systems – a direct expansion system where a refrigerant is circulated through the heat pump and the collector system; and an indirect system where a mix of water with antifreeze circulates through the collector system and the refrigerant circulates only in the heat pump. In this case it is an indirect system (Steimle n.d.). The distribution heat system of the conference room consists of four series connected fan-coil units, the heat exchangers in which water is circulated and heated or cooled air is transferred to the room by a fan (Tóth, 2008).

The system is assisted by four solar photovoltaic cells for electricity generation to run the circulation pump. System is shown in Figure 2.

3. The components of the heat pump system

The system consists of two parts - primary and secondary side which are separated by the heat pump (www.hidro-geodrilling.hu). The primary side consists of a vertical ground collector system with the heat exchanger and the secondary part consists of a heat distribution system with four fan-coils. A compressor of the heat pump works periodically by turning on and off. To maintain a correct and constant flow a heat reservoir is used. In summer in cooling mode the heat exchanger is used without using heat pump. Figure 3. and Figure 4. shows a schematic and components of the installed heat pump system.

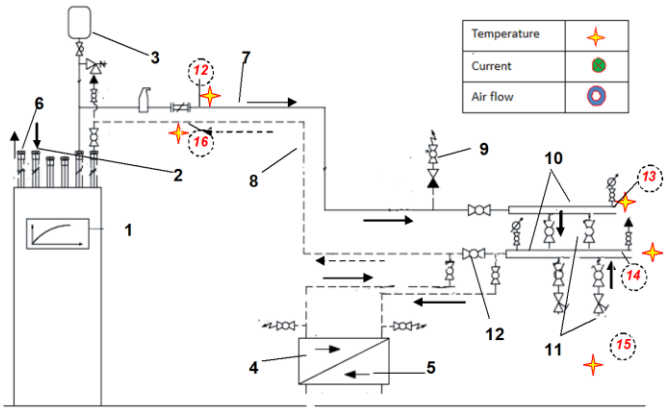


Figure 3. Primary side with measuring points
1. Heat pump 2. The secondary side (return pipe) 3. Compensator container 4. The primary side of the refrigeration heat exchanger 5. The secondary side of the refrigeration heat exchanger 6. The secondary side (outgoing pipe) 7. The branch of the primer side (outgoing pipe) 8. The secondary side (return pipe) 9. Recharging tap 10. Boreholes return pipes 11. The boreholes outgoing pipes 12. Relay tap in case of refrigeration in summer

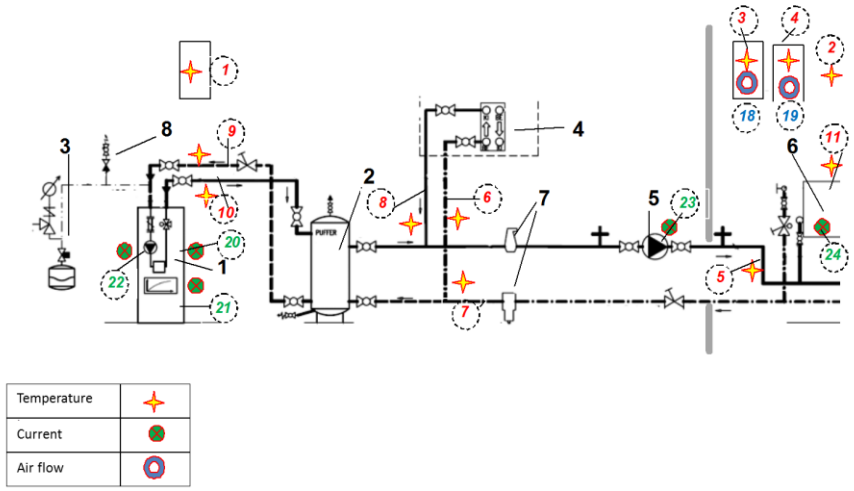


Figure 4. Secondary side with measuring points
1. Heat pump 2. Heat reservoir 3. Compensator container 4. Heat exchanger for refrigeration in summer 5. Circulation pump 6. Fan-coil units 7. Cleaning units

In heating mode heat is collected from the soil by a special liquid – mix of water and glycol running through U-type tubes situated in boreholes. In heat pump the heat is transferred to a refrigerant by heat exchanger and the temperature of it is increased by a compressor and the heat is transferred to the secondary loop by the heat exchanger and is delivered to a heat reservoir where the heated water is stored. It equalizes the temperature of the water because the compressor works periodically. The water flow is circulated by a pump and transferred to four fan-coils. In cooling mode the process is reversed but a ground heat exchanger is used for giving the heat to the soil directly. The heat exchanger is in the 100 m deep boreholes. By using cooling heat exchanger the power requirement is lower.

4. Measuring system

The measuring points of the system can be seen on the Fig. 3. and 4. The numbers of the measuring points are in dotted circles, the notation is the following:

1. open-air temperature in the shade,
2. air temperature in the heated or cooled room,
3. air temperature from the air mixer No. 4.,
5. temperature of the outgoing pipe to the fan-coils,
6. temperature of the incoming pipe of the “summer heat exchanger” in case of cooling,
7. temperature of the a fan-coils return pipe,

8. temperature of the return pipe of the "summer heat exchanger",
9. temperature of the water in return pipe from the heat reservoir,
10. temperature of the water in the outgoing pipe from the heat pump to the heat reservoir,
11. air temperature from fan-coil units,
12. glycol temperature in outgoing pipe from system to the bore holes collector,
13. glycol temperature in outgoing pipe at the divisors,
14. glycol temperature in return pipe at the divisors,
15. soil temperature in 2,5 m depth,
16. glycol temperature in return pipe from bore holes collector to the heat pump.

5. Results

To analyse the performance of the system monitoring was carried out. Measurements were made for a period of 20 days starting from 30th of October 2010 to 18th of November 2010. Temperatures in 16 different points were measured. Most significant results of minimal, maximal and average temperatures of 24 hour measurement in 6th of November are shown in Table 1.

Table 1. Measured temperatures

	Max T °C	Min T °C	Average T °C
Temperature of the soil (16m deep)	13,4	11,7	12,6
From the heat pump (out)	37,5	24,9	27,7
Back to the heat pump (in)	30,1	24,9	27,1
Air temperature in the room	25,0	23,4	24,3

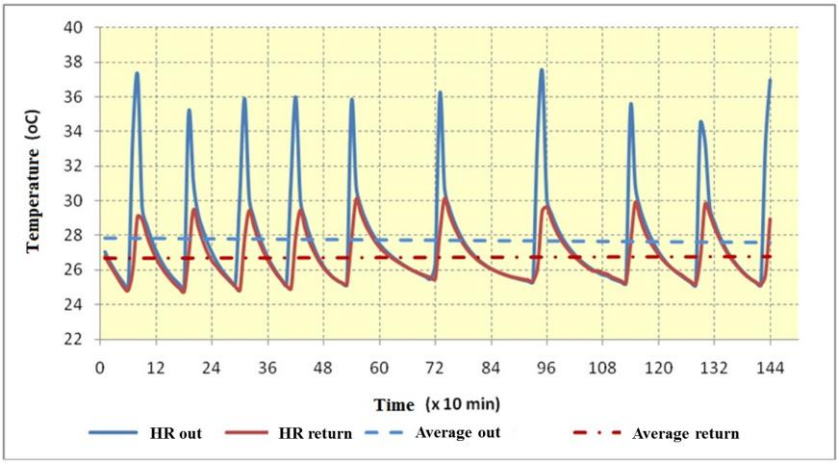


Figure 5. Outgoing and returning water temperatures of the heat reservoir (HR)

Temperature parameters are closed to design ones. The difference between the heat pump minimal and maximal temperature is because the compressing process of the heat pump doesn't happen constantly. The temperature sampling time was 10 minutes. Figure 5. shows 24 hour (144 × 10 min) temperature diagram of outgoing and returning water of the heat reservoir.

24 hours period of temperatures of borehole exchangers and the soil in 16 m depth are shown in Figure 6. The green line shows the soil temperature in 16 m depth which is between 11,8 and 13,4°C.

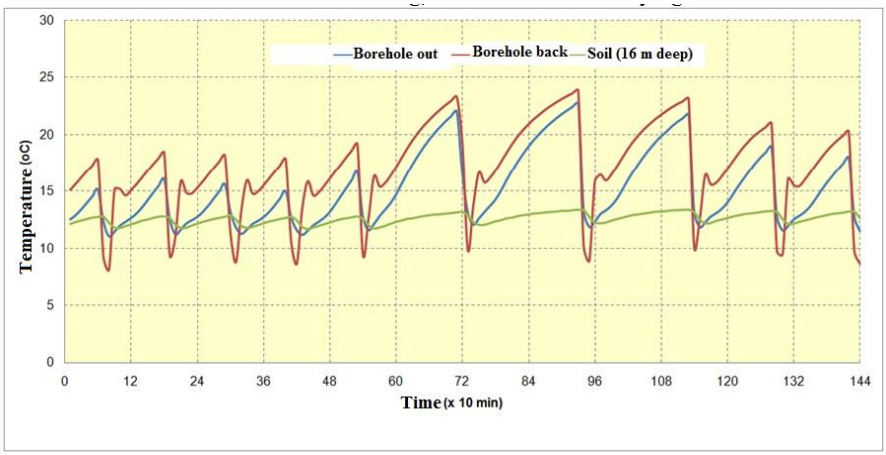


Figure 6. Soil and U-tube-type borehole exchanger temperatures

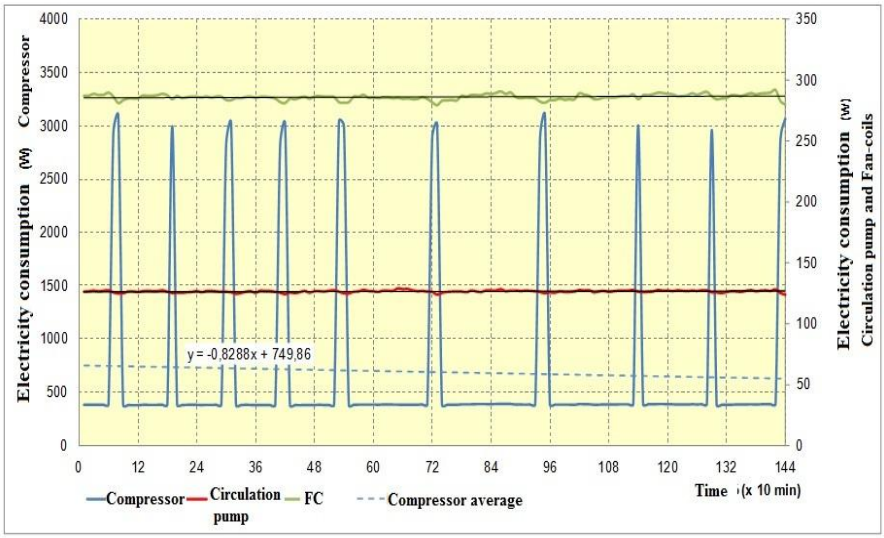


Figure 7. Electric consumption of the system

To analyse the performance of the hall system it is significant to analyse the consumption of the electricity. The system provides electrical energy for compressor, fan-coil units and the circulation pump. Performance of the system in a cold winter day when average outdoor temperature is -3 °C is shown in Fig. 7.

During a 24 hour period the compressor of the heat pump works with interval 17 – 25 min. The electric power consumption of the compressor unit during compressor turn-off only about 380 W, in turn-on state the electric consumption increases up to 3100 W. Electric consumption of fan-coils varies from 279,7 W to 292,2 W and the circulation pump's consumption is between126,5 and 129,9 W.

A variation of the supplied water temperature to fan-coils and the supplied air temperature to the room is shown in Figure 8. The supplied water temperature varies from 24,5 °C to 31,5 °C. And the supplied air temperature varies from 21,0 °C to 23,8 °C.

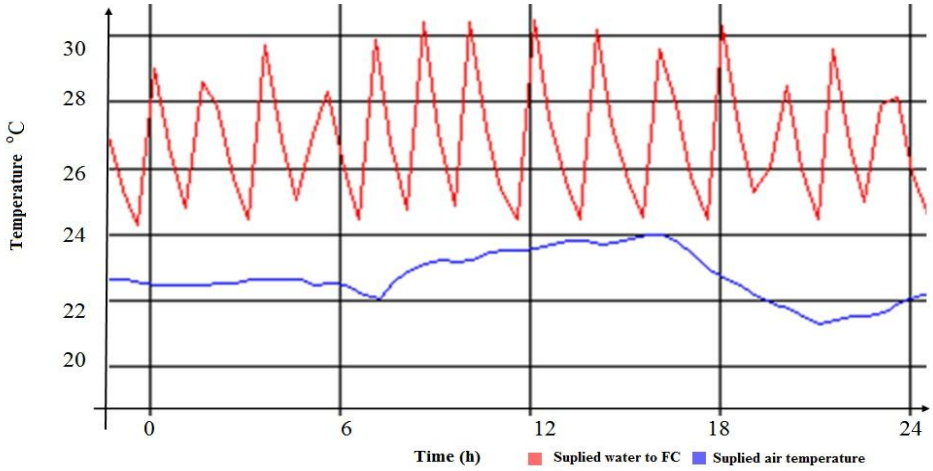


Figure 8. Fan-coil performance

Measured data proves that the project system meets the requirements. The operation was satisfactory and there was no trouble during a long period. Measured temperatures are close to designed temperatures and a comfort temperature +22 °C (± 2°C) was provided in the conference room even in the coldest months of the year.
Energy consumption of the running system – heat pump, circulation pump and fan-coils is shown in Table 2. It is compared with energy consumption of a natural gas boiler heating system combined with a separate climate control which would be required to heat and cool the same room. Savings of energy and carbon emissions of one year are shown in the table.

One year energy consumption was 12679,2 MJ (Heat pump 10368 MJ, circulation pump 1728 MJ, fan-coil units 583,2). By using a traditional heating system with boiler and climate control this number could be 4,6 times bigger and saves about 45236 kg carbon emission and € 135 708 with it during one year. Energy consumed by a circulation pump is provided by solar panels on the roof. Parameters and the performance of solar panel such as size, collected energy of one m^2 and a factor of a performance, working hours of a year and produced energy in one year are shown in Table 3.

Table 2. Energy consumption of the heat pump system compared to the energy consumption of a traditional system

	kWh/year	MJ/year		kWh/ year	MJ/ year
Natural gas burner	11550	51975	Heat pump	2880	10368
Split clime	1650	5940	Circulation pump	480	1728
			Fan-Coils	162	583
Together	**13200**	**57915**	**Together**	**3522**	**12679**

Saving	**45236**	MJ/year
CO_2 Saving	**3871**	kg/year

Table 3. Performance of solar panels

m^2	kW/m^2	η	h/year	kWh/year
1,8	1000	0,16	2100	604,8

As shown in table solar panels provide 604 kWh/year but the circulation pump requires only 480 kWh/year. Cost of the heat pump system was € 13 868 (1 drilling of 100 m deep borehole in Hungary costs € 2190).

6. Conclusions

By installing a ground source heat pump system the energy consumption and CO2 emissions are reduced. This solution for heating and cooling provides comfort for users as temperatures are easy to regulate and the system does not require maintenance. The system works error free and provides heating and cooling just by changing a position of a switch. The consumed energy is reduced by 4,6 times which makes up to € 135 708 savings in one year (according to the

electricity and natural gas tariff in 2011) making the system environmental friendly and cost effective for the business. The results of measured data shows that system works without any disorders and appropriate to designed parameters.

References

Ochsner, K. (2007): Geothermal heat pumps: a guide for planning and installing, London.

Sarbu, I., Bura, H. (2011): Thermal tests on borehole heat exchangers for ground–coupled heat pump systems, International Journal of Energy and Environment, Iraq.

Komlós, F. (1997): Épület - Energia - Élettér - Ember: hőszivattyús technika. Magyar Épületgépészet, XLVI. évf. 12. szám, p. 5 – 7

Steimle, F.: Wärmepumpentechnik - Einführung. Gas Wärmepumpen Praxis - Wärmepumpentechnologie, Band IV.

Büki, G. (2001): Magyar energetika a századfordulón számokban. Magyar Energetika, 2001/1. szám, p. 3– 11

Tóth, L., Fogarasi, L., Bihercz, G., Ménesi, A.: (2008): Utilization of Geothermal Heat for Air-Conditioning of Different Small Buildings. Journal of Agricultural Machinery Science. Volume 4, Number 3. 241-246. p. ISSN: 1306-0007

http://www.hidro-geodrilling.hu/index.php

MODELLING OF TERRAIN AND TOWED VEHICLE INTERACTION

László GURMAI, Péter KISS, Lajos LAIB

Department of Automotive Technology, Institute for Process Engineering

Abstract

The purpose of the research is to design and test a mathematical model to determine the stresses on vehicles moving on different types of terrain. Vehicles towed across terrain are acted upon by forces originating from the towing vehicle and from the ground-vehicle relationship, the latter comprising dynamic forces due to soil unevenness and forces due to rolling resistance. The impact of the ground on the vehicle depends on the geometrical features of the terrain and mechanical properties of the soil. The activating force affecting the moving vehicle is related to the effective profile, and the damping effect is the result of the friction among soil particles. A complex mathematical model will be developed using analytical mathematical models and the numerical results of field experiments on a test field. The activating function derived from the model will be tested on a purpose-designed test bench. For certain problems of forces acting on moving vehicles, the equation describing terrain conditions permits a more exact definition of forces which hitherto have only been estimated.

Keywords

effective profile, terrain-vehicle relationship, micro- and macro-obstacles, purpose-designed test field, terrain condition, test bench

1. Introduction

There are several accelerated fatigue testing methods for vehicles travelling on terrain, but compared with long-term operational tests under real conditions (Kiss, 2004), bench durability tests – even at identical running times – are frequently found to result in different failure modes. The reasons for the discrepancies have not yet been established. It may be concluded from a survey of published research that no method of testing has been devised which is capable of modelling real terrain and real soil (Kiss, Laib, 2007).
Many test systems have been designed to model extreme terrain conditions. These are aimed at gathering the most possible information, in the least possible time, on vehicle faults associated with motion across terrain. Analysis of the faults permits modifications to the structure, and the process is repeated until the required performance is achieved under the test conditions.

2. Objectives

The purpose of the present research is to model the ground–vehicle relationship. Specifically, the aim is to describe the activating forces (Komándi, 1966) to which the vehicle is subject by consequence of the terrain profile, i.e. micro- and macro obstacles, and the inhomogeneity and mechanical properties of the soil. The mathematical model will permit the determination of loads arising from the ground-vehicle relationship in various terrain conditions. The model, by means of a purpose-built test bench, will permit a towed vehicle to be subjected to loads simulating any of the extreme terrain and soil conditions which the vehicle might encounter in reality. The model may be used to address all problems involving the relationship between the soil and the towed vehicle.
Existing terrain modelling procedures are restricted to a very small number of profile variations and lead to a narrow range of outcomes. Since the failures during the tests arise only from the particular terrain modelled, modifications lead to a construction which withstands only the test conditions. Real terrain conditions generate a much wider range of activation forces, and machinery developed with previous test methods does not necessarily fit its intended purpose. A satisfactory model of terrain profile must incorporate the unlimited variations of activating frequencies which occur in reality.

3. Comparison of field durability test and rolling test bench

A comparative durability test was performed on a towed vehicle (header trailer). It involved durability tests on two identical vehicles, one running under real conditions, and the other on a rolling test bench. The two methods led to different failure modes. Under real conditions, large amplitude swings caused the lampholders at the end of the long frame structure to break off. The much more intensive rolling test bench test caused cracks at the king pin and the rear axle welds. In both cases, the structure had to be strengthened.
The comparative tests demonstrate that the rolling test bench does not simulate the forces occurring under real conditions.

4. Previous test methods and their limitations

There are four different procedures most commonly used in the testing of towed vehicles.
The field durability test (I) involves running under real terrain conditions for a period equal to the lifetime. This method gives a good approach to real loads on the vehicle structure, but is very time-consuming and is difficult to extend to all possible terrain conditions

Figure 1. Field durability test

The next method involves a purpose-designed test field (II). This has the advantage of shortening of the time to failure. An artificial obstacle field, however, can only include a very small number of terrain obstacles, is difficult to convert, and carries the great disadvantage of damaging the tractive structure together with the towed vehicle.

The most commonly-employed test method uses a rolling test bench (III). The towed device is fixed to the end of a drawbar and it wheels are set into rotation by eccentric rollers. The eccentricity and the rotation induces vibration in the structure, thus modelling vibrational accelerations from field obstacles. This is a much lower cost test and leads to failure in a short time. Further advantages are that the amplitude and intensity of vibrations can be altered by adjusting the eccentricity and speed of rotation, and that it permits comparative tests, because the conditions can be repeated. The method has several drawbacks, however. The most severe is that it does not model real conditions. It is not capable of simulating many of the activating forces to which vehicles (Fekete, 1978) moving on terrain are exposed. A vehicle traverses combinations of micro- and macro-obstacles on real terrain, and is therefore subject to activating forces which vary suddenly, dynamically and within wide limits. The eccentricity on a rolling test bench is much smaller in scale than the natural variation of terrain profile, and so only to a very small extent models the dynamic effects of gravitational forces acting on the axles when a vehicle traverses large terrain obstacles. The real activating forces are closer to impulse loads than constant-intensity and constant-amplitude vibrations. A further major defect of the rolling test bench is that it does not simulate the large dynamic forces generated by braking and acceleration.

The fourth method is the shaker rig (IV), on which the vehicle is set into vibration by computer-controlled hydraulic activators. This is the method used to test off-road vehicles, because it best models the forces acting on a vehicle traversing terrain. It is very costly, however, and the outcome very much depends on how well the program simulates the terrain conditions. To be able to

test devices towed on terrain, the forces arising from the terrain profile and the traction characteristics would have to be determined and modelled.

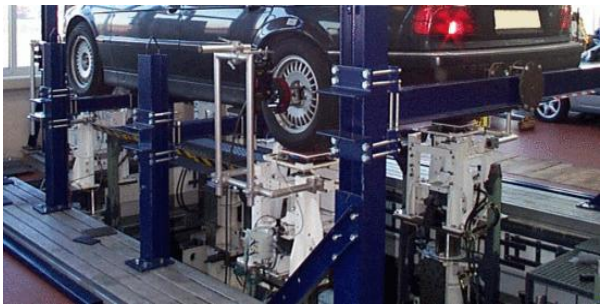


Figure 2. A test bench in operation

5. Micro- and macro-obstacles

Obstacles are categorised by height: macro-obstacles are larger than 250 mm, and micro-obstacles are smaller than this. Macro-obstacles divide into step-, trench- and barrier obstacles. The interaction of vehicles with such obstacles is the subject of terrain-vehicle interaction theory. The first and most important criterion for determining the mobility of a vehicle is whether the geometrical design of the vehicle permits it to cross the obstacle. Important design criteria for the test field are that the towed structure crosses the terrain obstacle without sliding off or falling over, and does not collide in any direction. The vehicle stoppage condition for a terrain obstacle may be described by a track curve obtained by approximation via the VSE function. To calculate the accelerations generated as the vehicle traverses terrain, a differential equation of dynamic balance may be written for each degree of freedom of the vehicle model.

$$\overline{M}\ddot{y} + \overline{K}\dot{y} + \overline{S}y = \overline{sx}(t) + \overline{k\dot{y}}(t) \tag{1}$$

The lifetime of the vehicle is to a large extent determined by the magnitude of the horizontal vibration accelerations arising from the terrain profile.
Vibrational accelerations are set up by macro-obstacles corresponding to the terrain relief and by micro-obstacles – those smaller than 250 mm – covering the surface of the soil. Investigation of the micro-obstacles and their effect on the structure is a much more complex task, because the corresponding profile types can only be described by statistical functions. The mathematical relations describing surface unevenness of the terrain are called stochastically distributed processes. Stochastic processes are a physically related group of events or forms with identical statistical parameters. To ensure that the model as far as possible approximates real conditions, the terrain profile curves describing the micro-

obstacles should be recorded on the type of terrain used by the vehicles. Test fields should be constructed with micro-obstacles whose distribution of size and density corresponds to real conditions.
Different soil types have different mechanical properties. These properties are to a large degree influenced by the inhomogeneity and moisture content of the soil and the ambient temperature The inhomogeneity derives from the mixture of different soil types, vegetation and other foreign matter. Vehicles traversing the terrain often contact the surface vegetation rather than the soil itself, and this also has differing mechanical properties. The inhomogeneity thus considerably affects the soil's mechanical properties. The adhesion of the soil depends on its moisture contact, which is expressed by the moisture uptake ability of the soil, pF. The temperature also influences the state of the soil and thus indirectly the moisture content. All of these factors greatly influence the mobility, and hence the lifetime, of a vehicle moving on the terrain. They must be taken into consideration in the construction and use of test fields, especially for comparative tests.
The influence of the soil and the terrain conditions on the vehicle structure depends on a great many factors. Models have been produced for wheels rolling on various soils, but at present there is no model which properly describes the relationship between the flexible soil and the flexible wheel.

6. Vehicles to be tested (header trailers)

Description of forces acting on towed vehicles requires more than theoretical equations of motion and activating functions. There are some forces generated by terrain conditions which can only be described by measurement. These have to be recorded via measurements on a suitable towed vehicle – in this case, a header trailer. Any vehicle, however, may only be tested on terrain conditions for which it was designed, i.e. those it will traverse when in use.

Figure 3. Header trailer

The running-wheeled towed structure (header trailer) is subjected to a variety of forces as it moves across the ground. Braking and acceleration exert forces on the vehicle via the drawbar. The header trailer's own braking system sets up braking forces at its wheels. The weight and the inertial force act at the centre of mass of the structure, as does the centrifugal force during turning. There is a force arise originating from rolling resistance in the vehicle's ground-vehicle relationship and dynamic forces originating from soil unevenness (micro- and macro-obstacles). The magnitude and intensity of these forces influence the vehicle's lifetime. Those which can only be determined by measurement are the dynamic forces arising from braking and soil unevenness.
In order to determine the relationship between towed vehicles and the soil, the geometrical and mechanical parameters of the header trailer must be determined, because these influence the results of the measurements. The vehicle's weight alters the effective profile, and the traction speed affects the intensity of the activating forces. The effects of these parameters must be properly taken into account before drawing any general conclusions.

7. Test field

The prime consideration in the design of the test field is that it should approach the size and distribution of real terrain obstacles. In order to determine this distribution, terrain profiles (Lee, 2009) must be measured under real conditions. There is an unlimited number of possible combinations of soil and terrain conditions in nature, and so the number of variations must somehow be restricted. Firstly, obstacles which would be impassable by virtue of the trailers' geometry are excluded. Secondly, the most relevant terrain conditions are those which the vehicle is likely to encounter during use. It is also important to exclude terrain profile combinations which are known from experience to make only a small contribution to reduction of the vehicle's lifetime. Another design consideration for the test field is the need for a higher frequency of terrain obstacles than that measured in real conditions, owing to the geometrical dimensions of the test field and the limitations of time available for testing.
There are several other circumstances to be taken into account in the design of the test field. The mechanical properties, vegetation and moisture content of the test field soil must not differ substantially from the real conditions. Then there are the damage which macro- and micro-obstacles may cause to the tractive vehicle and the need to prevent health hazards to the person driving it.
Given these boundary conditions, the optimum form is the "merry go round" test field. This configuration permits the tractive vehicle to travel on a different path than the towed vehicle and therefore be spared many of the stresses acting on the towed vehicle.

The vibrational accelerations induced by the surface profile are measured as the trailer traverses the test field. Measurements are made over pre-determined distance. Any failures are recorded for purposes of comparison.

Figure 4. Merry-go-round test field

8. Construction of the model

The mathematical description of the external forces arising as the towed vehicle moves across terrain comprises the dynamic relationships presented above and the measured forces ascribable to terrain unevenness and traction. The measured and calculated forces are combined into a complex mathematical model for the towed vehicle. Activations functions are then derived from the model and used to simulate, on a purpose-built test bench, the forces which act on the vehicle under real conditions. The magnitude and intensity of the forces depend on the velocity of travel and weight of the vehicle, the accelerations occurring during operation, the geometry of the terrain and the vehicle and the mechanical properties of the soil type.

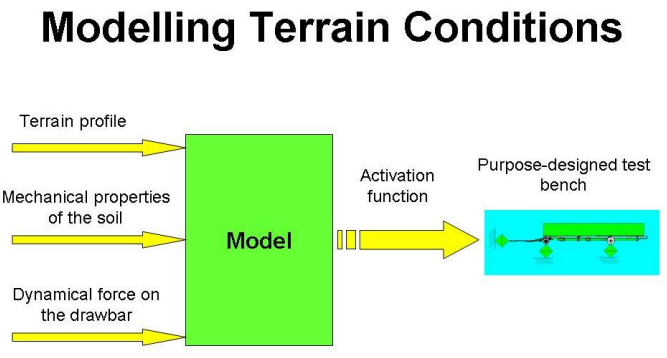


Figure 5. Modelling terrain conditions

The activation forces given by the model must meet certain boundary conditions. First of all, the maximum force acting on the structure may not exceed the force required for the immediate failure of the vehicle, i.e. no terrain profile may be formed which the vehicle cannot traverse without failure. Neither may the intensity of the activating forces exceed the critical sudden-failure value. Another boundary condition is that the minimum activating force must be greater than that which the tyre absorbs through its damping properties, because such activating forces have a negligible effect on vehicle lifetime.
By building a test bench to simulate an arbitrary combination of terrain conditions, it will be possible to generate activating forces equivalent to those occurring under real conditions. The test bench will permit different vehicle constructions to be compared.
The goodness of the model is most effectively determined by a comparative test. Under identical conditions, two identical towed vehicles should display the same failure mode. The comparative test will involve a durability test of one vehicle in the field, and of another on the purpose built test bench. If similar faults occur in both tests, the model may be pronounced satisfactory.

9. Expected results and conclusion

- A mathematical model which describes forces acting on vehicles traversing terrain, and which permits forces equivalent to real forces to be modelled.
- The ability to test the effects of various terrain conditions on the vehicle structure without the use of test fields.
- The possibility of simulating terrain conditions which in reality occur only under extreme conditions, enabling greater safety to be designed into vehicle structures.
- Incorporation into the model of how the damping properties of various soil types compare to the properties of hard road surfaces.
- The characterisation, via the model, of the mechanical processes which occur as a pneumatic tyre traverses the soil, thus providing better insight into the tyre-soil interaction.
- Applications in comparative testing, since the simulation may be repeated at any time.
- The determination of the effects of different soil types and terrain profiles on vehicle structures.
- More precise definition of several problems for which the forces acting on vehicles traversing terrain have hitherto only been approximated.
- Production of extreme terrain effects independently of environmental effects on the vehicle.
- Validatability. Every test may be repeated at any time, implying applicability to proving procedures.
- Considerable reduction in the duration and cost of vehicle testing.

Nomenclature

$\overline{\overline{M}}$	mass matrix
$\overline{\overline{K}}$	damping matrix
$\overline{\overline{S}}$	stiffness matrix
$\overline{y}$	displacement vector
$\overline{x}$	activating vector

References

Fekete, A. (1978): Some effects of vehicle speed on soil compactions. Proc. 6th Int. Conf. ISTVS, Vienna, Austria Vol. III. pp. 1021-1033.

Komándi, Gy. (1966) Establishment of physical soil parameters on the basis of the relationship between the shear diagram and tractive force characteristics. Landtechnische Forschung 4.

Kiss, P. Laib, L. (2007) Soil deformation and rolling resistance in the soil-tire interaction, proc. of the joint north America, Asia-Pacific ISTVS Conf. and annual meeting of Japanese society for terramechanics, CD SPNo: 2007-52-0339, Fairbanks, Alaska.

Kiss, P. (2004) The tractor's performance balance under non-stationary conditions, book of abstracts AgEng 2004 Conference Engineering the Future, p. 352-353. ISBN 90-76019-258, D 2004/0277/4, Leuven, Belgium.

Lee, R. C. (2009) Stochastic terrain and soil modeling for off-road mobility studies. Blacksburg, Virginia.

3. INSTITUTE FOR ENVIRONMENTAL ENGINEERING SYSTEMS

PROFESSOR DR. ISTVÁN FARKAS
DIRECTOR OF THE INSTITUTE

Dear Reader,

The Institute for Environmental Engineering Systems consists of two departments as Department of Environmental and Building Engineering and Department of Physics and Process Control. There are several research activities going on in the Institute having interrelations in order to find new scientific solutions concerning to the environmental engineering requirements. It gives a strong opportunity for sustainable development and allowing common way of approach in the solution of different nature of problems.
Form the results of the current research activity the following topics are selected out to publish in the recent issues of the Journal:

Technical background of very high efficiency GaAs based solar cell preparation:

- This work represents the electronic and mechatronic control-system for molecular source of beam epitaxial equipment. The control of motors and temperature of molecular sources happens with a PLC-based control system. This work was conducted by an external expert carrying on his PhD in our Institute.

Characterizations based an experimental of two photovoltaic module technologies:

- This paper concerns on characterizations of two different photovoltaic module technologies, i.e. polycrystalline silicon and amorphous silicon. Further outcome is that the improvement of both PV modules can be studied and achieved. This work was conducted by an Indonesian PhD student of the Institute.

Surface waves at convection-diffusion processes through porous media:

- The partial differential equation describing simultaneous convection and diffusion through porous media is discussed. It is shown, that the Riccati-type ordinary differential equation playing a crucial role in some contemporary modeling procedures.

Analysis of energy behavior of buildings:

- During the building operation time one of the possible ways of energy-saving is reduction of heating during the time the building is not in use, while suspending the cooling. Simulations were carried on using different building structures considering thermal mass.

SURFACE WAVES AT CONVECTION-DIFFUSION PROCESSES THROUGH POROUS MEDIA

Csaba MÉSZÁROS[1], Klaus GOTTSCHALK[2], István FARKAS[1],
Bernadett GYARMATI[3], Ágnes BÁLINT[3]

[1] Department of Physics and Process Control, Institute for Environmental Engineering Systems
[2] Leibniz Institute Für Agrartechnik, ATB Potsdam Bornim, Germany,
[3] Department of Chemistry and Biochemistry, Szent István University

Abstract

The partial differential equation describing simultaneous convection and diffusion through porous media is discussed in detail, particularly from the point of view of its general nonlinear character. It is shown, that the Riccati-type ordinary differential equation, playing a crucial role in some contemporary modelling procedures of the simultaneous convection-diffusion problems may be treated in a refined manner, by taking into account the genuine dispersive character of the porous bulk.

Keywords

Convection, diffusion, Riccati-type ordinary differential equation, percolation

1. Introduction

The powerful methods of the non-equilibrium thermodynamics have always played and play even nowadays a role of continuously increasing importance at establishing of fluid mechanics and continuum mechanics in general [Verhás, 1997]. They also represent the most effective tools for accurate describing of crucial physical properties of dissipative structures (to which belong e.g. the Rayleigh-Bénard convection cells, different types of reaction-diffusion systems etc. [Argyris et al., 2010]), whose theoretical and experimental investigation is in the forefront of research activity in the field of complex systems in general.

Since mathematical modeling of the convection-diffusion processes (which also belong to the circle of hierarchically organized complex transport phenomena indicated above) plays a crucial role in mathematical modeling of many important engineering problems [Pascal, 1993], [Kirschner et al., 2004] elaboration of novel and more accurate models of this problem represents a permanent task, whose complexity is reflected in the nonlinear character of the

ordinary differential equations (ODEs) and partial differential equations (PDEs) to be solved. The basic PDE is

$$\frac{\partial c}{\partial t} - \nabla \cdot \left(D(c) \nabla c \right) - \frac{dK}{dc} \cdot \frac{\partial c}{\partial z} = 0, \tag{1}$$

where $c = c(\vec{r}, t)$ denotes the concentration distribution function to be determined, $D = D(c, T, ...)$ is the (usually) thermodynamic state-dependent diffusion coefficient and $K = K(c)$ is the concentration-dependent hydraulic conductivity coefficient while z-axis corresponds to the direction of the gravitational acceleration. The solution of the nonlinear PDEs of type (1) may be even of solitonic character and requires application of the most effective calculation methods developed within framework of the Painlevé theory giving the most general description of such phenomena [Conte and Musette, 2008]. As it is well-known, solitons may exist due to balance between dispersion effects, which try to expand the initial localized wave packet, and nonlinearities trying to localize it. By careful analyses of these (and similar) opposite effects Fan [Fan, 2000] proposed a new general method for effective solving of nonlinear PDEs, which are of importance for the whole nonlinear physics. The essence of the powerful method applied in this study can be explained as follows: the linear term of highest order must be balanced with the nonlinear terms in the initial PDE to be solved. Therefore, the travelling-wave form of the general solution of the convection-diffusion problem is obvious. Then, using d'Alembert-type independent variables, i.e. $c = c(\zeta = x + y - \lambda t); \lambda = const.$, the general form of the solution will be [Fan, 2000], [Elwakil et al., 2004]:

$$c(\zeta) = a_0 + \sum_{i=1}^{q} \left(a_i \omega^i + b_i \omega^{-i} \right), \omega = \omega(\zeta) \tag{2}$$

$$a_0, a_i, b_i = const., (1 \le i \le q), \tag{3}$$

where $q \in \mathbb{N}^+$ and the component solution functions obey the Riccati-type ODE $\frac{d\omega}{d\zeta} = k + \omega^2$, with „$k$" as a parameter depending on the actual experimental conditions and to be determined subsequently.

2. Relevance of the Riccati-type ordinary differential equations at modeling of the convection-diffusion processes

In the case of the simplest type of simultaneous convection-diffusion (in agreement with the equ. (1)) we will have:

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} + v\frac{\partial c}{\partial x}, \tag{4}$$

where $\vec{v}$ denotes the macroscopic fluid flow velocity, which is frequently assumed to be of constant value at modelling of the convection-diffusion problems. It is well-known (e.g. [Eck et al., 2008]), that the simultaneous convection-diffusion processes result in appearance of travelling waves of concentration changes in fluids undergoing macroscopic flow. It is also known, the elastic surface waves (usually called Rayleigh-type waves [Landau et al., 1999]) may be represented by solutions of the wave equation:

$$\frac{\partial^2 u}{\partial t^2} - c^2\nabla^2 u = 0, \tag{5}$$

where *u* now denotes a component of the longitudinal or transversal deformation vector and *c* is the relevant longitudinal or transversal velocity. If we use a rectangular coordinate-system and assume, that the continuous elastic medium fills the infinite half-space z < 0, the travelling waves propagating into x-direction can be given by the expression

$$u = e^{i(kx-\omega t)} f(z), \tag{6}$$

whose insertion into (5) gives immediately:

$$\frac{d^2 f}{dz^2} = \left(k^2 - \frac{\omega^2}{c^2}\right) f. \tag{7}$$

Another well-known example from continuum mechanics is related [Landau and Lifshitz, 2000] to propagation of gravity waves on the free fluid surfaces. They form is (the meaning of symbols is analogous to those used at the calculation procedure leading to (7)):

$$\frac{d^2 f}{dz^2} - k^2 f = 0, \tag{8}$$

i.e. both (7) and (8) are de facto Riccati-type ODEs in linearized form (see the first part of the Appendix in this respect).
The dispersion phenomenon can be very naturally assigned to the convection-diffusion equation [Eck et al., 2008], which, as it has been presented previously is of the form (4). Even in linear approximation discussed briefly in the introductory part, this equation is able to reflect the most essential features of travelling waves with dispersive character. Namely. if we use the usual solution form of travelling waves $u(x,t)=u_0\cdot e^{i(kx-\omega t)}$ (u_0 denotes the maximal value of the oscillating vector component) its direct substitution into (4) gives directly:

$$\omega = vk - iD|k|^2. \tag{9}$$

Then, following some of the newest developments in the general theory of diffusion-convection processes e.g. [Fan, 2000], we may state, that extensive application of the Riccati ODE has its roots in the above cited ([Landau et al., 1999], [Landau and Lifshitz, 2000]) classic modeling methods leading to accurate descriptions of the Rayleigh-type surface waves. According to our knowledge, this aspect has not been emphasized in the literature. The above described simple example may be directly generalized to the case of simultaneous convection and diffusion of several components and the Lagrangian representation of the continuum mechanics has proven to be particularly suitable from this point of view [Mészáros and Bálint, 2011].

3. A novel-type application of the Riccati's ordinary differential equation

In accordance with the relations (A,1) and (A,3) from the Appendix, we are also going to use here the linearized variant of the Riccati ODE by specifying $h(x)=\sum_n a_n(p)\cdot x^n$, where the coefficients are of stochastic character, i.e. in the relation $a_n = a_n(p)$ the independent variable *„p"* will be defined as the actual percolation probability. Firstly, let us assume, that there is only one „dominant term" in this expression, and according to the novel solution method we propose in this paper, the Riccati equation will be redefined in the form of:

$$y' - y^2 - a_n(p)\cdot x^n = 0. \tag{10}$$

In this case, the solution can be written (again by use of the MAPLE computer algebra system) as:

$$y(x)=\frac{C\left[J\left(\frac{n+3}{n+2},\frac{2\sqrt{a_n(p)}x^{\frac{n}{2}+1}}{n+2}\right)\sqrt{a_n(p)}x^{\frac{n}{2}+1}-J\left(\frac{1}{n+2},\frac{2\sqrt{a_n(p)}x^{\frac{n}{2}+1}}{n+2}\right)\right]+\ldots}{x\left[CJ\left(\frac{1}{n+2},\frac{2\sqrt{a_n(p)}\cdot x^{\frac{n}{2}+1}}{n+2}\right)+Y\left(\frac{1}{n+2},\frac{2\sqrt{a_n(p)}\cdot x^{\frac{n}{2}+1}}{n+2}\right)\right]}$$
$$\frac{\ldots+\left[Y\left(\frac{n+3}{n+2},\frac{2\sqrt{a_n(p)}\cdot x^{\frac{n}{2}+1}}{n+2}\right)\sqrt{a_n(p)}\cdot x^{\frac{n}{2}+1}-Y\left(\frac{1}{n+2},\frac{2\sqrt{a_n(p)}\cdot x^{\frac{n}{2}+1}}{n+2}\right)\right]}{x\left[CJ\left(\frac{1}{n+2},\frac{2\sqrt{a_n(p)}\cdot x^{\frac{n}{2}+1}}{n+2}\right)+Y\left(\frac{1}{n+2},\frac{2\sqrt{a_n(p)}\cdot x^{\frac{n}{2}+1}}{n+2}\right)\right]},\tag{11}$$

where *C* is an integration constant, while J(k,x) and Y(k,x) denote the Bessel-functions of k-th order, which are of the first kind, and second kind respectively. It is obvious, that the basic invariance property of the Riccati ODE explained by relations (A,5 - 9) is reflected in the form of the solution (11), too. Similarly, if we perform linearization of the ODE (10), we have to solve:

$$\frac{d^2u}{dx^2}-a_n(p)\cdot x^n u(x)=0.\tag{12}$$

The solution of this ODE can be obtained directly (again, we used the MAPLE 10 system), and its final form is more concise compared to (12):

$$u(x)=C_1\sqrt{x}\cdot J\left(\frac{1}{n+2},\frac{2\sqrt{-a_n(p)}\cdot x^{\frac{n}{2}+1}}{n+2}\right)+C_2\sqrt{x}\cdot Y\left(\frac{1}{n+2},\frac{2\sqrt{-a_n(p)}\cdot x^{\frac{n}{2}+1}}{n+2}\right),C_1,C_2=const.,\tag{13}$$

i.e. the final result is explained again by Bessel-functions of the first and second kind, respectively, and of the same order. The effective treatment of the general convection-diffusion problem by using $h(x)=\sum_n a_n(p)\cdot x^n$ instead of the single homogeneous term in (10) is then straightforward.

4. Conclusions

It is shown, that the surface waves appearing at convection-diffusion processes can be described in a much more refined manner by use of Bessel-type special

functions compared to the earlier descriptions. Moreover, this modelling method could also be incorporated into general formalisms of the „wandering moisture level zones" at such kinds of drying processes.

Appendix - Some basic invariance properties of the Riccati-type ordinary differential equations

(1.) Linearization

The Riccati-type ODE as a typical nonlinear differential equation, whose general forms is $y' = f(x)y^2 + g(x)y + h(x)$, by the integral transformation $u(x) = e^{-\int f(x)y(x)}dx$ can be directly prescribed as

$$f(x)u'' - \left(f'(x) + f(x)g(x)\right)u'(x) + f^2(x)h(x) = 0, \tag{A,1}$$

which is already a higher order, but linear ODE. Particularly, for $f(x) \equiv 1$ (accompanied by $g(x) \neq 0$ and $g(x) \equiv 0$ respectively) we will have:

$$u'' - g(x)u' + h(x) = 0, \tag{A,2}$$

$$u'' + h(x) = 0. \tag{A,3}$$

(2.) Invariance with respect to the homography group transformations

As it is well-known, in general case the Riccati's ODE (which we use in the present sub-section in the form of $y' + f(x)y^2 + g(x)y + h(x) = 0$) can not be solved by the method of quadratures. In the case, when we know a particular solution ($y_1(x)$) of its, it is possible to represent the general solution as

$$y(x) = y_1(x) + \frac{1}{u(x)}, \tag{A,4}$$

where $u = u(x)$ denotes a general solution of a Bernoulli-type ODE (to which the initial Riccati - type ODE is reduced via use of the particular solution $y_1(x)$).

Therefore, knowing the particular solution $y_1(x)$, the complete solution of the Riccati's ODE may be written e.g. [Conte and Musette, 2008] as

$$y(x) = y_1(x) + \frac{1}{u(x)} = \frac{CF(x) + G(x)}{CH(x) + K(x)}, C = const. \tag{A,5}$$

Since the general solution of the latter can be derived by two quadratures, the above-given simple solution form (A,5) is particularly suitable for detailed studies of the basic invariance properties of the solutions of the Riccati-type ODEs [Mitrinović, 1939]. Namely, eliminating the constant *C* from this equation and the one obtained by its derivation, we arrive again at a Riccati-type ODE:

$$\begin{vmatrix} F & G \\ H & K \end{vmatrix} y' = \begin{vmatrix} F & G \\ F' & G' \end{vmatrix} + \left\{ \begin{vmatrix} G & H \\ G' & H' \end{vmatrix} + \begin{vmatrix} K & F \\ K' & F' \end{vmatrix} \right\} y + \begin{vmatrix} H & K \\ H' & K' \end{vmatrix} y^2. \tag{A,6}$$

Then, the transformation of the unknown function

$$y(x) = \frac{Q(x)z(x) + R(x)}{S(x)z(x) + T(x)}, (QT - RS \neq 0), \tag{A,7}$$

(defining the Möbius group, widely applied at conformal mappings e.g. [Marvel et al., 2009]) leads again to a Riccati-type ODE in the form of:

$$z'(x) + F(x)z^2(x) + G(x)z(x) + H(x) = 0, \tag{A,8}$$

where:

$$F = \frac{Q'S - QS' + fQ^2 + gSQ + hS^2}{QT - RS}, H = \frac{R'T - RT' + f\,R^2 + gTR + hT^2}{QT - RS},$$
$$G = \frac{R'S - RS' + Q'T - QT' + 2f\,RQ + gSR + gTQ + 2hST}{QT - RS}. \tag{A,9}$$

Acknowledgement

Research was supported/subsidized by the TÁMOP-4.2.2.B-10/1-2010-0011 „Development of a complex educational assistance/support system for talented students and prospective researchers at the Szent István University" project.

References

Argyris, J., Faust, G., Haase, M., Friedrich, R. (2008): Die Erforschung des Chaos (Eine Einführung in die Theorie nichtlinearer Systeme) 2. Aufl., Springer-Verlag: Berlin-Heidelberg.

Conte, R., Musette, M. (2008): The Painlevé Handbook. Springer Science + Business Media B.V.: Dordrecht, The Netherlands and Bristol, UK: Canopus Publishing Limited. XXIV, pp. 256.

Eck Ch., Garcke H., Knabner P. (2008): Mathematische Modellierung, Springer-Verlag: Berlin-Heidelberg.

Elwakil, S. A. El-Labany, S. K., Zahran, M. A., Sabry, R. (2004) Exact traveling wave solutions for a diffusion-convection equation in two and three spatial dimensions. Computer Physics Communications, Vol. 158, pp. 113 – 116.

Fan, E. (2000): Extended tanh-function method and its applications to nonlinear equations. Physics Letters A, Vol. 277, pp. 212-218.

Kirschner, I., Mészáros, Cs., Bálint, Á., Gottschalk, K. and Farkas, I. (2004): Surface changes of temperature and matter due to coupled transport processes through porous media, Journal of Physics A: Mathematical and General, Vol. 37, pp. 1193-1202.

Landau, L. D., Lifshitz, E. M. (2000): Fluid Mechanics, 2nd Ed. Oxford-Boston-Johannesburg-Melbourne-New Delhi-Singapore: Butterworth-Heinemann.

Landau, L. D., Lifshitz, E. M. (1999): Theory of Elasticity, 3rd Ed. (by E. M. Lifshitz, A. M. Kosevich, L. P. Pitaevskii) Oxford-Boston-Johannesburg-Melbourne-New Delhi-Singapore: Butterworth-Heinemann.

MAPLE 10 (2005): A Symbolic Computation System; Waterloo Maple Inc.

Marvel, S. A., Mirollo, R. E., Strogatz, S. H. (2009): Identical Phase Oscillators with global Coupling evolve by Möbius Group Action. Chaos. Vol. 19. pp. 043104 – 1-11.

Mészáros, Cs., Bálint, Á. (2011): Transient transport processes in deformable porous media, Chinese Physics B 20 110507.

Mitrinović, D. S. (1939): Théorème sur l'équation de Riccati. C.R. Acad. Sci. Paris, Vol. 208, pp. 156-157.

Pascal, H. (1993): On a nonlinear convection-diffusion equation. Physica A, Vol. 192, pp. 562–568.

Verhás, J. (1997): Thermodynamics and Rheology (Fluid Mechanics and its Applications, Vol. 38.), Budapest, Akadémiai Kiadó and Dordrecht: Kluwer Academic Publishers.

USE OF CONTINUOUS AND DISCONTINUOUS ENERGY MODEL FOR OFFICE BUILDING

Attila KEREKES[1], Györgyné HALÁSZ[2]

1 Dryvit Profi Bt.
[2] Department of Environmental and Building Engineering, Institute for Environmental Systems

Abstract

During the building operation time one of the possible ways of energy-saving is reduction of heating during the time the building is not in use, while suspending the cooling. We are examining an office building model using a TRNSYS program, determining how the Annual energy consumption is formed based on different operating schedules, inner heat loads and ventilation conditions. We also ran the simulation using different building structures considering thermal mass.

Keywords

Discontinuous heating, thermal mass, energy-saving, thermal model

1. Introduction

One of the most important questions nowadays is possibilities of reducing energy consumption while increasing the energy efficiency. One of the possibilities for energy-saving in existing buildings lies using discontinuous operation times. The comfort parameters which are required for the building users are guaranteed only during the operating time of the building. The building heating, cooling and mechanical ventilation systems are switched off for the longest time possible. Because of the building's thermal mass, all heat emitting equipment is to be switched on before the operating time, in order to ensure that required comfort parameters will be provided. Because of the conservation reasons during the heating period they usually limit the cooling down by defining the lowest inner temperature allowed (Fekete I., 1985).
In the case of the given building it is important for the operator to know which various use modes will require which amount of energy. During the designing process it can be important to know the structure of the building because of the impact of the building structure on the energy efficiency.
Regarding the above mentioned questions, we have created a model of an existing building using TRNSYS in two versions: one with a small thermal mass construction and one with a big thermal mass construction (Petrus Tri Bhaskoro, 2011). We have created different Annual simulations varying operational modes.

2. Description of the model

2.1. Building, building structures

The building is a two level, flat roofed, free-standing office building. Building's useful floor area is 713,90m^2. The building is located in m above the Debrecen. Ground floor is located at +0,17 m and first floor at +3,77m above the pavement level The building's longitudinal axis forms a 45° angle with the direction of North. Offices, meeting rooms, tea kitchens, sanitary facilities, machinery room and a corridor are located on the floor, while the server room can be found on the ground floor level. The floor level and the ground level are connected by a staircase. Figure 1 shows the overview layout plan of the building.

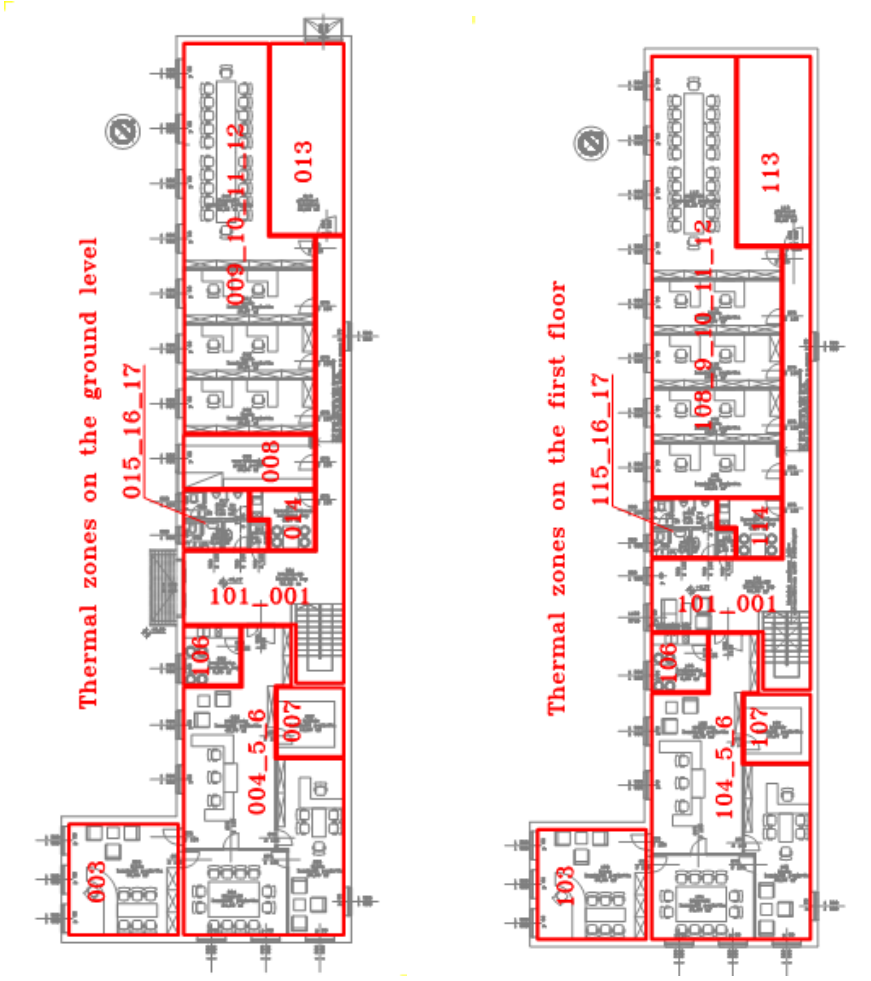


Figure 1. The building overview layout plan

The characterictics of the bordering materials are summarized in the following Table:

Table 1. Characteristics of the bordering structures

Structure	Name	Type	Overall heat transfer factor [W/m^2K]	Specific thermal mass [kg/m^2]
The small specific mass version (Building with light structure)				
LS	**External wall _Light-structure**	External wall	0,26	32,00
LS	**Holding wall _ Light-structure**	Internal wall (between heated areas)	0,16	32,00/ 32,00
LS	**Bulkhead_ Light-structure**	Internal wall (between heated areas)	0,54	32,00/ 32,00
LS	**Slab_ceramics**	Inner slab	0,19	32,00/35,00
LS	**Slab_ laminate parquet**	Inner slab	0,19	32,00/11,00
LS	**Roof (final ceiling)**	Roof	0,21	32,00
The big specific mass version (Ferro-concrete wall structure with continuous exterior side insulation)				
FC	**External wall_FC**	External wall	0,26	514,00
FC	**Holding wall_FC**	Internal wall (between heated areas)	2,08	385,00/ 385,00
FC	**Bulkhead_FC (=PTH)**	Internal wall (between heated areas)	1,70	68,00/ 68,00
FC	**Slab_ceramics**	Inner slab	0,75	146,00
FC	**Slab_ laminate parquet**	Inner slab	0,73	132,00
FC	**Roof (final ceiling)**	Roof	0,21	305,00
Both structure versions of the same building				
Identical	**Window**	Window (external, wood and PVC)	1,60	
Identical	**Door**	Glazed door (external, steel)	1,80	
Identical	**Internal door**	Door (interior, to an unheated space)	1,80	
Identical	**Ceramic floor**	Floor (laid on soil)	0,85	146,00
Identical	**Floor laminate parquet**	Floor (laid on soil)	0,85	132,00

In both structure versions the premises' inner dimensions are identical, we are not taking into consideration the differences in wall thicknesses. The windows was built with glazing's solar factor of 0,6 and the inner side shading's solar factor of 0,55. We were not taking into consideration differences in linear overall heat transfer factor which exist because of the structure differences, although they exist in reality. This way, the bordering structures in both versions of the building structures will have identical heat loss factor.

2.2 Thermal zones

As the Figure 1. shows, building was divided into 18 zones. Each zone consists of the rooms of the same expected thermal behavior (the same orientation, the same function, the same required internal air conditioning parameters, etc.). Neighboring premises are in thermal connections with each other. The internal walls within the zones on the floor level have also been defined because they must be counted with during the calculation because they play a role in the heat storage process.

2.3 Mechanical equipment

Heating and cooling of the premises are provided by using individually regulated convector units. Heating and air conditioning units capacities are determined by MSZ 04-140/3 and MSZ 04-140/4 standards (heat demands and cooling heat loads).

The ventilation of the premises can be either by infiltration, based on defined number of air changes, or mechanical ventilation based on defined fresh air flow rate. In the case of mechanical ventilation temperature of the supply air during the heating season is 22°C=const., otherwise it is 24°C=const. Mechanical ventilation is installed with a recuperator, with defined heat transfer efficiency. The mechanical installation provides the fresh air flow rate according as defined by the MSZ CR 1752:2000 standard.

The heating season covers 0-2520 and 6239-8760 hours periods during the year.

2.4. Use modes

The use of the building is described by the following parameters.

The required inner temperatures values during the use period and the values out of the operating time (reduced) are determined according to the following

Defined temperatures:

Temperatures during the heating period:

Machinery rooms:		15°C
Archives:	operating time	19°C,
	non - operating time	15°C
Other premises:	operating time	22°C,
	non - operating time	15°C

Temperatures during the cooling period:

Machinery rooms:		-- (no cooling)
Sanitary rooms:		-- (no cooling)
Server:		24°C
Other premises:	operating time	24°C,
	non - operating time:	-- (no cooling).

In the case of discontinuous heating, the whole heating system in the model restarts automatically depending on the outer temperature and the time elapsed

since the switch off. We have determined the control parameters in such a way that at the start of the operating time the zone temperatures usually reach the defined values.
The inner heat loads originating from men and computers can take maximal values during the operating time, and equal zero out of the operating time.
The mechanical ventilation can take maximal values during the operating time, and are always zero out of the operating time. While the mechanical ventilation is in progress there is no infiltration.
There are two options regarding the operating time: the first is continuous operation 0-24h (seven days a week), the other is discontinuous operation – 12h use per working day (7-19h during the day) and no use during the weekends.

3. Simulations

When creating the models we were running a Annual simulation, with meteorological data provided by the TRNSYS program for Debrecen, Hungary. We are one by one taking into consideration the specified use parameters (heat load, mechanical ventilation) during individual simulations as the Table 2. shows.

Table 2. Simulation combinations

Sequence number	Heat load		Mechanical ventilation		Operating time	
	0	100%	0	100%	Continuos 0-24h	Discontinuous (12h per working day)
1.	x		x		x	
2.	x		x			x
3.		x	x		x	
4.		x	x			x
5.		x		x	x	
6.		x		x		x

We ran the simulation using the six combinations for big and small thermal mass buildings.
During the course of the simulations the TRNSYS program collected the data about the Annual heating and cooling energy consumption.
In order to make a comparison, we have converted the obtained data to the consumption of the primary energy. Heating coefficient of performance Ck=1,01, primary energy conversion factor 1,0 (condensation, natural gas fueled boiler). Cooling EER=3, electrical primary energy conversion factor=2,5.
These data are valid only in Hungary (see 7/2006. (V.24.) TNM regulation), but it is necessary to express all the energy uses through the use of the primary

energy (Zöld A., 2006). (In another country meteorological may data differ as well).

4. Results

We have summarized the results in the Table 3. and Table 4. The "Zone heating" as well as the "Zone cooling" include individually regulated heat emitters energy consumption. The "Ventilation heating" and the "Ventilation cooling" include ventilation energy consumption.

Table 3. Simulation results: Annual energy consumptions

Thermal mass	Heat load	Mech. Vent.	Op. time	Zone heating	Zone cooling	Ventilation heating	Ventilation cooling	Heating sum	Cooling sum
				[kWh/a]	[kWh/a]	[kWh/a]	[kWh/a]	[kWh/a]	[kWh/a]
Small thermal mass	0	0	con.	86 750	6 433			86 750	6 433
			dis.	65 180	4 340			65 180	4 340
	100%		con.	33 940	75 940			33 940	75 940
			dis.	43 090	35 260			43 090	35 260
		100%	con.	14 010	89 970	271 599	1 776	285 609	91 746
			dis.	16 960	45 920	82 545	1 102	99 505	47 022
Big thermal mass	0	0	con.	86 880	3 831			86 880	3 831
			dis.	64 050	2 962			64 050	2 962
	100%		con.	31 550	72 210			31 550	72 210
			dis.	40 560	31 310			40 560	31 310
		100%	con.	11 480	86 610	271 599	1 776	283 079	88 386
			dis.	14 630	43 160	82 545	1 102	97 175	44 262

Table 4. Annual primary energy consumption

Thermal mass	Heat load	Mechanical vent.	Operating time	Heating primary energy	Cooling primary energy	Sum primary energy	Achievable energy savings – discontinuous operating time
				[kWh/a]	[kWh/a]	[kWh/a]	[%]
Small thermal mass	0	0	con.	87 618	7 720	95 337	
			dis.	65 832	5 208	71 040	25,5%
	100%		con.	34 279	91 128	125 407	
			dis.	43 521	42 312	85 833	31,6%
		100%	con.	288 465	110 095	398 560	
			dis.	100 500	56 427	156 927	60,6%
Big thermal mass	0	0	con.	87 749	4 597	92 346	
			dis.	64 691	3 554	68 245	26,1%
	100%		con.	31 866	86 652	118 518	
			dis.	40 966	37 572	78 538	33,7%
		100%	con.	285 910	106 063	391 973	
			dis.	98 147	53 115	151 261	61,4%

Additionally: the inner heat load (people and computers heat emission) in the case of continuous operating time (0-24h) will be 153100 kWh/a, while in the case of discontinuous operating time will be 71630kWh/a.

5. Conclusions

Without taking the mechanical ventilation and heat loads into consideration, the discontinuous operation provides heating and cooling energy savings. In practice, that would represent an office building which is not crowded, in which there is no use of computers and has no mechanical ventilation system.
In a case of not having mechanical ventilation, but with internal heat loads we can see that the discontinuous operation is less favorable considering the heating energy use. That means two things. One, that the inner heat load heats the building during the operation time and therefore lowers the heat consumption. Out of operation there is also need for heating in order to achieve a temperature which is lower than prescribed, therefore that lower temperature has to be achieved through the use of the heating system. The previously described impact for the continuous use operates throughout the heating season, while if the operation is discontinuous it operates only during the operating time. The other thing, that the continuous operation is not the same as the case of a discontinuous operation without lowered heating. So, it doesn't mean that discontinuous operation with heating reduction will require more energy, but it means that the continuous operation doesn't provide a possibility to lower the heating, but that the inner heat load continuingly lowers the heating energy consumption. During the discontinuous operation we have lower the heating consumption, but out of the operation time we cannot count on the inner heat loads. In reality, that kind of building could be a regular crowded office building, with computers and cooling system, but still without mechanical ventilation.
The previously described impact can be also applied in those buildings which have mechanical ventilation and big amount of internal heat loads. The heating energy consumption is bigger in the case of discontinuous operation, but in this case we also have to take the mechanical ventilation energy consumption into consideration. The "Sum-up of heating" shows us that in this case the discontinuous operation will be more favorable. This tells us that the use of the ventilation heating system will consume much more energy than the individually controlled convective heat emitters. In practice this would represent an "A" category office building.
In any case the consumption of the primary energy is lower in the case of the discontinuous operation (calculated by the Hungarian primary energy conversion factors) (Fekete I., 1985).

The value of the primary energy saving will be increased if the amount of the inner heat loads increases and mechanical ventilation is in use. It also increases the consumption of the primary energy.
The building with a big thermal mass structure uses less energy in all the examined cases, and bigger energy savings can be achieved with discontinuous operation.
We would like to draw your attention to the following:

- The heat loss factors are identical for both building structures
- The TRNSYS is calculating with the solar energy heat gains and with the heat stored within the structure during the course of the simulation

6. Summary

The results of the simulations clearly show that the more detailed calculation of the effects of the inner heat loads, ventilation and the environmental effects can lead to different results than expected. For example: heating energy consumption in case of a building with big inner heat load, without mechanical ventilation and discontinuous operation can be bigger than in a case of a continuous operation. The further detailed model examination helps us to evaluate and interpret the results. A building with big inner heat loads and continuous operation is not the same as the case of a discontinuous operation without lowered heating, therefore the continuous operation is not necessarily equal with the use without heat lowering.
Different kind of operation modes of a building can be examined in details by thermal modeling. In a case of a given building the energy consumption can be determined in advance which can provide help when making technical, economical and business decisions.

Acknowledgement

Research was supported/subsidized by the TÁMOP-4.2.2.B-10/1-2010-0011 „Development of a complex educational assistance/support system for talented students and prospective researchers at the Szent István University" project.

References

Fekete, I.: (1985) Handbook of Building Physics, Technical Book Publisher, Budapest (in English)
Zöld, A. editor: (2006) The new Building Energy Regulations, Bausoft Pécsvárad Ltd., Pécs (in English)

Petrus Tri Bhaskoro and Syed Ihstham Ul Haq Gilani: (2011) Transient Cooling Load Characteristic of an Academic Building, using TRNSYS. Journal of Applied Sciences, 11: 1777-1783.

TRNSYS 16 a TRaNsient SYstem Simulation program, Volume 6: Multizone Building modeling with Type56 and TRN Build, Solar Energy Laboratory, University of Wisconsin-Madison 1500 Engineering Drive, 1303 Engineering Research BuildingMadison, WI 53706 – U.S.A.

CHARACTERIZATIONS BASED AN EXPERIMENTAL OF TWO PHOTOVOLTAIC MODULE TECHNOLOGIES

Dani RUSIRAWAN[1], István FARKAS[2]

[1]Department of Mechanical Engineering, Institut Teknologi Nasional (ITENAS) Bandung, West Java – Indonesia
[2]Department of Physics and Process Control, Institute for Environmental Engineering Systems

Abstract

This paper concerns on characterizes of availability (exergy) of two different of Photovoltaic (PV) modules technologies, i.e. polycrystalline silicon (included crystalline silicon technology) and amorphous silicon (included thin-film technology), under Gödöllő, Hungary climatic conditions. The PV modules placed on surface orientation position 5^o to East for facing to South (azimuth angle, γ) with a tilt angle 30^o (β). The actual operational data, such as solar irradiation (G), current (I), voltage (V) and electrical power (P) are used in order to characterizes both of above module technologies. To characterizes the exergy efficiency of PV modules, firstly the exergy of the total solar irradiation for considered locations is needed. Exergy losses, as the result of the irreversibilities in the process, also will be found in these characteristics. Further outcome from this works, the real possibility improvement of both PV modules (i.e. electrical efficiency) can be studied, observed and achieved.

Keywords

Exergy, polycrystalline technology, amorphous silicon technology, empirical data and electrical efficiency.

1. Introduction

As fossil fuel prices have risen and concerns over greenhouse gases (GhGs) and global climate change have increased, alternative technologies for producing electricity have received greater attention. Among the technologies that may help to address these concerns is solar photovoltaic (PV) systems, which capture solar radiation and convert it directly into electrical energy. In PV applications, solar cell/module is a basic building block to construct PV systems (panel/array) through its connection both in series or parallel, in order to obtain the desired main parameter such as current (I) or voltage (V). Currently, direct cost of solar PV system is widely acknowledged to be much greater than fossil fuel generation or many other renewable energy sources (Borenstein, 2008).

The output power of PV system (solar cells/modules/panel/array) can be affected by many factors, such as irradiance, temperature and material. The raw material of solar cells mainly can be classified into silicon and compounds, meanwhile the PV system technology can be categorized as thin-film and crystalline, as can be seen in Table 1.

Table 1. PV categories based on technology and material

	Technology	Material	Sub-material
PV	Wafer based crystalline silicon	Silicon	1. Mono/single crystalline silicon, sc-Si
			2. Poly/multi – crystalline silicon, mc-Si
			3. Heterojunction with intrinsic thin layer, HIT
	Thin film	Silicon	1. Amorphous silicon, a-Si
			2. Microcrystalline silicon, μc-Si
			3. Multi Junction (Tandem), example: a-Si/a-Si and micromorph silicon (a-Si/μc-Si)
		Compound	1. Copper - Indium - (Gallium) - Selenide, CI(G)S
			2. Cadmium telluride, CdTe

It is well known that most of the radiation (solar energy) absorbed by a PV system is not converted into electricity (electrical energy) but contributes also to increase the temperature of the module (thermal energy), thus reducing the electrical efficiency.

Conversion efficiency, defined as the ratio between the produced electrical power (electrical energy per second) and the amount of incident solar energy per second, is one of the main performance indicators of PV cells and modules. Table 2 provides the current efficiencies of different commercial PV modules (International energy agencies, 2010).

Table 2. Current efficiencies of different PV technology commercial modules

Wafer based c-Si		Thin film		
sc-Si	mc-Si	a-Si; a-Si/μc-Si	CdTe	CIS/CIGS
14-20%	13-15%	6-9%	9-11%	10-12%

Testing (conducted to an experimental data) and modelling efforts are typically to quantify and then to replicate the measured phenomenon of interest. Testing and modelling of PV system performance in the outdoor environment is very complicated and influenced by a variety of interactive factors related to the environment and solar cell physics. In order to effectively design, implement, and monitor the performance of photovoltaic systems, a performance model must be able to separate and quantify the influence of all significant factors. In view of this, it is now becoming essential to look for various aspects in order to increase the PV energy conversion into electricity on its application in the field.

State of the art about research and development in PV field in order to get the PV solar cells/modules with high efficiency and low cost can be seen in Figure 1.

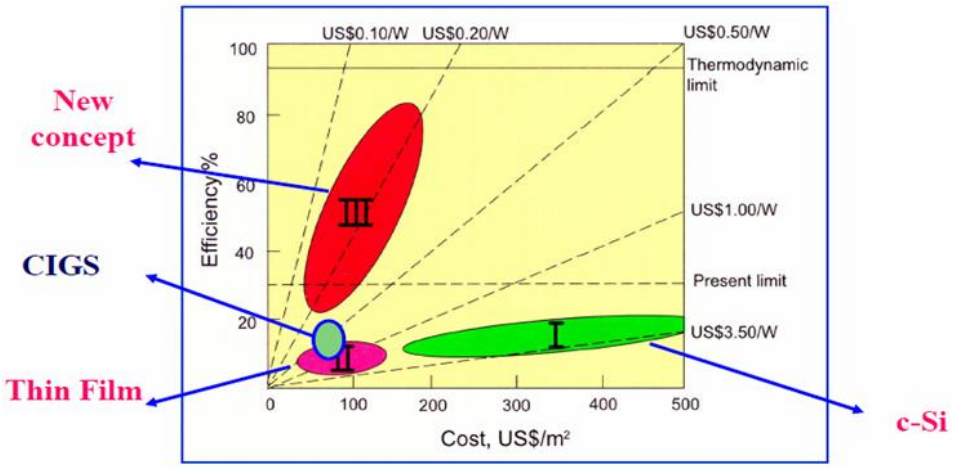


Figure 1. Cost-efficiency analysis for first-generation (I), second-generation (II), and third-generation (III) photovoltaic technologies

The so-called "third-generation PV" will be based on nanostructures. An important advantage for nanostructures solar cells is that they can be used to incorporate new physical mechanisms that allow the efficiency greater than that of a one-junction solar cell (Razykov et al., 2011).

In thermodynamic point of view, photovoltaic (PV) system (cell/module/panel/ array) performance can be evaluated in terms both energy and exergy (availability). Energy analysis (energetic) is based on the first law of thermo-dynamics meanwhile exergy analysis (exergetic) is based on both the First and the Second Laws of Thermodynamics. Unlike energy, exergy is not subject to a conversion law (except for ideal or reversible processes). Exergy is consumed or destroyed, due to irreversibility in any real processes. The exergy consumption during a process is proportional to the entropy created due to irreversibility associated with the process.

In the framework of the PV Enlargement project of EU, since October 8, 2005, a fixed frame 10 kWp grid-connected PV array system has been installed at Szent István University (SZIU), Gödöllő, Hungary. The surface orientation for this system is 30^{o} for tilt angle (β) and 5^{o} to East (South facing) for azimuth angle (γ). The systems used two different of PV technology i.e. polycrystalline PV technology (ASE-100) and amorphous silicon PV technology (DS-40). The schematic diagram of the installation can be seen in Figure 2 (Farkas et al., 2008 and Seres et al., 2009).

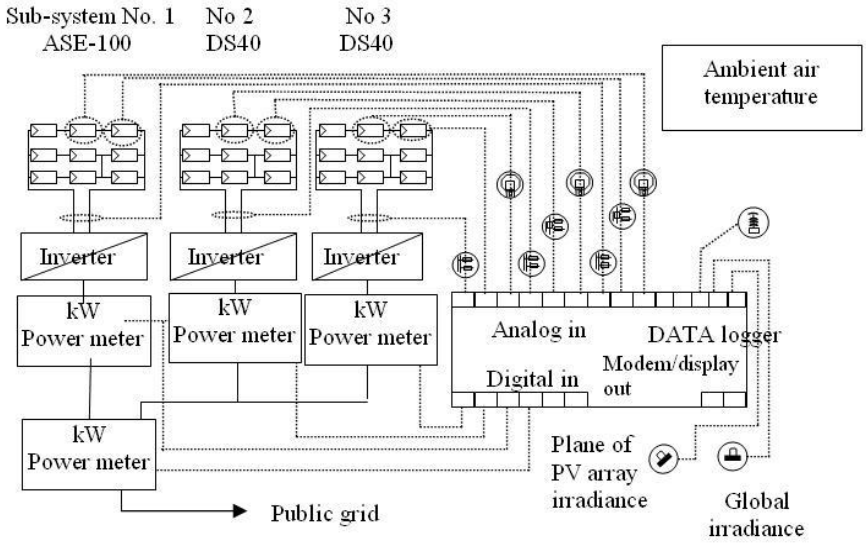


Figure 2. Schematic diagram of a 10 kWp grid-connected PV array at Szent István University, Gödöllő

Previously, exergy evaluation (yearly bases) of both above PV modules technologies have beeen performed theoretically by "solar energy parameters" and "photonic energy" model. A monthly data for analysis were taken from PV*SOL 3.0 software packages, which acquires data from MeteoSyn, Meteonorm, PVGIS, NASA SSE, SWERA (Rusirawan et al., 2011).
In this research, an exergetic evaluation of PV module, as a basic component of a 10 kWp grid-connected PV array system at Szent István University, Gödöllő - Hungary, will be presented based on experimental data refers to general exergetic model, which involves simultaneous both of The First Law and The Second Law of Thermodynamics. As a long term target of this research, a new method to optimize and increase the overall performance of grid-connected PV array system at Szent István University, as shown in Fig. 2., can be studied, observed and proposed.

2. Availability (exergy) analysis of PV modules

The other name of availability is available energy or available work, but currently all these terminology has been formulated as an exergy (standard definition). The exergy of a thermodynamic system is the maximum work that can be done by the system when undergoes reversible processes that bring the system into complete thermodynamic equilibrium with a defined reference environment. Exergy is always destroyed when a process involves a temperature change. This destruction is proportional to the entropy increase of the system together with its surroundings. The destroyed exergy has been called anergy

(Agrawal et al., 2011). Generally, correlation between energy, exergy and anergy can be written as follow:

$$\text{Energy} = \text{exergy} + \text{anergy} \tag{1}$$

Exergy analysis identifies the location, magnitude and the source of thermodynamic inefficiencies in a system. This information, which can not be provided by other means (e.g., energy analysis), is very useful for the improvement and cost-effectiveness of the system. The exergy analysis is based on the second law of thermodynamics, which includes accounting the total exergy inflow, exergy outflow and exergy destructed from the system.
Exergy analysis of PV model based on terminology fluxes energy and entropy, can be seen in Figure 3. All in and outgoing fluxes in the thermodynamic balance equations for energy and entropy must be known. Radiation energy flux densities are denoted e ($W.m^{-2}$) , and entropy flux densities s ($W.m^{-2}.K^{-1}$). The arrows indicating the energy fluxes show the respective content of available energy (Labuhn et al., 2001).

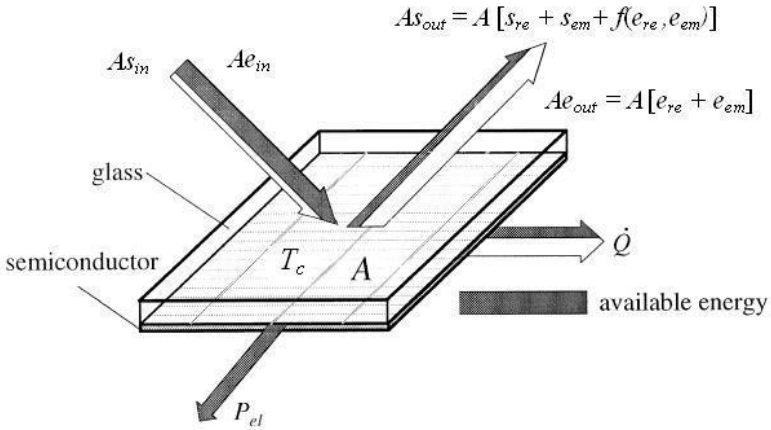


Figure 3. Energy fluxes Model of PV system

Based on Figure 3, the electric power P_{el} (W) is calculated from the energy balance equation:

$$0 = A(e_{in} - e_{out}) - \dot{Q} - P_{el} \tag{2}$$

and the entropy balance equation:

$$0 = A(s_{in} - s_{out}) - \frac{\dot{Q}}{T} + A\dot{s}_{ir} \tag{3}$$

After elimination of the heat flux $\dot{Q}$, finally P_{el} can be determined as follow:

$$\frac{P_{el}}{A} = (e_{in} - e_{out}) + T(s_{out} - s_{in}) - T\dot{s}_{ir} \tag{4}$$

where $\dot{s}_{ir}$ is the entropy production rate per surface area (W.m^{-2}.K^{-1}), A is the photovoltaic area of the module (m^2), $\dot{Q}$ is the heat flux (W.m^{-2}), e_{re} is the flux energy – reflection (W.m^{-2}) and e_{em} is the flux energy – emission (W.m^{-2}).

3. Method of evaluation

To simplify the concept such as shown in equations (1) – (4), the "Sankey diagram" as can be seen in Figure 4, can be implemented as a guidance in order to get all the parameters values (Kabelac, 2008).

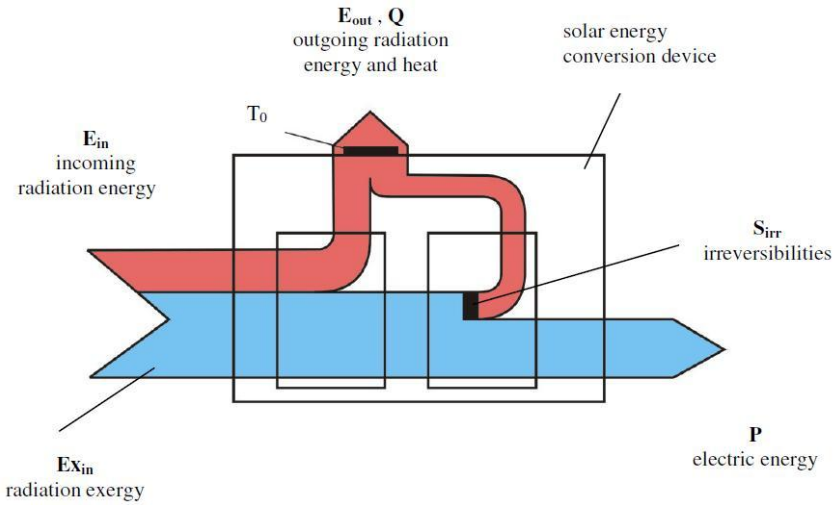


Figure 4. Energy and exergy flow diagrams in solar energy conversion device

It is known that the main energy source of PV system and other solar devices is the solar irradiance incident on the modules. To evaluate the exergy efficiency of PV modules, therefore, firstly the exergy of the total solar irradiation (W) for considered locations are needed, and can be expressed as follow (Petela, 2003):

$$\dot{E}x_{solar} = G \times A \times \left(1 - \frac{4}{3}\frac{T_a}{T_s} + \frac{1}{3}\left(\frac{T_a}{T_s}\right)^4\right) \tag{5}$$

The exergy electrical efficiency of a system in general can be given as:

$$\eta_{ex} = \frac{\dot{E}x_{out}}{\dot{E}x_{in}} = \frac{\dot{E}x_{electrical} + \dot{E}x_{thermal} + \dot{E}x_{dest.}}{\dot{E}x_{solar}} = \frac{\dot{E}x_{electrical} + I_r}{\dot{E}x_{solar}} \quad (6)$$

$$I_r = \sum Ex_{dest.} = \dot{E}x_{dest.thermal} + \dot{E}x_{dest.electrical} \quad (7)$$

I_r contains both internal and external losses. Internal losses are electrical exergy destruction, $\dot{E}x_{dest,electrical}$ and external losses are heat losses, $\dot{E}x_{dest,thermal}$, which is numerically equal to $\dot{E}x_{thermal}$ for PV system analysis (Joshi et al., 2009).

In practice, equation (6) can be replaced by the following expression (Calderón et al., 2011):

$$\eta_{ex} = \frac{\dot{E}x_{out}}{\dot{E}x_{in}} = \frac{P_{el}}{\dot{E}x_{solar}} = \frac{V \times I}{\dot{E}x_{solar}} \quad (8)$$

where P_{el} is actual electric power generated by PV modules and $\dot{E}x_{solar}$ is exergy of the solar radiation (the maximum electrical power obtainable from the radiation).

Based on equation (8), the value of destroyed exergy (anergy) can be evaluation as well.

System in Figure 2 is equipped with data logger, which recording and collecting the data. Not all data are recorded. The software system calculates 10 minutes average values from the measured data and these averages are recorded to the hard drive of the PC. The screenshot of the data logger during the data collection is presented in Figure 5.

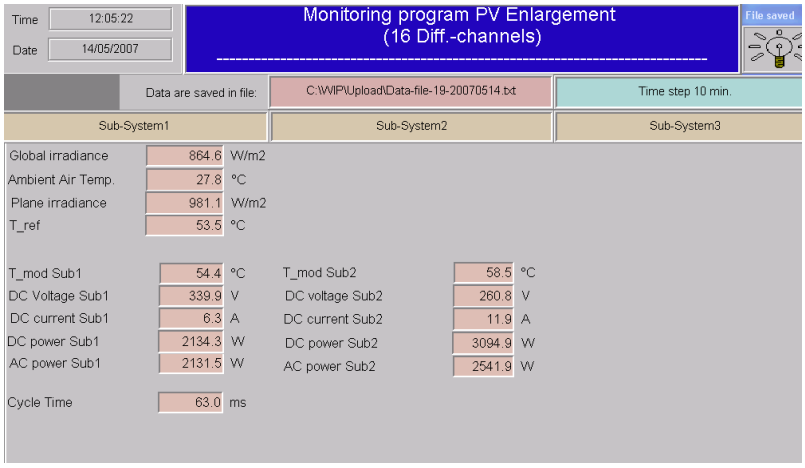


Figure 5. Screenshot of the data logger PC during data acquisition process

4. Results and discussion

Figures 6 and 7 show the values of the overall radiation incident on the inclined plane of the PV modules, *G* (W/m^2); the exergy of the radiation incident on the inclined plane, *Ex-solar* (in W/m^2); the power supplied by the PV modules, *P* (in W/m^2); the major loss exergy (exergy destroy) as a consequence of the irreversibilities of the process, *Ex_loss* (in W/m^2); actual electrical efficiency, *Eff_act* (%); and exergy efficiency, *Eff_ex* (%); both for ASE – 100 and DS – 40. All data for evaluation is taken from database in May 02, 2007, and evaluation have performed at constant temperature, i.e. average temperature in May (16.48^{o}C).

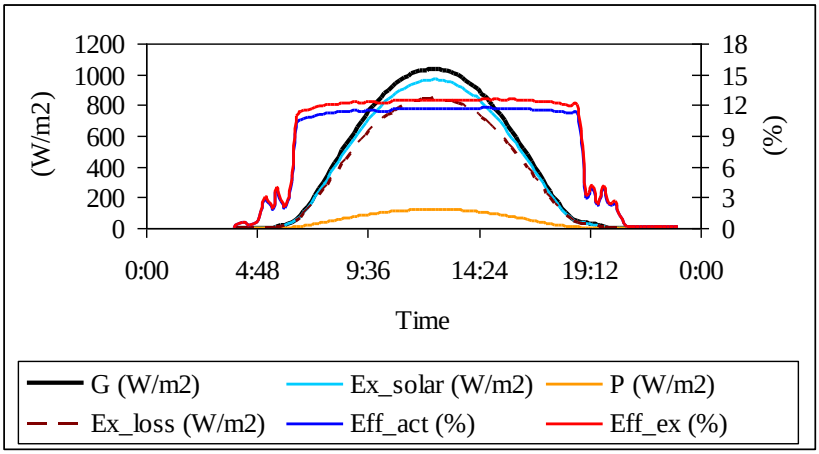


Figure 6. Daily based of exergetic characteristic of ASE-100 (poly-crystalline modules)

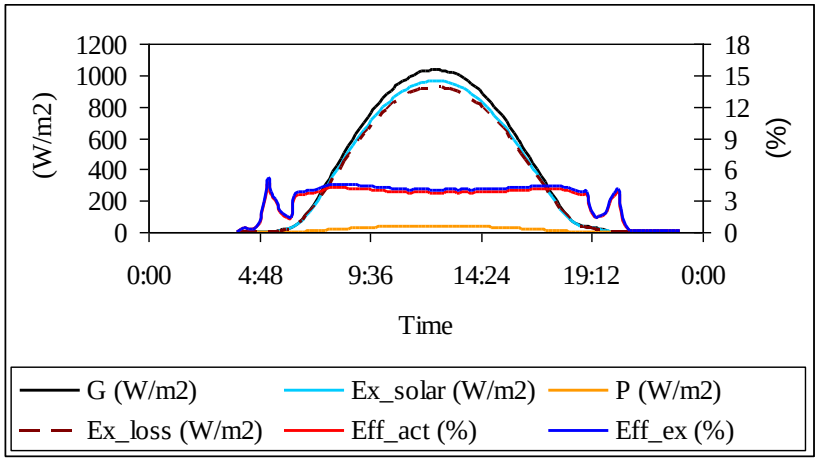


Figure 7. Daily based of exergetic characteristic of DS-40 (amorphous silicon modules)

In Figures 8 and 9, the percentage of exergy destroyed (anergy), both in PV modules are shown, refers to the exergy efficiency values.

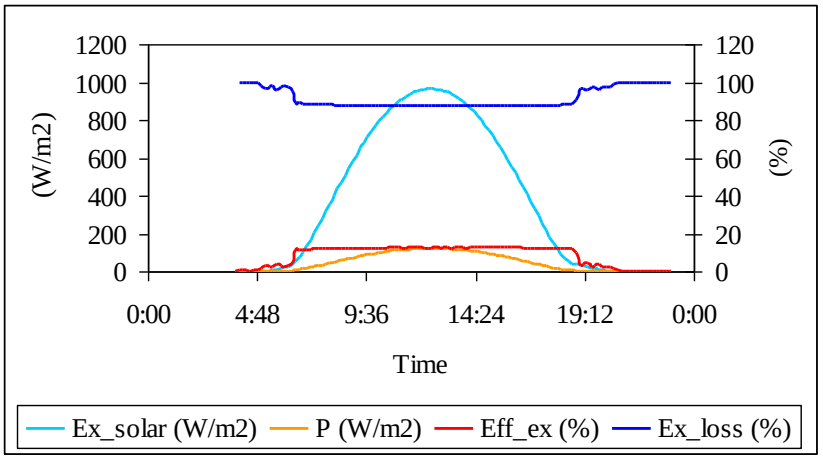


Figure 8. Exergy and anergy (exergy destroy) of ASE-100 (poly-crystalline modules)

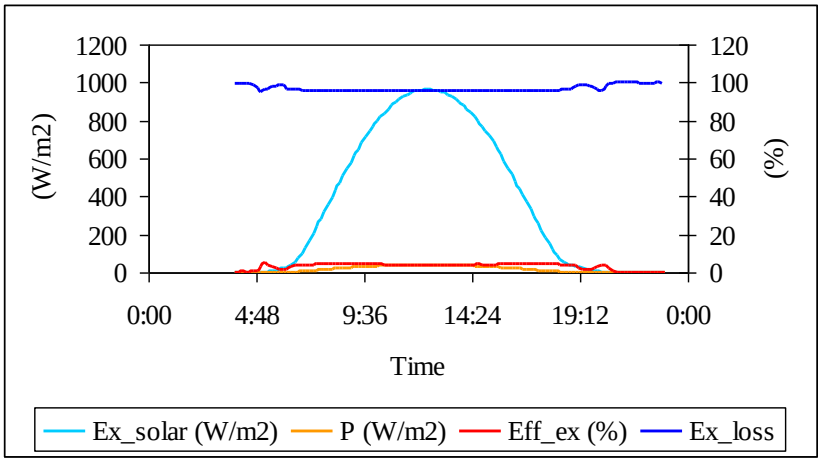


Figure 9. Exergy and anergy (exergy destroy) of DS-40 (amorphous silicon modules)

Based on Figures 6-9, it's clear that the actual power extracted from the PV modules (*P*) over the day is much less than could be extracted (*Ex_solar*). The characteristics showed that both the PV modules, in average, have a low exergy efficiency, *Eff_ex*, i.e. 12% and 4% (equal with anergy 88% and 96%) for ASE-100 and DS-40, respectively (slight different with an actual electrical efficiency, *Eff_act*).

Because of the solar radiation energy (G) is always larger than the solar radiation exergy ($\dot{E}x_{solar}$), and the electrical energy and electrical exergy are equal, the actual (energy) efficiency of the PV modules is always smaller than the exergy efficiency.

5. Conclusion

An availability (exergy) analysis, as energy behavior of the two photovoltaic modules, i.e. ASE – 100 (polycrystalline silicon) and DS – 40 (amorphous silicon), has described in this paper based on experimental data. The empirical characteristics, both exergy efficiency and exergy destroy (anergy) that occur in the system as a results of reversibilities of the PV modules, have been evaluated and calculated throughout the day (daily bases). As expected, the exergy efficiency characteristics of ASE-100 module (which included crystalline materials type) are higher than DS-40 module (which included thin film materials type). Further deep analysis and studies are still needed in order to get possibility improvement of PV efficiency.

Acknowledgements

This research is carried out with the support of OTKA K 84150 project, Hungarian Scholarship Committee and the Ministry of National Education of the Republic Indonesia.

References

Agrawal, S. and Tiwari, G.N. (2011): Energy and exergy analysis of hybrid micro-channel photovoltaic thermal module. Solar Energy (85) 356–370.

Borenstein, S. (2008): The Market Value and Cost of Solar Photovoltaic Electricity Production. Centre for the study of energy market (CSEM), University of California Energy Institute, Berkeley – California. Available from: <http://www.ucei.berkeley.edu/PDF/csemwp176.pdf>, Accessed on February 17, 2012.

Calderón, M., Calderón, A. J., Ramiro, A., González, J. F. and González, I. (2011): Analysis of the performance of the photovoltaic array through the exergy efficiency. International Conference on Renewable Energy and Power Quality, 2011 April 13-15, Las Palmas, Spain. Available from: <http://www.icrepq.com/icrepq%2711/280-calderon.pdf>, accessed on September 23, 2011.

Farkas, I. and Seres, I. (2008): Operational experiences with small-scale grid-connected PV system. Szent István University Faculty of Mechanical Engineering, R&D in Mechanical Engineering Letters (1), 64-72.

International energy agency (2010): Technology road map solar photovoltaic energy. Available from: <http://www.iea.org/papers/2010/pv_roadmap.pdf>, accessed on February 19, 2012.

Joshi, A.S., Dincer, I. and Reddy B.V. (2009): Performance analysis of photovoltaic systems: a review. Renewable and Sustainable Energy Reviews (13) 1884–1897.

Kabelac, S. (2008): Thermodynamic basic of solar radiation. 5th European Thermal-Sciences Conference 2008, The Netherlands.Available from: <http://www.eurotherm2008.tue.nl/>, accessed on November 17, 2011.

Labuhn, D. And Kabelac, S. (2001): The spectral directional emissivity of photovoltaic surfaces. International Journal of Thermophysics [22] No. 5, September 2001, 1577-1592.

Petela, R. (2003): Exergy of undiluted thermal radiation, Solar Energy (74) 469-488.

Razykov, T.M., Ferekides, C.S., Morel, D., Stefanakos, E., Ullal, H.S. and Upadhyaya, H.M. (2011): Solar photovoltaic electricity: current status and future prospects. Solar Energy [85] 1580–1608.

Rusirawan, D. and Farkas, I. (2011): Exergetic assessment of polycrystalline and amorphous photovoltaic modules in different methods, Proceeding of the 2nd International Conference in Agricultural Engineering, Synergy 2011, October 9-15, 2011, Gödöllő, Hungary.

Seres, I. and Farkas, I. (2009): Energy production issues of a 10 kWp photovoltaic system. Proceeding of the International Conference in Agricultural Engineering, Synergy 2009, August 30 – September 2, 2009, Gödöllő, Hungary.

4. INSTITUTE FOR MECHANICAL ENGINEERING TECHNOLOGY

PROFESSOR DR. GÁBOR KALÁCSKA
DIRECTOR OF THE INSTITUTE

Dear Reader,

The Institute for Mechanical Engineering Technology (consists of three departments: Department of Material and Engineering Technology, Department of Maintenance of Machinery, Department of Mechatronics) faced an unusual year during 2011.
The year was a success of PhD research activities concerning four defended dissertations based on a systematic research carried out at the Institute in the field of materials, technologies and maintenance.

- Concerning small aircrafts in order to increase the aviation safety as well as to improve the quality of technical and using services a new model of the service system was developed. The model provides theoretical basis for a software development connected to the theme, too.
- The abrasion resistance as a newly arisen load of the multilayer hot-dip galvanized coatings was investigated. Performing rapid comparative laboratory abrasion tests the coatings with different composition hot-dip galvanized were compared to each other and the empirical characteristics of wear resistance in the function of frictional length, of speed and of abrasion medium pressure were determined.
- Extended research of machining and tribological applicability of ZrO_2-MgO ceramic surfaces were carried out. Different technological parameters were analyzed having PCD and CBN cutting tools.
- In the field of renewing technologies the "cold metal spraying" was investigated. The connection between the technology and the properties of cold flame-sprayed porous metallic coatings were determined with the selected powders.

In Mechatronics:

- New results of ER fluids composed with different additives were successfully investigated as it is shown in the article.

Beside the research activities the whole infrastructure of the laboratories and workshops were overhauled resulting better condition for the further scientific work at Institute for Mechanical Engineering Technology.
More information about the Institute: www.geti.gek.szie.hu

THE EXAMINATION OF ELECTRORHEOLOGICAL PROPERTIES OF TiO_2/MINERAL OIL SUSPENSIONS UNDER FLOW MODE

László FÖLDI, Eszter SÁRKÖZI, László JÁNOSI

Department of Mechatronics, Institute for Mechanical Engineering Technology

Abstract

A claim has come up in the research of the mechatronical systems to develop special materials that can get information from a computer and change their attribution accordingly. From a control technology point of view using electric or magnetic field is the most obvious. One group of these materials is electro-rheologic (ER) fluids that change their shear strength feature according to the electric field applied. Electrorheological (ER) fluids composed of TiO_2 particles suspended in mineral oil (SN 150 A) are studied in this work. The rheological response has been characterized as a function of field strength, flow rate and volume fraction.
The main goal of this work is to build up a test station for determining the effects and the relationships between the ER fluid parameters and the electric field applied. The unique ER valve construction and the test apparatus are shown and also the function tests of this unit have been carried out.

Keywords

Electrorheological effect, Flow control, Hydraulics

Introduction

The electro-rheological effect
The change in apparent viscosity of electro-rheological (ER) fluid is dependent on the applied electric field, i.e. the potential divided by the distance between the plates. The change is not a simple change in viscosity, hence these fluids are now known as ER fluids, rather than by the older term Electro Viscous fluids. The effect is better described as an electric field dependent shear yield stress. When activated an ER fluid behaves as a Bingham plastic (a type of viscoelastic material), with a yield point which is determined by the electric field strength. After the yield point is reached, the fluid shears as a fluid, i.e. the incremental shear stress is proportional to the rate of shear (in a Newtonian fluid there is no yield point and stress is directly proportional to shear). Hence the resistance to motion of the fluid can be controlled by adjusting the applied electric field.

Properties of electro-rheological fluids

Electro-rheological fluids are suspensions of extremely fine non-conducting particles (up to 50 micrometers diameter) in an electrically insulating fluid. The apparent viscosity of these fluids changes reversibly by an order of up to 100,000 in response to an electric field. For example, a typical ER fluid can go from the consistency of a liquid to that of a gel, and back, with response times on the order of milliseconds. The effect is sometimes called the Winslow effect, after its discoverer the American inventor Willis Winslow.

Challenges and advantages of ER phenomenon

A major problem is that ER fluids are suspensions, hence in time they tend to settle out, so advanced ER fluids tackle this problem by means such as matching the densities of the solid and liquid components, or by using nanoparticles, which brings ER fluids into line with the development of magneto rheological fluids. Another problem is that the breakdown voltage of air is ~ 3 kV/mm, which is near the electric field needed for ER devices to operate. An advantage is that an ER device can control considerably more mechanical power than the electrical power used to control the effect, i.e. it can act as a power amplifier. But the main advantage is the speed of response, there are few other effects able to control such large amounts of mechanical or hydraulic power so rapidly. Unfortunately, the increase in apparent viscosity experienced by most electro-rheological fluids used in shear or flow modes is relatively limited. The ER fluid changes from a Newtonian liquid to a partially crystalline "semi-hard slush". However, an almost complete liquid to solid phase change can be obtained when the electro-rheological fluid additionally experiences compressive stress (MONKMAN, 1995). This effect has been used to provide very effective clutches or hydraulic valves.

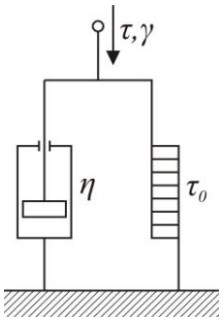


Fig. 1. The Bingham model

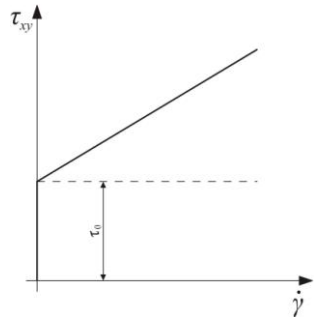


Fig. 2. The shearing tension in the function of shearing gradient

Development, structure and operation of an electro-rheological valve

The ER phenomenon as a material model is described by the Bingham model related to ER liquids, which is based on the mathematical model describing the non-newton liquids (LEE, 2002). The Bingham model is a complex viscoplastic rheological model (CSIZMADIA, 2003). As a material model it can be divided into an ideally ductile an ideally viscous member (ZRÍNYI, 1999). (Fig. 1)

Research objectives
The aim of testing flow properties of the ER liquids is to establish a model, which can be generally applied of testing the behaviors of ER liquids made of different materials considering the physical data of the applied materials. By computer simulation the parameters of this mathematical model can be determined, which can be generated by minimum search. As the first step of the model identification the approximate search of the minimum of the established target function is accomplished by genetic algorithm, then the refinement of the result with the known numerical methods.

The ER valve
During the further material and application tests it is practical to use a device which has application possibilities too. In this consideration for further investigations we need the design of a hydraulic ER valve, which can be operated built in hydraulic circuits.

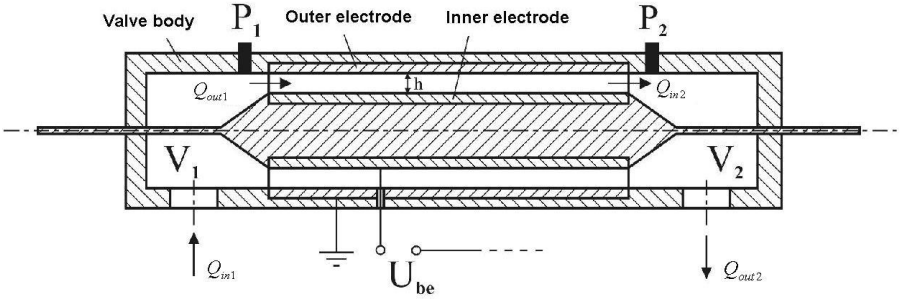


Fig. 3. The ER valve

Development of measuring system for investigating the features of electro-rheological fluids

Applied test principle
Consider two concentric pipes (external and internal electrodes) and constant fluid flow is flowing between them. An electric field will be developed according to the electric potential applied on the electrodes and the gap between

the electrodes and all of these causes the change of shearing strengths of fluid. The magnitude of change can be definitely determined from the pressure drop. During the test series the fixed parameters relating to the individual set of measurement are: fluid flow, oil temperature, fluid concentration. The only variable parameter is: the electric field. The parameter to be tested is the pressure drop and the calculated one is the shearing strength.

Design of hydraulic measuring circuit

The system consists of three main parts, like: the hydraulic power unit, ER circuit, data acquisition and control system. The hydraulic unit ensures the flow of ER fluid by a driving servo cylinder. The ER valve with flow meter and the pressure difference meter are fitted into the ER circuit.

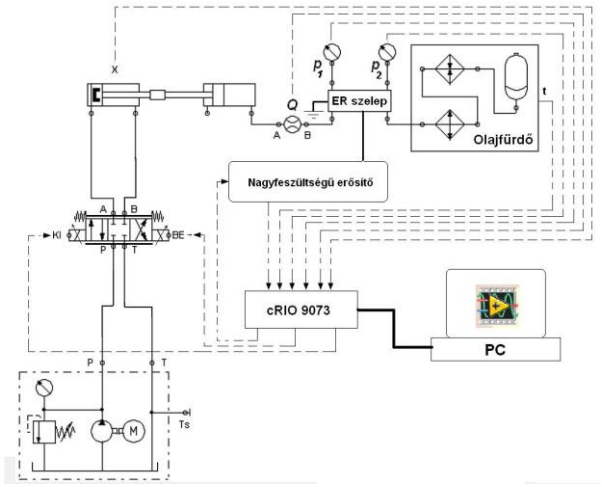


Fig. 4. Experimental setup

Test control and data acquisition

The applied NI CompactRIO™ programmable automation controller is a modular system; out of its modules we used the analogue-to-digital converter (NI 9201), for a dual purpose. On the one hand we applied it in the controlling process to measure the voltage signal (which is in proportion to the displacement) provided by the displacement encoder. On the other hand we used it in collecting data about the voltage signals corresponding to pressure values provided by the analogue pressure sensors. The communication between the CompactRIO™ and the computer was ensured by an Ethernet connection. We realised the real-time control by applying the FPGA module of CompactRIO™ programming it in the LabVIEW 2009 software.

Test results

In my test system flow mode takes place. Unlike the pressure increase process seen in other professional studies, the cases that I examined showed different pressure increase. In the following pages I am presenting a typical measurement result with analysing its course.

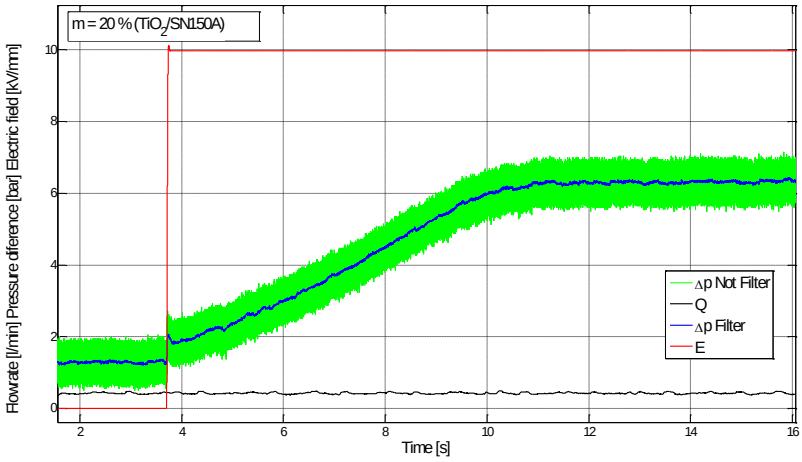
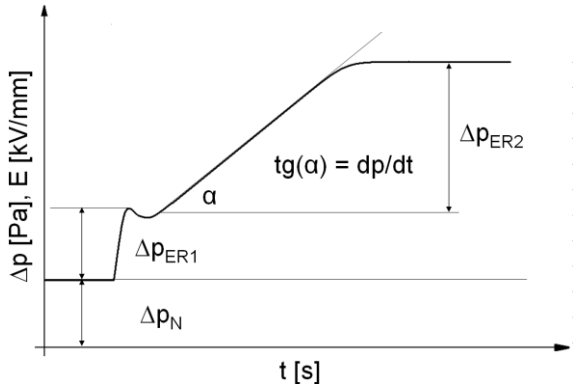


Fig. 5. Pressure increase process (m = 20 %, E = 10 kV/mm, Q = 0,42 L/min)

The pressure increase consists of several parts. The pressure of the section that preceeds the switching on of the electric field is given from the basic viscosity of the ER fluid (Δp_N). After switching on the electric field a rapid pressure increase can be observed, whose response time (τ_1) in this measurement was 28 ms. The pressure drop decreases from this point on, reaches a minimum value then it increases again. Hereinafter the mean value between the minimum and maximum values of this section will be designated as **primary pressure increase (Δp_{ER1})**.
The secondary pressure increase goes on to a maximum value. The intensity of this secondary pressure increase is far smaller than that of the primary pressure increase. Its profile can be approached by a linear trend, hereinafter I will name this section as **intensity of the secondary pressure increase** (**dp/dt**). The pressure increase reaches a maximum value, the mean value of this section will be referred to as **secondary pressure increase** (Δp_{ER2}).



Δp_N - pressure increase given from the basic viscosity, [Pa]
Δp_{ER1} - pimary pressure increase given from the ER effect, [Pa]
Δp_{ER2}- secondary pressure increase given from the ER effect, [Pa]
dp/dt - intensity of the secondary pressure increase, [Pa/s]
E -electric field, [kV/mm]

Fig. 6. The sections of pressure increase in flow mode

Experiment plan
In the tests in permanent flow rate circumstances I switch the electric field step-response-wise and I measure the pressure drop at the valve. I examine the change of the measured pressure drop in the function of time. The changed parameters are the ER fluid quantity (Q) flowing through the valve, the electric field (E), which can be counted from the distance between the electrodes and the voltage controlled by the high voltage amplifier.

Flowrates applied in the experiments:
1st chart: Experimental settings (Flowrate)

Q_1	Q_2	Q_3	Q_4	Q_5
[litre/min]	[litre/min]	[litre/min]	[litre/min]	[litre/min]
0,42	0,65	0,85	1,05	1,3

I modify the electric field as follows:
2nd chart: Experimental settings (Electric field)

E_1	E_2	E_3	E_4	E_5	E_6
[kV/mm]	[kV/mm]	[kV/mm]	[kV/mm]	[kV/mm]	[kV/mm]
0	2	4	6	8	10

Primary pressure increase in the function of the electric field and the flowrate
Having tested the dependency of the primary pressure increase from the variables (E,Q), I obtained an idea about the mathematical relations of the pressure increase process. My objective is to define the behaviour of the given

additive material contentrated (TiO_2/SN150A) ER fluids in the function of the electric field and the flowrate. For this I set a two-variable function in the case of every observed parameter. As the first step we need to find a suitable function and by identifying its parameters we are able to create the right function, which is the best fitting. It could be seen from the results of the measurements that the electric field dependency of the primary pressure increase can be well approached by power functions. However, as a result of the tests we can also see that the value of the primary pressure increase changes in the function of the flowrate to certain extent. In view of this, I chose the following function for approaching the electric field and flowrate dependency of the primary pressure increase:

$$\Delta p_{ER1} = (a \cdot Q + b) \cdot E^{c \cdot Q + d} \tag{1}$$

where:
- Q flowrate[L/min]
- E electric field [kV/mm]
- a, b, c, és d parameters

I then fitted this function on the mean values of the primary pressure increase measured at several experimental settings. I defined the parameters by the least squares method.

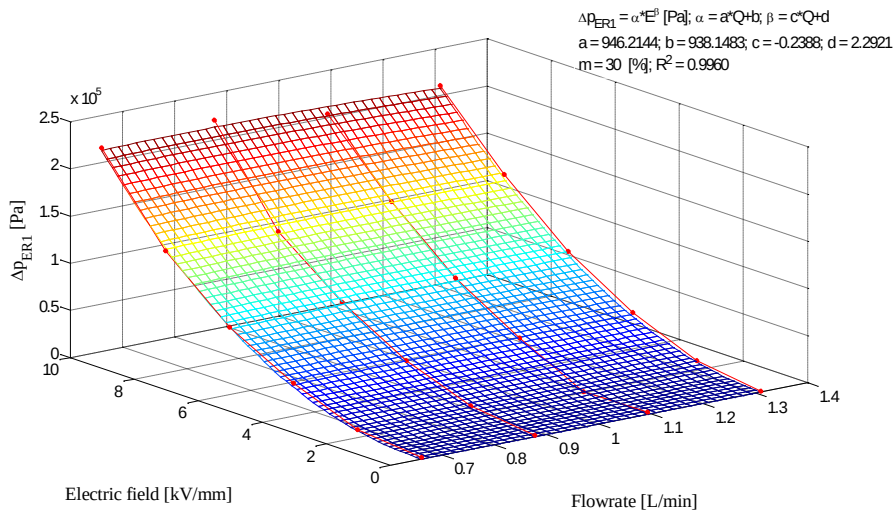


Fig. 7. Primary pressure increase in the function of the electric field and the flowrate (m = 30% additive material concentration TiO_2/SN150A blend)

The figure shows the mean values of the primary pressure increase at several experimental settings (Electric field: E = 0, 2, 4, 6, 8, 10 kV/mm, Flowrate: 0,66; 0,81; 1,06; 1,31; L/min) To each and every setting five repetitions belong. The figure also shows the function mentioned above ($\Delta p_{ER1} = (a \cdot Q + b) \cdot E^{c \cdot Q + d}$), whose parameters (a, b, c, d) were defined by the least squares method. Fixed parameter at the measurement: additive material concentration 30%. The value of the correlational coefficient: R^2 = 0,9960. Parameters: a = 946,2144; b = 938,1483; c = -0,2388; d = 2.2921.
The approaching function, by which the value of the primary pressure increase in the function of the electric field and the flowrate (at p = 99 % probability level), in the case of m = 30 % additive material concentration blend, in Q = 0,66 – 1,31 L/min and E = 0 – 10 kV/mm validity scope:

$$\Delta p_{ER1} = (946{,}2144 \cdot Q + 938{,}1483) \cdot E^{-0{,}2388Q+2{,}2921} \pm 16010 \quad [\text{Pa}]$$

Secondary pressure increase in the function of the electric field and flowrate
After examining the dependency of the secondary pressure increase from the individual variables (E, Q), I got an image about the mathematical relations of the pressure increase process. My objective is the definition of the behaviour of the given additive material concentration (TiO_2/SN150A) ER fluids in the function of the electric field and the flowrate. To do this I set a two-variable function in the case of every tested parameter. As the first step we must select a suitable fuction, and by the identification of its parameters we can create the best fitting spatial function. From the test results it could be seen that the electric field dependency of the secondary pressure increase could be approached by autocatalytic functions. As a result of the tests we could also observe that the secondary pressure increase changes autocatalytically too, in the function of the flowrate. In view of this, I chose the following function for approaching the electric field and flowrate dependency of the secondary pressure increase:

$$\Delta p_{ER2} = (a \cdot E + b) \cdot \left(\frac{1}{1 + e^{c + d \cdot Q}} \right) + (\alpha \cdot Q + \beta) \cdot \left(\frac{1}{1 + e^{\gamma + \delta \cdot E}} \right) \quad (2)$$

Where:
- Q flowrate[L/min]
- E electric field [kV/mm]
- a, b, c, d, α, β, γ, and δ parameters

I fitted this function on the mean values of the secondary pressure increase measured at several experimental settings. I defined the parameters by the least squares method.

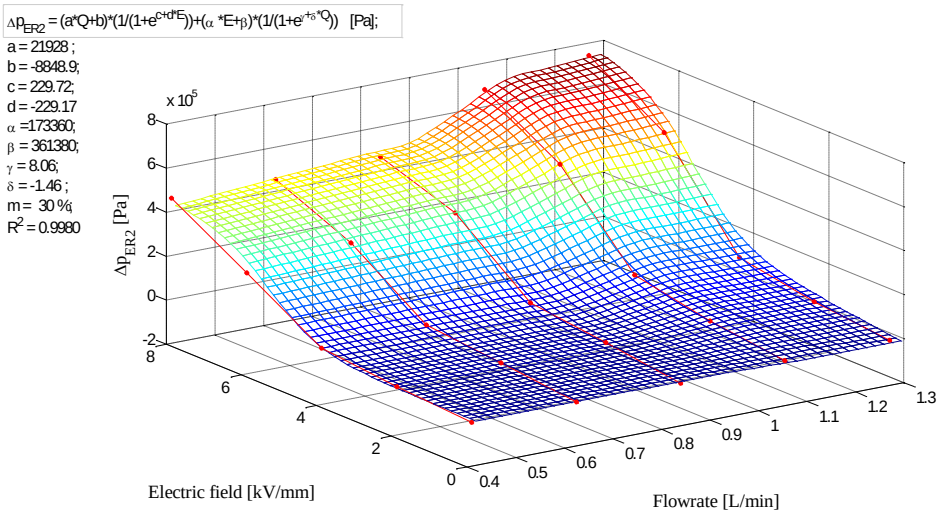


Fig. 8. Secondary pressure increase in the function of the electric field and the flowrate (m = 30 % additive material concentration TiO_2/SN150A blend)

The figure shows the mean values of the secondary pressure increase at several experimental settings (Electric field: E = 0, 2, 4, 6, 8 kV/mm, Flowrate: 0,42; 0,66; 0,81; 1,06; 1,31; L/min) To each and every setting five repetitions belong. The figure also shows the function mentioned above $\left(\Delta p_{ER2} = (a \cdot E + b) \cdot \left(\frac{1}{1+e^{c+d \cdot Q}}\right) + (\alpha \cdot Q + \beta) \cdot \left(\frac{1}{1+e^{\gamma+\delta \cdot E}}\right)\right)$, whose parameters (a, b, c, d, α, β, γ, és δ) were defined by the least squares method. Fixed parameter at the measurement: additive material concentration 30%. The value of the correlational coefficient: R^2 = 0,9980. Parameters: a = 21928; b = -8848,9; c = 229,72; d = -229,17; α = 173360; β = 361380; γ = 8,06; δ = -1,46.

The approaching function, by which the value of the secondary pressure increase in the function of the electric field and the flowrate (at p = 99% probability level), in the case of m = 30% additive material concentration blend, in Q = 0,42 – 1,31 L/min és E = 0 – 8 kV/mm validity scope:

$$\Delta p_{ER2} = (21928 \cdot Q - 8848{,}9) \cdot \left(\frac{1}{1+e^{229{,}72-229{,}17 \cdot E}}\right) +$$

$$+ (173360 \cdot E + 361380) \cdot \left(\frac{1}{1+e^{8{,}06-1{,}46 \cdot Q}}\right) \pm 91117 \ \text{[Pa]}$$

The intensity of the secondary pressure increase in the function of the electric field and the flowrate

Having tested the dependency of the intensity of the secondary pressure increase from the variables (E, Q), I obtained an idea about the mathematical relations of the pressure increase process. My objective is to define the behaviour of the given additive material contentrated (TiO_2/SN150A) ER fluids in the function of the electric field and the flowrate. For this I set a two-variable function in the case of every observed parameter. As the first step we need to find a suitable function and by identifying its parameters we are able to create the right function, which is the best fitting. It could be seen from the results of the measurements that the electric field dependency of intensity of the secondary pressure increase can be well approached by power functions. However, as a result of the tests we can also see that the value of the secondary pressure increase intensity changes in the function of the flowrate to certain extent. In view of this, I chose the following function for approaching the electric field and flowrate dependency of the secondary pressure increase intensity:

$$\frac{dp}{dt} = (a \cdot Q + b) \cdot E^{c \cdot Q + d} \tag{3}$$

Where:

- Q flowrate[L/min]
- E electric field [kV/mm]
- a, b, c, és d parameters

I fitted this function on the mean values of the secondary pressure increase intensity measured at several experimental settings. I defined the parameters by the least squares method.

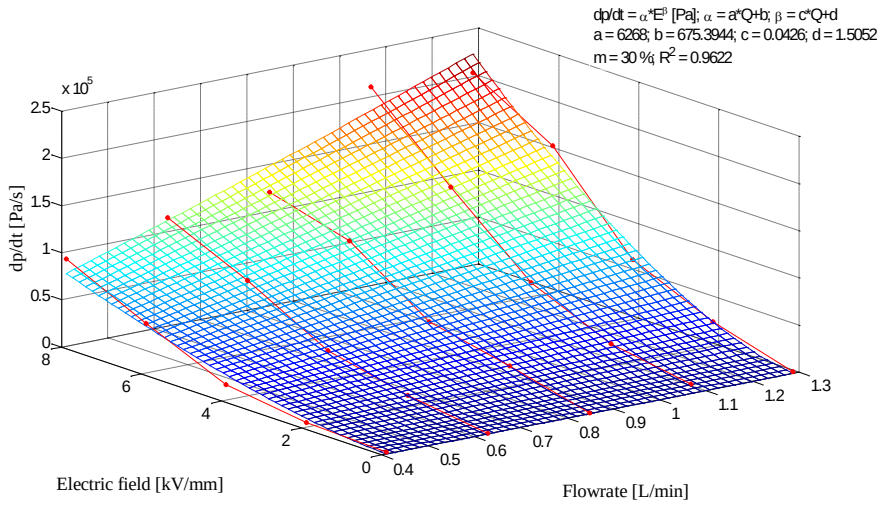


Fig. 9. Secondary pressure increase intensity in the function of the electric field and the flowrate (m = 30 % additive material concentration TiO_2/SN150A blend)

The figure shows the ($\frac{dp}{dt}$) mean values of the secondary pressure increase intensity at several experimental settings (Electric field: E = 0, 2, 4, 6, 8 kV/mm, Flowrate: 0,42; 0,66; 0,81; 1,06; 1,31; L/min) To each and every setting five repetitions belong. The figure also shows the function mentioned above ($\frac{dp}{dt} = (a \cdot Q + b) \cdot E^{c \cdot Q + d}$), whose parameters (a, b, c, d) were defined by the least squares method. Fixed parameter at the measurement: additive material concentration 30%. The value of the correlational coefficient: R^2 = 0, 9622. Parameters: a = 6268; b = 675,3944; c = 0,0426; d = 1,5052.
The approaching function, by which the value of the secondary pressure increase intensity in the function of the electric field and the flowrate (at p=99% probability level), in the case of m=30% additive material concentration blend, in Q = 0,42 – 1,31 L/min és E = 0 – 8 kV/mm validity scope:

$$\frac{dp}{dt} = (6268{,}0 \cdot Q + 675{,}3944) \cdot E^{0{,}0426Q+1{,}5052} \pm 43020[Pa/s]$$

Conclusions

I carried out these tests with a blend of SN 150 A mineral oil and 5 μm average particle size TiO_2 . In order to develop a stable blend I used Komad 309A-type dispergent. The magnitude of the applied electric field is 0-10 kV/mm. I varied the flowrate in a range of 0,41 - 1,3 L/min. The test temperature of the ER fluid is: 25^{o}C.

I worked out a new measuring system and test method for examining the ER fluids in flow mode, which is suitable to unfold some ER attributes such as the secondary pressure increase and the intensity of the secondary pressure increase. These attributes **cannot be measured** with the test systems functioning on the principle of rotation, which systems are generally used in experiments. With this test equipment the shear stress increase (the apparent viscosity increase) originated from the effect of the ER fluid can be determined dynamically from the pressure drop and the flowrate.

I established that in flow mode the pressure increase given from the change of the shear stress, which is changing because of the step response-shape electirc field variation consists of two sections: a rapid primary pressure increase right after the swithing on the electric field, and a secondary, slower (longer response time) section. I also found out that the intensity of the primary pressure increase is at least an order of magnitude bigger than that of the secondary one.

Based on the test results I determined the flowrate and the electric field dependency of the primary and secondary pressure increase and of the secondary pressure increse intensity in reference to a given additive material concentration ER fluid.
The connection between the **primary pressure increase** and the flowrate and the electric field – in the case of a given additive material concentration ER fluid – can be approached by the following function (the parameters of the equation should be determined in the function of the applied additive material concentration):

$$\Delta p_{ER1} = (a \cdot Q + b) \cdot E^{c \cdot Q + d}$$

The relationship between the **secondary pressure increase** and the electric field and the flowrate – in the case of a given additive material concentration ER fluid – can be approached by the following function (the parameters of the equation should be determined in the function of the applied additive material concentration):

$$\Delta p_{ER2} = (a \cdot E + b) \cdot \left(\frac{1}{1 + e^{c + d \cdot Q}} \right) + (\alpha \cdot Q + \beta) \cdot \left(\frac{1}{1 + e^{\gamma + \delta \cdot E}} \right)$$

And finally, the relationship between the **intensity of the secondary pressure increase** and the electric field and the flowrate – in the case of a given additive material concentration ER fluid – can be approached by the following function (the parameters of the equation should be determined in the function of the applied additive material concentration):

$$\frac{dp}{dt} = (a \cdot Q + b) \cdot E^{c \cdot Q + d}$$

References

Csizmadia B., Nándori E. (szerk.) (2003): Modellalkotás, Nemzeti Tankönyvkiadó, Budapest.

Lee H.-G., Choi S.B. (2002): Dynamic properties of an ER fluid shear and flow modes, In: Materials and Design, XXIII. Évf.

Monkman. G.J. (1995): The Electro-rheological Effect under Compressive Stress - Journal of Physics D: Applied Physics - Vol. 28, pp 588-593 - Institute of Physics.

Zrínyi M. (1999): Intelligens anyagok, Magyar Tudomány, 1999., 6. sz., 682.

ZIRCONIUM-DIOXIDE CERAMICS TURNING

Gellért FLEDRICH, Róbert KERESZTES, István PÁLINKÁS,
Lajos PELLÉNYI

Department of Material and Engineering Technology,
Institute for Mechanical Engineering Technology

Abstract

The engineering ceramics are such constructional materials which can be used at essentially higher temperature comparing to materials used till now, at under heavy physical and chemical load.
The claim to structural ceramics (for example: zirconium-dioxide) is continuously increases. The zirconium-dioxide [Salmang, Scholze (2007)] as basic material having lower hardness and other characteristics is suitable to machine by tool with regular edge, so in case of piece production or in case of small and medium series manufacturing, at quick prototype-manufacturing it can become potential material alike. Its cutting [Mészáros, Szepesi (2005)] and machinability characteristics has to be known that this could be ensue. My research work concentrates to one part of this, in accordance with the recommendation of a company producing and developing zirconium-dioxide semi-finished product.

Keywords

Zirconium-dioxide, turning, cubic boron-nitride, polycrystal diamond, surface topography, friction

1. Introduction

Greater material removal capacity can be reached by tools with regular edge than by grinding mainly in case of more complicated surfaces. The advantages can be perceptible in case of hole machining in particular. The sources of literature provide minimal reference to the machining with turning of ceramics, thus to the zirconium-dioxide and to the aluminium-oxide. Nowdays the grinding is the machining after the generally wide-spread sintering.

2. Material and method

During tests we have machined the surface of ceramic with different settings. Based on preliminary measurings we established that it is expedient to choose the depth of cut between 0,01 and 0,05 mm, the feed rate should be between 0,01 and 0,05 mm/rev. immediately at the tool edge. We also carried out preliminary measurings concerning the approximate value of the cutting speed, this was around 50 m/min. After turning we made surface topography as well as electron microscopic exposures from the surfaces machined and we analyzed those. We have measured the dry frictional characteristics of the ceramic / steel material pair with comparing character as a further test of the surfaces turned and grinded of the ceramic.

2.1. The materials tested

Zirconium-dioxide ceramic specimen [CeramTec, (2008)]
The characteristics of the ZN 40 engineering ceramic is that it has got favourable physical and chemical characteristics at high temperature range. It has got high hardness (1250 HV), because of this it can be cut only with polycrystal diamond and with cubic boron nitride tools. The material tested is a zirconium dioxide ceramic stabilized with magnesium. The diameter of the cylindrical specimens was 16 mm used at turning tests.

Steel specimen
To measure the frictional force we placed steel specimen to the ceramic surface. The material of the steel plates was St37F dimension 20x20 mm, thickness 1,5 mm. The contact surface was grinded, its average surface roughness was: R_a=0,8.

Single-point cutting tool
The cutting tool was a tipped turning tool with 12x12 –shank. We have carried out my tests with two kinds of inserts (tips) with cubic boron nitride (CBN) and with polycrystal diamond (PCD). It is expedient to carry out such hard cutting with these materials.

2.2. Cutting tests

To measure the axial and tangential components of the cutting force we applied resistance force-meter-tool head between the tool-head and tool shank developed and manufactured by us. It makes possible to measure the feed rate force (F_f) and the cutting force (F_c) in case of straight turning in this form.

2.3. Surface topographical tests

Because of the increasing requirements created against surfaces machined there is a need to evaluate many-sided the surface micro-topography which means the surface characteristics by writing up data collected from the sampling surface. I have made electron microscopic exposures from the ceramic surface machined with different cutting parameters as well as from surface grinded.

2.4. Friction tests

We have made an equipment to carry out tribological tests. During the test we pressured the steel counterface with determined normal direction force the casing surface of the rotating ceramic specimen and in the meantime we measure the value of the friction force with force meter cell. We have calculated the friction coefficient characterizing the system from the normal direction force and the friction force as well as we measured the wear of the steel specimen and its deformation.

2.5. Complementary tests with thermo-camera

The heat developing during cutting has got a great effect on the material removal process as well as it effects strongly the tool service-life. To study the heat affected zone developed we made exposures with thermo-camera during cutting. Considering that the basic material chosen by me for to be machined is significantly more rigid the heat deriving from internal friction presents oneself probably in smaller amount.

3. Results

3.1. Results of cutting tests

During tests we have determined the characteristic content and the exponents to given parameter of the empirical connection (feed rate, depth of cut, cutting speed) taking into also account the cutting speed from the values of main cutting force measured by using mathematical statistical methods.

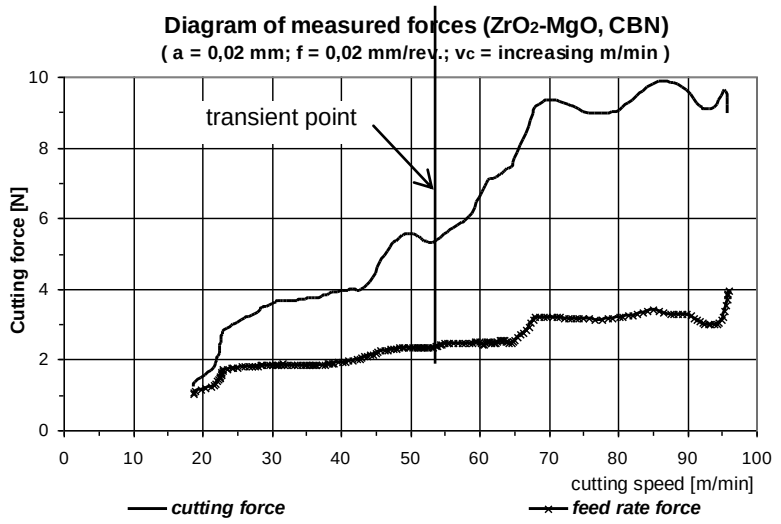


Figure 1. Diagram of main and feed rate forces.
(v_c= changing m/min., a= 0,02 mm, f= 0,02 mm/rev.; ceramic: zirconium dioxide, tool: CBN)

Cutting tests results of ceramic

The Figure 1. shows the characteristic values of measurings carried out with CBN – tool at uniformly increasing cutting speed. The cutting speed (v_c) measured along the horizontal axis the cutting forces ($F_{c;}$ F_f) can be found on the vertical axis.

In case of CBN-tool material the amount of feed rate forces are approximately the half of the amount of main cutting force. The value came about the cutting force with diamond (PCD) tool. The two forces at the two kind of tools show similar characters in their trends.

I also show the change of the cutting forces in the next diagram (Figure 2). I set the cutting speed to 25m/min. value. The value of the depth of cut was 0,02 mm. I changed the values of the feed rate according to f=0,01;-0,02;-0,03;-0,04;-0,05 mm/rev.

The main cutting force increases with the increase of the feed rate, however the value of the feed rate force almost hardly changes. At a=0,04 mm depth of cut the great degree deviation of the main cutting force allows to conclude to the damage of the cutting edge. That was proved later by microscopic exposures.

It is also significant the running up of the feed rate force in the beginning range. Both the ceramic and the tool are at ambient temperature in the beginning first point. During cutting significant amount of heat developed and warning up starts. As an effect of this the value of the force increases to a certain time then sets in a near constant value it can be seen in the diagram.

I have also measured the changes of the main cutting force in accordance with the formers to the polycrystal diamond tool. The value of the cutting speed was, v_c=25m/min. The value of the depth of cut was, a=0,02 mm. The values of the feed rate were, f=0,01;-0,02;-0,03;-0,04;-0,05 mm/rev.

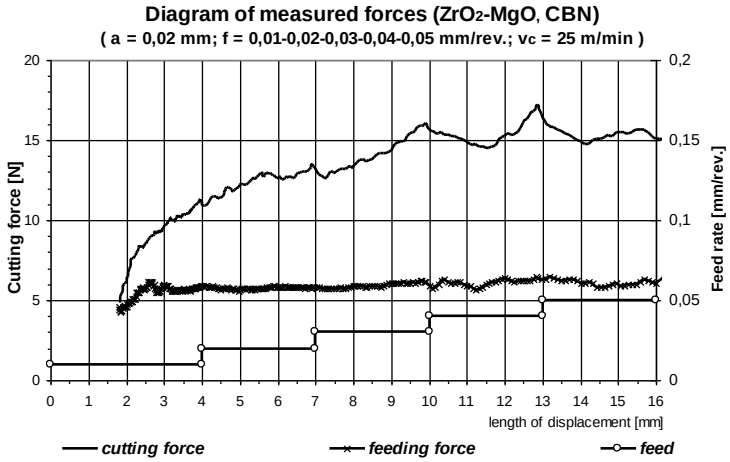


Figure 2. Diagram of main and feed rate forces.

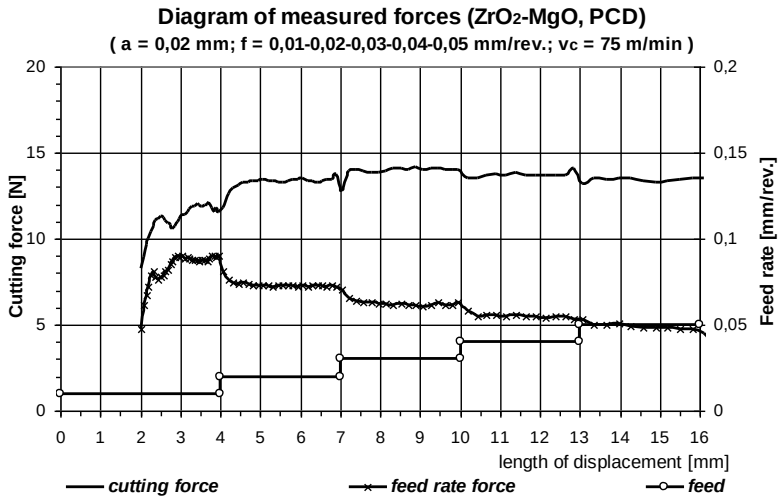


Figure 3. Diagram of main and feed rate forces

The trends are similar to the cubic boron nitride tool. The cutting force shows increasing tendency, while the feed rate force set in to a certain value, then showed slight decrease.
The cutting force at higher cutting speed (Figure 3) set in a nearly constant value following a steep running up which was lower than at smaller (v_c=25 m/min.) cutting speed. The feed rate force following the running up or decreased or had got the same value. The tool edge caused the deviation of the feed rate force with great probability.

3.2. Determining the cutting force in the function of parameters
The sources of literature to the cutting force [Horváth, Markos 1995], [König, Klocke (1997)], use a theoretical relation worked out first of all to steels which is the following in the function of cutting parameters:

$$F_c = C_v \cdot f^x \cdot a^y \cdot v_c^z \ [N]$$

My aim is with the planned tests to be carried out to decide the usefulness of the preceding equation at cutting ceramic and to determine the necessary parameters with multivariate linear regression.
The mathematical function matched to the results of tests planned:

$$F = 477{,}183 \cdot f^{0{,}1646} \cdot a^{0{,}4397} \cdot v_c^{-0{,}2994}$$

It is important to mention concerning the usefulness of the equation that this relation describes the formation of the measuring results within the parameter – range determined in the test plan.

3.3. Results of topographic surface tests

Results of microscopic examinations and 3D-al topographic surface tests
The quality of the surface reflects well the cutting quality. Microcracks develop on the surface during machining deriving from the britleness of ceramics. It is undesirable the developing of cracks because it is indispensable the proper choosing of cutting parameters.
The 3D-al topographic measurings made numerical the surface roughness extending in space, the Figure 4. shows such a sample surfaces. The surface characteristics were nearly as good as the surface characteristics grinded in case of certain cases of the cutting settings.

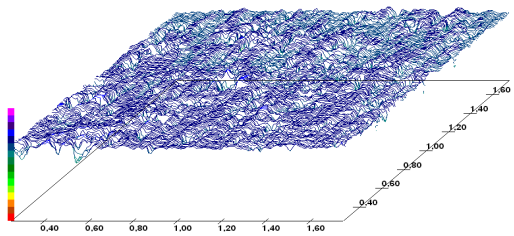


Figure 4. Ceramic surface turned with PCD-tool (v_c=25m/min., a=0,02mm, f=0,05mm/rev.)

Results of surface check with scanning electron microscope
For the sake of greater manification we have made scanning electron microscopic exposures in 100x, 500x then 2000x-times magnifications from the ceramic surface machined.

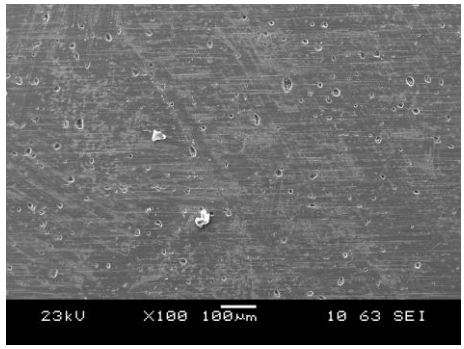


Figure 5. Ceramic surface grinded

The original grinded surface electron microscopic exposure can be seen in Figure 5. The amount of craters as a result of machining can be considered identical with the machining with PCD-tool. However the dimension of craters formed are 30-40% greater. This dimensional difference can increase the lubricant keeping capacity of the surface. The grinding grains made also the surface into serrated but because of the smaller dimension of grains the ditches, grooves, scratches dimension are also smaller:

The exposures with 2000x magnification prove the producing plastic (ductile) chip removal (Figure 6.) which shows favourable surface-continuity. This proves the applicability of turning with polycrystal diamond.

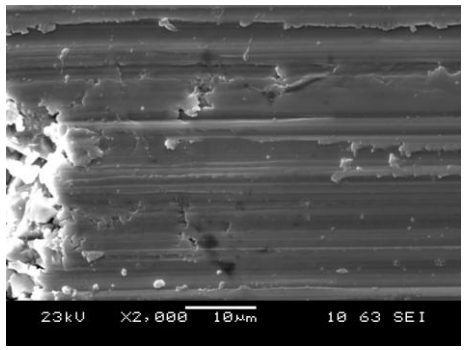


Figure 6. Surface turned with PCD-tool v_c=75m/min., f=0,04 mm/rev., a=0,02 mm, N=2000x

3.4. Results of frictional model tests and their interpretations

We analyze in the followings the values of the friction coefficient, wear and deformation got in the friction model testing system produced during my research work. The polycrystal diamond tool material proved to be suitable from the two tool materials used during machining ceramic. Because of this we have carried out the frictional tests with the specimens cut with PCD tool as well as with the original surfaces grinded.

Results of friction and wear and their evaluation

On the casing surface of ZrO_2 ceramic tested at v_c=25 and 75 m/min. cutting speed and 5 different feed rate (f=0,01;-0,02;-0,03;-0,04;-0,05 mm/rev.) machined surface can be found. The width of these are 3 mm one by one. Figure 7. shows the value of arising friction force developed between the ceramic and the steel specimen. We distinguished the different surfaces with various colours and marks. The “k” marks the grinded surface while the numbers “1, 2, 3, 4, 5” mark the feed rates set during turning in hundredth millimeter.

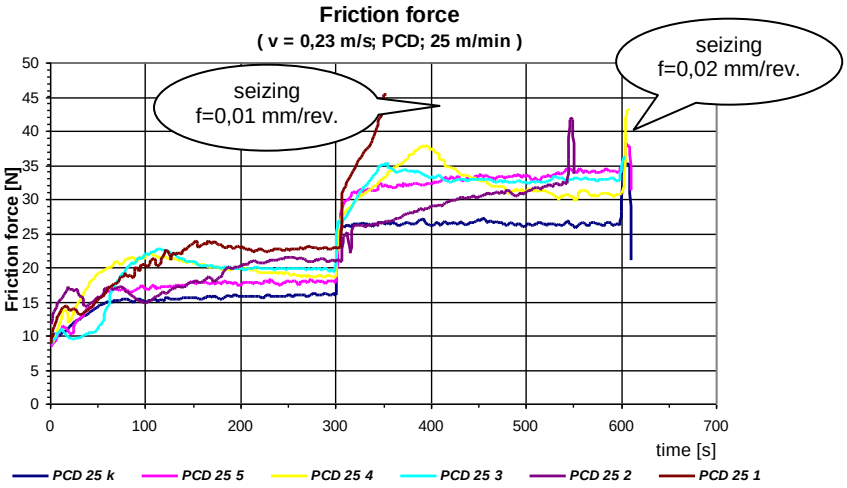


Figure 7. Friction force diagram (between steel/grinded and ceramic surface cut)

The friction force didn't show significant change on the surface grinded within identical load section. It set in a nearly constant value within short time (50 s). The grinded surfaces also seized with imposing maximum load. It is striking on the diagram that the friction force increased in such a great extent on the 2. load level on the ceramic surfaces cut with 0,01 and 0,02 mm/re. feed rate, that the surfaces seized in this section. The seizing ensued on the 3. load level on surfaces cut with greater feed rate.

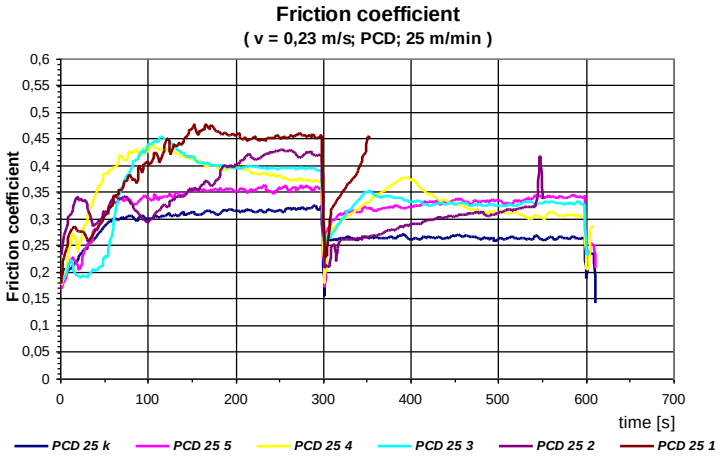


Figure 8. Diagram of friction coefficient (between steel/grinded and ceramic surface cut)

It is evident from the values of friction coefficients (Figure 8.) that higher μ values characterize the surfaces machined with smaller feed rate. The value of

friction coefficient on the surface turned with f=0,05 mm/rev. feed rate remained at nearly constant value in the sections tested similar to the grinded surface. Fluctuating friction coefficient, characterizing seizure can't be observed.

3.6. Results of complementary tests with thermo-camera

The thermo-camera exposure of the heat affected zone developed during cutting can be seen in Figure 9. The exposures prove unambiguously that because of the good heat insulation of ceramic at the tool point on the workpiece a heat ring is formed increasing significantly the thermal load of the tool point. In the lower two exposures the red colour shows the heat ring developed on the workpiece and the tool point, that is the hottest point.

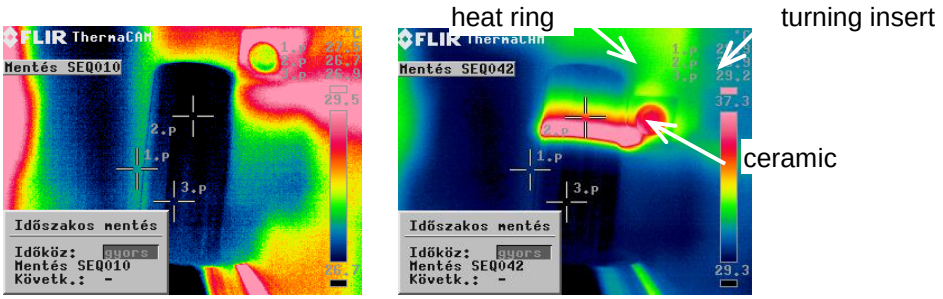


Figure 9. Exposures made by thermo-camera after cutting.

We have evaluated these tests only with comparing characters, I didn't take into account exact temperature results.

Conclusions

The demand increased in great scale to ceramics with special characteristics by satisfying environment protection standpoints, too. Different parts can be produced in average ±0,5-1,5% accuracy to gauge after sintering [Fritz, Schulze, (2007)]. So in many cases surfaces given has to be also machined in their hard (~1200-1400HV) condition. Economical use of semi-finished products during part production is only expedient in case of developing methods with greater material removal. The forming of microcraks is a serious problem of machining in hard condition. Reducing these presumes the more favourable machining beside the suitable material selection.

Based on our research work we have made the following scientific statement:

– We establish in the test field that in case of ZR40 (ZrO_2-MgO) ceramic basic material with the CBN and respectively PCD tool types having regular edge geometry chosen by me can be reached plastic (ductile) chip removal during turning in the cutting parameter field set (a=0,02 mm; f=0,01-0,02-0,03-0,04-0,05 mm/rev.; v_c=25-75 m/min.).

- We have proved with my measurings that in the upper range of the cutting speed of test field the polycrystal diamond tool is more favourable at turning zirconium-dioxide ceramics stabilized with magnesium it can be used with smaller cutting force than the cubic boron nitride tool. Furthermore I have established that the polycrystal diamond (PCD) resulted cutting force similar to the cubic boron nitride at lower (v_c=25 m/min.) testing cutting speed. The occurrence experienced at different frictional processes. I verified with thermo-camera exposures the forming and different heat rings taken place.
- We have proved with mathematical-statistical methods that the connection concerning the main cutting force $F_c = C_v \cdot f^x \cdot a^y \cdot v_c^2 [N]$ to be found in the technical literature in the tested parameter field (f=0,02-0,04 mm/rev,; a=0,02-0,04 mm; v_c=25-75 m/min.) can be extended to turning ZrO_2-MgO ceramic with PCD tool (with the edge geometry defined). We have determined with my measurings the values of the constant and exponents – C_v=477,183 - - x=0,1646- - y=0,4397 - - z=-0,2994 – and we proved that the connection can be applied with 95% probability with on the parameter interval tested.
- We established with digital picture – analyses based on microscopic exposure that at cutting with cubic boron nitride (CBN) type tool ≈ 57% surface damage (microcrack, shell like pitting, etc.) rose on the surface machined – concerning to unit zirconium-dioxide surface – already in that case when the depth of cut was only 0,02 mm, the feed rate however was only 0,04 mm/rev. Greater depth of cut and feed rate than this resulted still more unfavorable surface. The turning of ZR40 ceramic with this type of tool material is only suggested restricted.
- We have established beside the conditions of the test system that the turning with polycrystal diamond (PCD) turning tool as well as the grinding process resulted shell-like pittings, microcracks alike on the ceramic surface. The amount of surface damages in unit surface during turning with PCD approximately was ≤ 8,5% and differs in character and dimension the 8-9% surface damage experienced at grinding. The effect of difference influences the sliding characteristics which we proved with sliding measurings.
- We have proved with SEM exposures (2000x) that the wrinkles indicating the phase transformation formed to the effect of the great passive force at grinding don't appear on the ZrO_2 ceramic surface machined in case of using polycrystal diamond (PCD) turning tool ($\gamma=0^0$).
- We have proved with friction tests (block-on-ring tribological system, St37F grinded steel (R_a=0,8mm) "block" surface, without lubrication, ceramic "ring" specimen) that at high turning cutting speed (v_c=75 m/min.) and at small feed rate (f=0,01-0,02 mm/rev.) the friction resistance is smaller on the ceramic surface. The surface grinded has got smaller dimension but into the great number shell pittings the steel worn particles

seat quickly, which transforms the ceramic/steel friction connection into steel/steel characteristic friction, which increases significantly the adhesive component of the friction force.

5. Suggestions

The literature sources don't refer to the turning machining concerning ceramics, thus zirconium dioxide and aluminium oxide. Nowdays the grinding is the machining after the generally wide-spread sintering. The development of the finished- and semi finished products' manufacturing requires cutting ever more complex surfaces. The more economical machining of 3D-al surface requires the further development of tools with regular edge. The zirconium dioxide as basic material is suitable to machine by tool with regular edge deriving from lower ceramic hardness and from other characteristics so in case of piece production or small – and medium series production, at quick prototype production can become potential material alike. To that this should be ensued its cutting, and machinability characteristics has to be revealed.
The cutting with traditional machine-tool in the mechanical engineering is followed by high speed (HSC) "hard" machining controlled by computer (CNC). Nowdays it is possible to machine complex 3D-as shape and profile with the newest 5-axis HSC milling for example with machining-centre on such materials as copper, graphite, plastic and steel up to 70 HRC hardness.
Two main directions can be drawn up to carrying on resarches.

- Extending the applicability by changing the technological parameters,
- Technology optimization to be suitable to application-technic standpoints.

Further comprehensive elaborating of these fields can be realized with the co-operation and initiating of companies manufacturing ceramic materials and parts. The attainable further results entail with extending the technological database significantly.

6. References

H. Salmang, H. Scholze (2007): Keramik, *Springer Verlag, Berlin*

CeramTec AG Innovative Ceramic Engineering Medical Products Division: *Broschüre* 2008

A. H. Fritz, G. Schulze (Hrsg.) (2007):Fertigungstechnik. 8. Aufl *Springer Verlag, Berlin* p. 279.

König W., u. Klocke F. (1997): Fertigungsverfahren. Bd. 1: Drehen, Fräsen, Bohren. 5. Aufl . *Springer-Verlag, Berlin, Heidelberg*.

Horváth M., Markos S. (1995): Gépgyártástechnológia, egyetemi jegyzet, Műegyetemi Kiadó

Mészáros I., Szepesi D. (2005): Edzett acélok nagypontosságú megmunkálása II. *Gépgyártás* XLV. évf. 4. sz. pp. 20-25.

ABRASIVE WEAR OF HOT-DIP GALVANIZED MULTILAYER COATINGS

László SZABADI, Lajos PÉK, Gábor KALÁCSKA

Department of Maintenance of Machinery, Institute for Mechanical Engineering Technology

Abstract

The hot dip – galvanizing is a technology used to protect durably the surfaces of iron – and steel constructions which protective effect depends strongly on the thickness and density of the surface layer. Adhesive sliding conditions and deformation capacity have been studied in case of comparatively limited conditions but regulated comparative measuring data concerning abrasive resistance there aren't at disposal neither in technical literature nor in the database of coat – producing and of developing companies. These data are indispensable to develop technology improving the abrasive resistance.

Keywords

Abrasive wear, abrasion resistance, hot-dip galvanized layer, multilayer coating

1. Introduction

The aim of the research work was to determine with rapid comparative laboratory abrasion tests that the hot-dip galvanized coatings with different composition what abrasive resistance got as compared to each other. To carry out tests based on such measurings which represent by properly scientifical sound foundations the differences between the abrasive resistance of coatings hot-dip galvanized having different compositions. Our further aim is to present empirical properties in the function of frictional length, speed and the pressure of abrasive medium for developing technologies improving abrasive resistance.

The aim of our research work was to determine with rapid comparative laboratory abrasion tests that the coatings with different composition hot-dip galvanized what abrasion resistance got as compared to with each other. Mechanical loads also affect the products beside corrosive affects at industrial, agricultural and public place pavement grids at industrial filters hot-dip galvanized spreading nowadays [Tardy, Grega (1997); Antal (1999); Varjas, Kőszegi, Göblyös (1999)]. It is already a utilizing demand nowadays to develop

an abrasion and rubbing resistance coatings at surfaces hot-dip galvanized exposed to abrasion, to sand - and breakstone scatter [Jang (2010)]. There are not at disposal neither in technical literature nor in database of coat-producing and of developing companies comparative regulated measuring data concerning abrasion resistance [Antal (2002, 2007); Bárczy (1959)]. These data are indispensable to develop technology improving the abrasion resistance.

2. Testing method and equipment

2.1. Coated specimens tested

We have chosen S235JRG2 steel as specimen material used for abrasive tests, which is the most definite base metal for example of pavement grid. As this material is desoxidized with aluminium- and not with silicium – and the mechanism of layer developing is determined first of all by the amount of silicium to be in steel, therefore it can be hot-dip galvanizing outstandingly. We have taken into account at deciding the specimens dimension the tool form of the abrasion tester, as well as those positions and the geometrical dimension of the container containing the abrasive medium ensuring abrasion. Figure 1 shows the dimension of specimen galvanized.

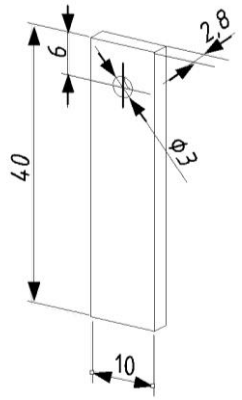


Figure 1. Specimen dimensions

Two types of coating were tested in the function of friction length, speed and pressure of abrasive medium.

Technigalva: The determining components are the zinc (Zn), aluminium (Al), lead (Pb) and nickel (Ni) of the zinc-bath. The coating was made by dry periodic technology.

Technigalva heat treated: The specimens were heat treated in order to the coating should have zinc-iron alloy phases in the total cross-section.

The abrasion resistance of coatings depends on its hardness, therefore it is needed to measure the micro-hardness. The micro-hardness tests showed that the Techigalva coatings had 47,1 HVM in average, while the heat treated coating had 106,6 HVM in average.
Reference Technigalva coatings resulted in average : HV M: 47,1
while the range:
R_{max} - R_{min} = 54 – 41,7 = 12,3

The heat treated Technigalva resulted in average: HV M: 106,6
while the range:
R_{max} - R_{min} = 114,2 – 89,5 = 24,7

Figure 2. and 3. show the SEM-pictures made from coatings, introducing the distinguished layers and EDS sampling areas. Certain layers, phases can be separated very good in the thousand-fold magnification. We have also carried out EDS (Electron Detector System) tests.
EDS spectra results are summarized in Table 1., a typical measured graphs – zeta phase - can be seen in Figure 4 and 5.

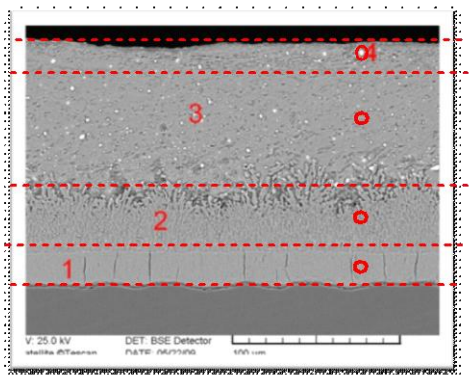


Figure 2. Technigalva layers
O : EDS measuring point

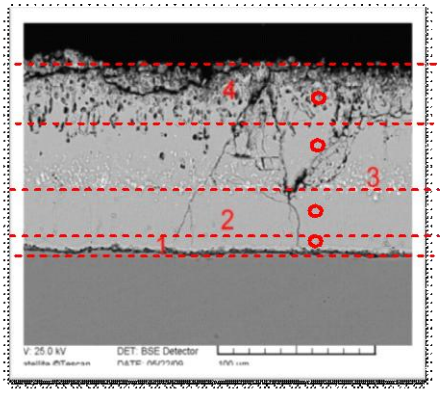


Figure 3. heat treated layers
O : EDS measuring point

Table 1. EDS spectra result of Fe, Zn and Pb values (%)

Technigalva „T”				Heat treated Technigalva „H”		
Layer	Fe (%)	Zn (%)	Pb (%)	Fe (%)	Zn (%)	Pb (%)
1	16,4	83,6	-	47,5	52,5	-
2	10,1	89,9	-	16,3	87,7	-
3	-	100	-	19,6	79,6	0,8
4	-	21,2	78,8	17,9	82,1	-

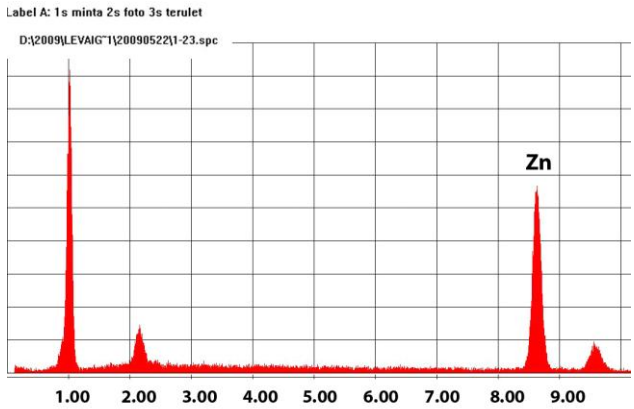


Figure 4. EDS spectra of Technivalva layer, phase Zeta (3)

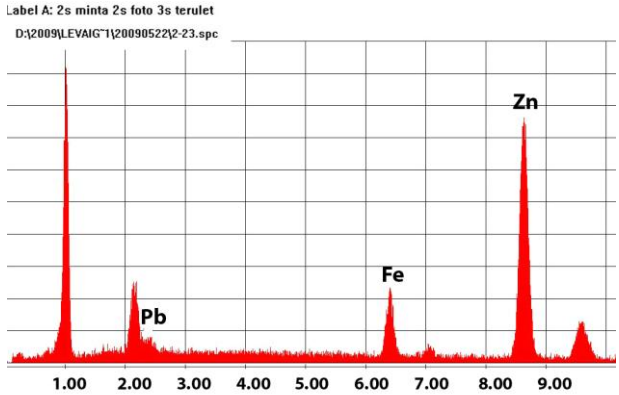


Figure 5. EDS spectra of heat treated Technivalva layer, phase Zeta (3)

2.2. Abrasive tribotester: modified "sand-slurry" equipment

The "sand-slurry" principle is well known in the VI. test category of tribological modeling. Great number of specimens can be measured at the same time in abrasive medium given as well as it can be well define but beside in different conditions with the abrasion tester.

Figure 6. shows the abrasion tester developed.

The working shaft can be found at the exit side of the worm-gear on which 3 pcs. arm cross-clamps can be found – in different heights related to the base plate. The specimens to be abraded can be fixed in suitable position on this. To one tool, to each arm 3 pcs. altogether 6-6 pcs. specimens can be fixed. Important characteristic of the cross-clamps is that the specimens can be fixed

with each other in 90^0 included position, at their sides in pairs altogether 6 various positions related to the centre of gyration.
The tools are set turned away to one another on the working shaft in top view the circle is divided to 30^0 sectors. The container containing the abrasive medium can be put into an outer container in case of demand, which can be filled with cooling-heating water to the thermal dynamics of measuring procedure can be regulated.
The abrasive medium was 0/8 OK – type ballast stone. Its average aggregation density in dry condition is 1,7 ton/m^3. There is no practically clay-sludge content as it is produced from washed, granulated gravels by knapping. The grain fraction is between 2 and 8 mm.

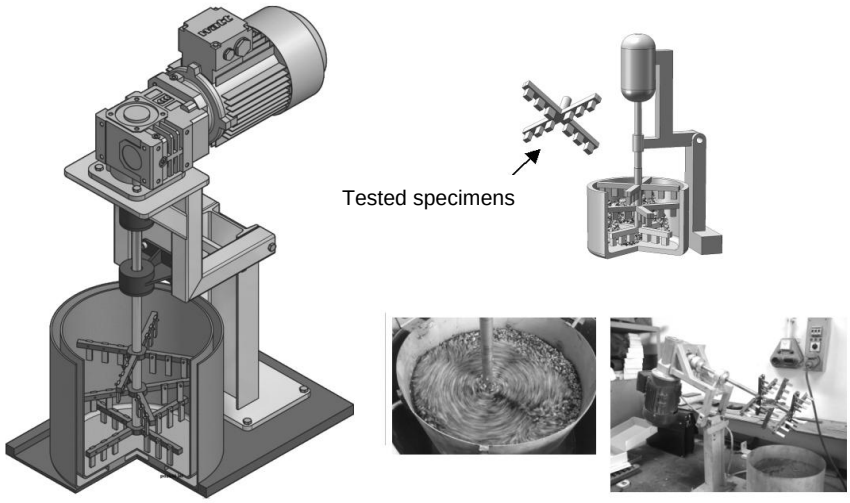


Figure 6. The "sand-slurry" tester with specimens mounted.

The specimens to be in different radiuses move with various different peripheral speed in the abrasive medium and the height position results different surface pressure rations. The abrasion tester makes possible exceptional complex evaluation in the function of these variables.

3. Results

3.1. The abrasion and speed connection

We have measured the abrasion of the 6 pcs. heat treated and 6 pcs. not heat treated specimens placed in all three levels (A,B,C) after seven various abrasion time. I have measured the abrasive wear as the decrease of the layer thickness in

seven different stage in time, thus getting different sliding distances for each sample in a given time. The surface pressures in standing (not rotating) condition:

p_1 (A level) = 153,83 Pa
p_2 (B level) = 521,82 Pa
p_3 (C level) = 902,41 Pa

Testing speed range: 14 – 40 m/min

The specimens covered different length with different peripheral speed moving on various radiuses. After averaging the abrasion results measured at each specimen, repeating three times the measuring series the same incline could be seen at the lines to be adapteable to the plot at heat treated and not heat treated specimens, too. Based on these the speed independence supposed were proved by mathematical – statistical methods, by covariance analyses at all three levels at heat treated and not heat treated specimens, too. The abrasion values of all specimens to be in the different levels can be presented with a single regression straight line.
This means that the abrasion values do not depend on the abrasive speed in the speed domain tested (Figure 7. and 8). There is no significant difference between the specimens moved with various speeds but placed at the same level.

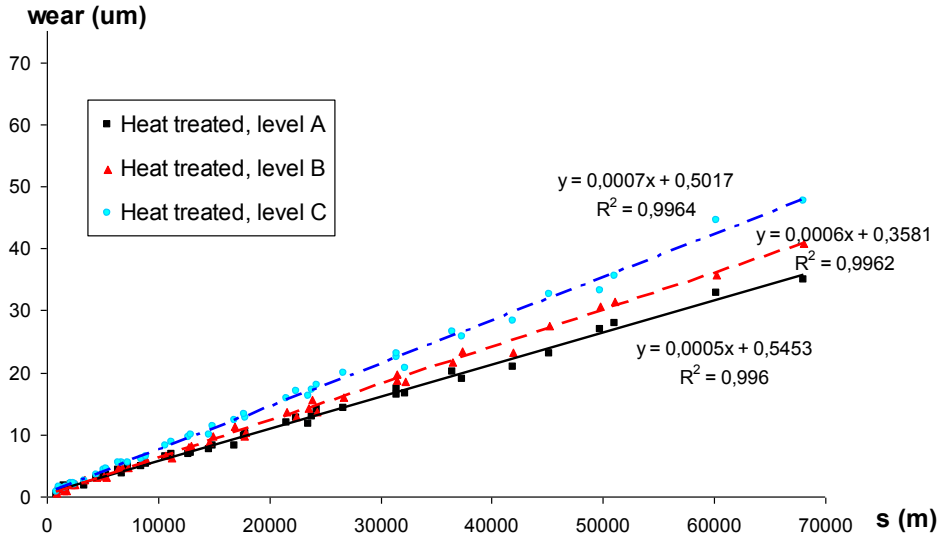


Figure 7. Abrasion values of heat treated specimens on "A", "B" and "C" level (h_1, h_2 and h_3 depth)

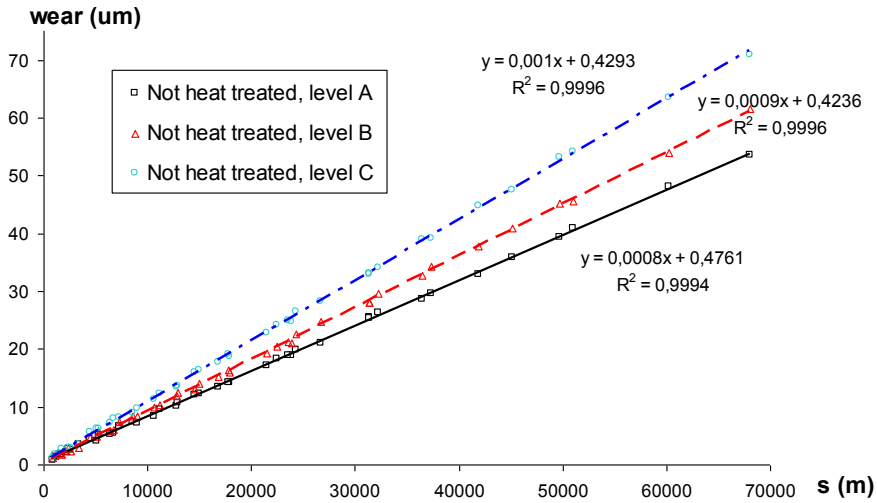


Figure 8. Not heat treated specimens abrasion values on "A", "B" and "C" levels.

3.2. Connection between surface load and abrasion

The specimens placed on the "A", "B" and "C" level get different surface loads because of this the abrasion values measured on "A", "B" and "C" levels has to be compared. According to the hypothesis the abrasion values depend on the load. It can be approximated with linear trend-line the appropriate data to different levels on the diagrams, the matching is close in each case. The data appropriate to different levels can be separated visibly however if there is significant difference between them it has to be examined by regression analysis. The worst case is where there is the smallest difference: this is the "A" and "B" level data of the specimens heat treated. The calculations carried out proved that there is significant difference between the abrasion values of specimens fixed on "A" and "B" level. This means that the amount of surface load has significant effect on the abrasion values in the system tested. In case of higher load the specimens have got higher abrasion.

3.3. Comparing the layer structure of coating and abrasion values

During the tests carried out the wear phenomenon occurs in different layers depending on the given stage – sliding distance - . Thus, the wear process reached Eta, Zeta, and Delta phases.

The abrasion measured as a resulting effect on the surface of specimens means the continuous decreasing of coatings with layer structure. In Figure 9. and 10. it can be seen how changes the percentage rate of decisive chemical elements in certain layers of heat treated specimens. The chemical composition's changing does not influence the abrasion intensity. According to data tested by EDS spectroscopy the percentage rate of chemical elements in the layers of not heat treated specimens is formed otherwise, but this compound does not influence the abrasion intensity of certain layers.

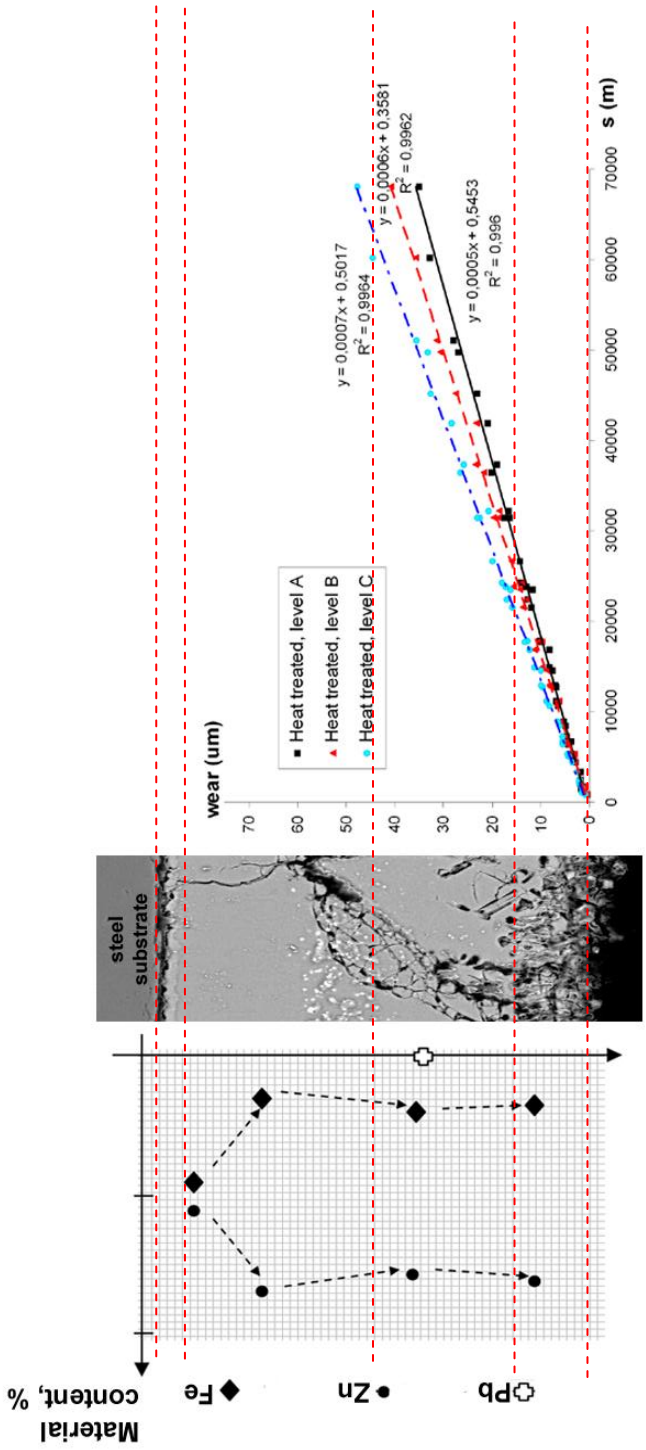


Figure 9. Heat treated specimens abrasion and the connection of layer structure.

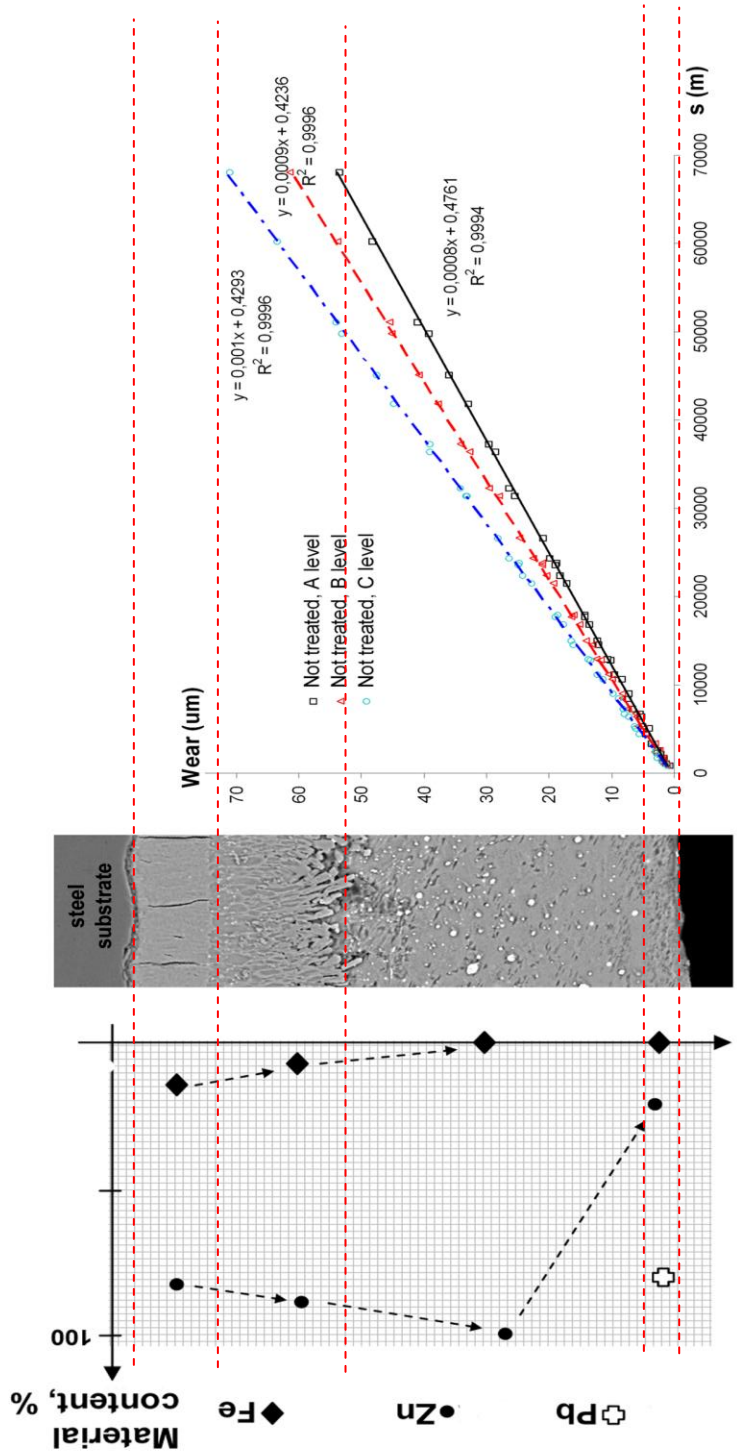


Figure 10. Not treated specimens abrasion and layer structure

4. Conclusions

We have verified with statistical methods that the intensity of abrasive wear doesn't depend on the frictional speed in the model testing system and conditions developed in case of multi-layer hot-dip galvanized coatings.
This statement is valid both to basic and to hest treated coatings in case of different compression conditions, too.
We have verified by my measurings that Rabinowicz's theory is also true in case of multi-layer hot-dip galvanized coatings, where the intensity of abrasive wear depends on the medium pressure and on the resistance of medium deriving from it. There is significant difference in the wear intensity of specimens to be in different pressure levels.
The heat treatment improves the abrasion resistance of the multi-layer Technigalva hot-dip galvanized coatings. The higher hardness resulted abrasive resistance increase in case of each speed and of pressure conditions in the testing system.
We have proved with SEM exposures and EDS spectroscopy that the heat treatment resulted different layer structures and compositions. We established that the gradient structure of heat treated layers differ substantially from the structure of basis coating.
We have established that the layer structure of coatings doesn't appear in the abrasive resistance, the abrasive resistance of certain layers don't differ. The linear abrasion dynamics of hot-dip galvanized coatings tested was independent from the layer – with given composition and hardness – to be in fictional contact with the abrasive medium.
I have established concerning the abrasive resistance that the resultant effect comes to full display of the multi-layer coatings having inner gradient structure. There was no effect of the inner gradient characteristic in the testing system onto the abrasion intensity measured but the resultant of the different gradient structure was the different abrasive resistance.
We have proved with my measurings that the heat treatment processes are suitable to modify the resultant abrasive resistance by changing the inner gradient structure.

7. References

Antal Á. (1999): A tüzihorganyzás alkalmazási területei és távlatai. Előadás VEKOR Konferencia. Balatonfüred. 1999. 04. 13-15.
Antal Á. (2002): A tüzihorganyzás technológia rövid története. Tüzihorganyzás. (1) 1.
Antal Á. (2007): Horganyzott termékek, évszázados emlékek. Tüzihorganyzás. (4) 4.
Bárczy Z. (1959): Fejezetek a Borsodnádasdi lemezgyár történetéből Borsodnádasdi Lemezgyár. Borsodnádasd.

Jang Y. (2010): The tribology and formability of zinc coated steel sheets subjected to different strain states. Deparment of Materals Science and Engineering, Case Western Reserve University. Dissertation. pp. 1-126.

Tardy P., Grega O. (1997): A bevonatolt lemezek várható fejlődése. DUNAFERR Műszaki Gazdasági Közlemények.

Varjas P., Kőszegi Sz., Göblyös B. (1999): A tüzihorganyzás hatása a hegesztési varratok tulajdonságai. HUNKOR '99 Konferencia Budapest.

5. INSTITUTE FOR SYSTEMS ENGINEERING AND MANAGEMENT

Associate Professor Dr. Miklós DARÓCZI
DIRECTOR OF THE INSTITUTE

Dear Reader,

Our Institute consists of three professional areas called departments:

- Department of Applied Management,
- Department of Engineering Economics,
- Department of Material Handling and Logistics.

The activity and the research fields of the Institute can be separated by these departments, and had the following main research activities during 2011:

Department of Applied Management:

- Development of manufacturing process with quality management techniques.
- The role of marketing in the innovation ability of the agricultural machinery manufacturers.

Department of Engineering Economics

- Developing the system of agricultural machinery management based on integrated enterprise management system.
- Technical and economic analyses of the farm machinery of arable farming.

Department of Material Handling and Logistics

- Developing logistical processes by mathematical modeling and simulation.
- Research of material handling machines with special respect to conveyors.

The results of our activities were published not only in Hungary but worldwide in different papers and conferences.

The educational activity is very important in our Institute. We have reviewed the curriculum of our Engineering Management Programmes. We developed different e-learning materials and several subjects taught also in English like

Quality Management, Enterprise Management, Farm Machinery Management, Project Management, Engineering Economics and Industrial Marketing.

More details about the Institute for Engineering Management are available at: http://mumi.gek.szie.hu/

ARENA MODULES FOR MODELING KANBAN-CONTROLLED MANUFACTURING

János BENKO

Department of Material Handling and Logistics, Institute for Systems Engineering and Management

Abstract

A company may reap the full benefits of kanban control only after determining an optimal or near-optimal system configuration. Finding such a configuration requires methods that can determine key performance measures, such as replenishment cycle, average fill rates, average inventory levels, etc. In industrial application the kanban-controlled systems are complicated multi-stages systems with restricted manufacturing capacity. Therefore, the existing analytical (mathematical) methods are only applied to determine just mentioned key performance measures fast and raw way. On the other hand the computer simulation is widely useable method and it may generally be used to analyze the true performance of a complex system

Keywords

Module, simulation model, Kanban system, Arena

1. Introduction

The production management approach Just-In-Time (JIT) gained worldwide prominence when the rest of the world noticed the increasing success of Japanese companies in the late 1970s and early 1980s. As one major operational element of JIT, the kanban control system became a popular topic in western research and industry. Manufacturing companies outside Japan began to use kanbans to control production and flow of material. The Kanban system was originally developed by Toyota Motor Corporation, and it was devised in 1954. At the beginning, the system had been introduced tentatively in one section of a factory. After that, the system had been expanded to the entire manufacturing system and was established as a form of the technique in 1970 (Toyota Motor Corporation 1988). Since then, the Toyota production system (TPS), which has only one element Kanban system, became well known worldwide. As for a Kanban system, it has attracted international attention because it differs completely from the traditional push production-control system. Kanban system, that is JIT manufacturing, adopts the pull production system where items are

processed at the upstream process, by receiving instructions from the downstream process.
A vast amount of studies on the Kanban system has been performed from the standpoint of production management. As for determining the number of kanbans, several approaches have been presented using techniques such as mathematical programming, Markov chain and simulation. Some analytical evaluation methods can be found in the literature, particularly for systems with a single product. Analytical (mathematical) evaluation methods are needed that can determine key performance measures quickly, even if these methods only approximate the true performance of the system. Finding the best configuration requires methods that can determine key performance measures, such as replenishment cycle, average fill rates, average inventory levels, etc. Computer simulation may generally be used to analyze the performance of a system, but to identify an optimal configuration, many different system variants may have to be evaluated.
In this study, a series of the modules are developed for developing simulation models on the flow-type multistage Kanban manufacturing system. This study focuses the Kanban system, i.e., the single-card and the dual-card kanban. By using the proposed modules for Kanban system, a simulation of multistage manufacturing systems can be developed and performed quickly and easily.

2. Classification of Kanban systems

The least complex variant of a kanban-controlled manufacturing system with multiple products is a system with a single multi-product manufacturing facility. Besides the production facility, the system contains a scheduling board, an output store for finished products, containers to store and carry finished items, and one set of kanbans for each product in the system.
Traditionally, a kanban is a tag-like card (kanban is Japanese meaning "card" or "visible record"). One kanban must be attached to each container in the output store. The number of kanbans is limited, restricting the maximum amount of finished items in the system. When a container is withdrawn, the accompanying kanban is detached from the container and placed on the scheduling board. Alternatively, the kanban may be detached when the last item is removed from the container (this is equivalent to using a fixed number of containers to limit the maximum inventory of a product). Also, removed kanbans may be put in a kanban collection box located in the output store before they are transferred to the scheduling board, either when a given number of kanbans has accumulated or when a specified amount of time has elapsed from the last transfer. A detached or "active" kanban authorizes manufacture of one standard container of the product indicated on the card. When a container has been filled with the prescribed number of items, the now "inactive" kanban is affixed to the container and the container is transferred to the output store.

Multi-stage kanban systems may be classified by the rules for transferring containers from the output store of one stage, say stage m, to the input store of the following stage, say stage *m*+l. At least four different set of rules may be found in the literature. In some systems, the output store of a stage is also the input store of the next stage. Consequently, there is no need for a transfer mechanism. In systems with separate output and input stores, the transfer of containers from the output store to the input store may be executed at different points in time. In *Table 1*, we summarize four different material transfer schemes. Each type is shortly explained in the following sections.

Table 1. Classification of Material Transfer Schemes

Output Store Stage *m* = Input Store Stage *m* + 1		Output Store Stage *m* ≠ Input Store Stage *m* + 1	
Type l	**Type 2**	**Type 3**	**Type 4**
Withdrawal immediately before start of production	Withdrawal immediately after activation of kanban	Fixed quantity, variable withdrawal cycle	Fixed withdrawal cycle, variable quantity
One-Card System		Two-Card System	

Type-1 material transfer. The output store of a stage is also the input store of the following stage, and material is withdrawn from the store immediately before start of production (*Figure 1*). This scheme has been labeled *late material transfer*.

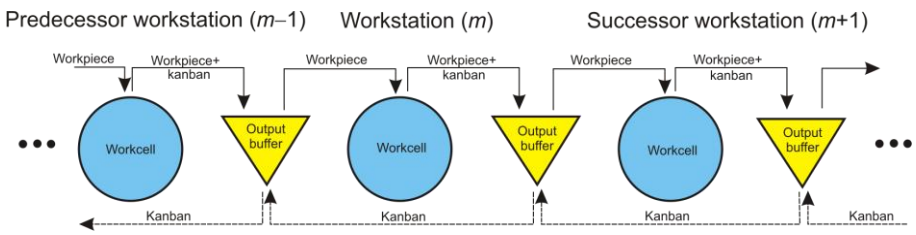


Figure 1 Single-card Kanban system (Type-1)

Type-2 material transfer. The output store of a stage is physically separated from the input store of the next stage. The material from the output store of stage m is withdrawn immediately after activation of a kanban in stage *m*+l (the active kanban authorizes the withdrawal of a container with input material). The kanban is attached to the container and both join the queue in front of the manufacturing facility of stage *m*+1. If the output store of stage m is empty upon activation of a kanban in stage *m*+l, then the transfer is delayed until the manufacturing facility of stage m completes a container with the appropriate parts. This scheme is called *immediate material transfer*.

Type-3 and type-4 material transfer. The output store of a stage is physically separated from the input store of the following stage, and an additional set of cards, called withdrawal, conveyance, delivery, move, or transportation kanbans, is used to organize the transfer of containers between the stages. In this study the name of these set of cards is *K-kanban*, and the kanbans organizing moving inside the production stage are called *B-kanban* (*Figure 2*). These systems are commonly referred to as two-card or **dual-card** kanban systems, in contrast to one-card or **single-card** kanban systems that only use a single-card type. A withdrawal kanban must be attached to each container in the input store of a stage. When a container is taken up for production, the withdrawal kanban is removed and put into a kanban collection box. Eventually, a carrier takes the withdrawal kanbans out of the box and moves to the output store of the preceding stage. There, he withdraws a full container for each withdrawal kanban in his possession, removes the regular kanban from each container, and attaches one of the withdrawal kanbans instead (the regular kanbans are often called production kanbans in two-card kanban systems). Then he carries the containers to the input store of stage *m*+l. The removed production kanbans are put into a box from which they are eventually collected by a worker who places them on the scheduling board of stage *m*. If the carrier finds fewer containers in the output store than he holds withdrawal kanbans in his possession, then he returns the extra kanbans to stage *m*+l and puts them back into the kanban collection box in the input store of stage *m*+l.

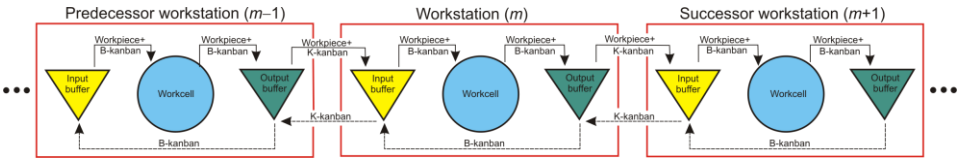


Fig. 2. Dual-card Kanban system

The point in time when the carrier removes the withdrawal kanbans from the kanban collection box is determined by one of two different schemes: (1) fixed quantity, variable withdrawal cycle (type-3 material transfer), or (2) fixed withdrawal cycle, variable quantity (periodic material handling, type-4 material transfer). In the first scheme, the carrier removes the withdrawal kanbans when a predetermined fixed number of cards has accumulated. The length of the withdrawal cycle, that is, the time between consecutive material transfers, may therefore vary. With the second scheme, the carrier removes the kanbans periodically, following a predetermined fixed schedule. Here, the withdrawal cycle is fixed and the number of cards may vary. Note that type-3 material transfer is equivalent to type-2 material transfer if the fixed withdrawal quantity is set to one.

3. Characteristics and application of the developed modules

The Arena Professional Edition is an advanced, hierarchical structured simulation system which provides an interactive environment for building, graphically animating, verifying, and analyzing simulation models. Due to the hierarchical structure with the Professional Edition, you can design a unique Arena template that is specific to your particular project, company, or industry. With the Professional Edition, you can create complete simulation building blocks, called modules and after translation modules can be integrated into template. In our case we will develop different type of kanban modules and will integrate them into template called Kanban.tpo.
The restricted extent of this study does not make possible the detailed presentation of the steps of the development and the background logic of the modules therefore we devote more time to outline the characteristics and application opportunities of the modules. The self-developed Kanban Process template consists of six modules (*Customer_S*, *Customer_D*, *Workstation_S*, *Workstation_D*, *Supplier_S* and *Supplier_D*) which are applicable for modeling both of single-card and dual-card Kanban systems. In the name of the modules the extension *S* and *D* refer to single-card and dual-card system. The common feature of the modules, that not only the train of thought of their development, but the logic of their application similar. To the function of all models need a customer and a supplier station that can be modeled with *Customer* and *Supplier* modules. According to the number of the production stages, required number of workstations can be inserted between the supplier and the customer stations. We suppose that a batch of parts or jobs is flowed in the model between stations and we do not examine its combination. Stations, rows and resource pictures are attached to the modules to utilize the animated opportunities of Arena model. On paths (*Route*) between the stations (*Station*) can be displayed the motion of the entities. The pictures of resources (*Resource Picture*) show the state of the resources (*Busy, Idle, Inactive*). The *Queues* represent the length of the queues inside the modules. The entity pictures that symbolize the motion of the workpieces and kanban cards can be chosen optionally. The time parameters like interarrival time, or route time may be deterministic ones and stohastic. To the last cases the available distribution functions can be picked from pull-down lists.
In the models the modules named *Customer_S* and *Customer_D* generate customer demand and provide the pull effect. The completed dialogue windows of these modules are displayed in *Figure 3*. The most important parameters of modules: the name of the stations, the interarrival time of the customer, the preparation time of an ordering, the number of the resources, predecessor station name, move time between the last workstation (predecessor station) and the customer station. The picture to animate kanban cards can be given optionally.

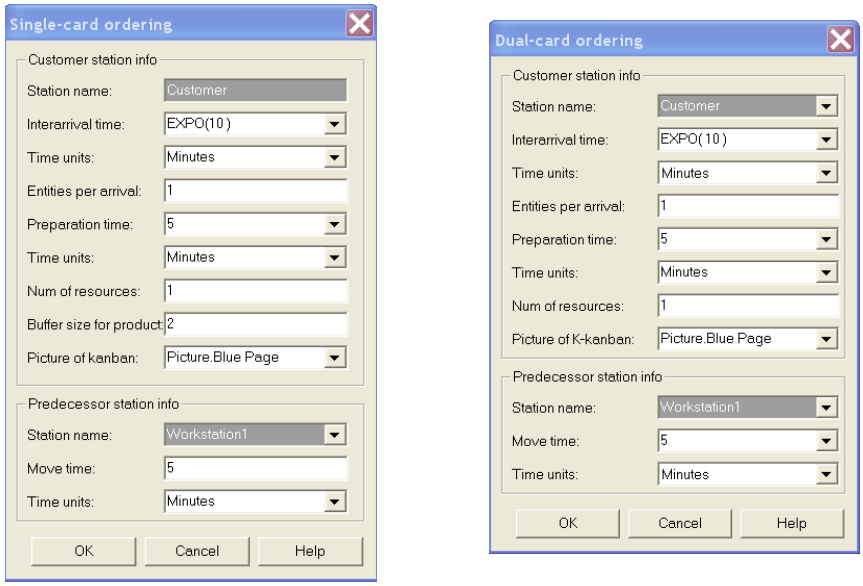


Figure 3 Dialog boxes of *Customer_S* and *Customer_D* modules

Workstation_S and *Workstation_D* modules which simulate the workstations significantly differ from each other regarding their complexity and their data requirement. In the single-card system the output buffer of the production stair is the input buffer of next production stair (*Figure 1*). Opposite this in a dual-card system the output buffer of the production stair separates from the input container of the next production stair physically and all production stair has own input and output buffer (*Figure 2*). The motions between the input and the output buffer are generated by the inside kanban (*B-kanban*) and between the output buffer and the input buffer of the next workstation are managed by the outside kanban (*K-kanban*).

Dialog window of single-card workstation module (*Workstation_S*) is depicted in *Figure 4*. Beside the station name, which is module and station identifier, the additional input parameters are the next. The process time and the number of resources determine the capacity of the workstation. The buffer size for material defines the initial inventory of the input buffer, which simultaneously is the output buffer of the predecessor station. In the dialog window the identifier of the previous and a next station are obligated field, and the transportation time can be given optionally. The detached or "active" kanban cards are streamed onto the previous station and the inactive kanban cards together with workpieces (products) are moved onto the next station.

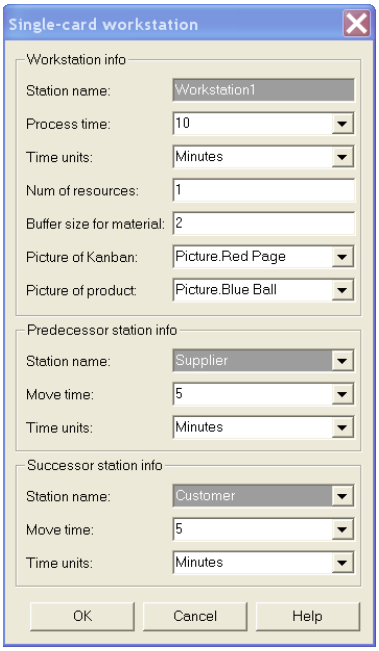


Figure 4 Dialog box of *Workstation_S* module

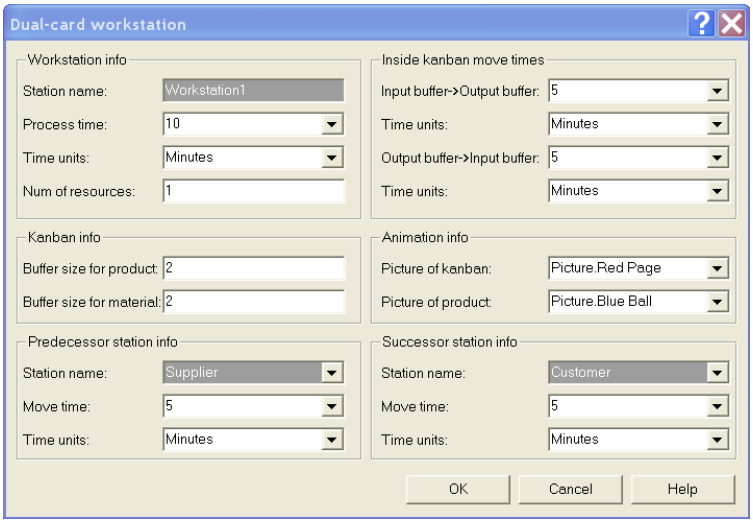


Figure 5 Dialog box of *Workstation_D* module

The dual-card workstation (*Workstation_D*) module has more input parameters (*Figure 5*). Beside the listed parameters of the single-card workstation (*Station name*, *Process time*, *Num of resources*, *Buffer size for material*, information of the predecessor and successor station) here has to grant the buffer size for product and the move times of the inside kanban yet.

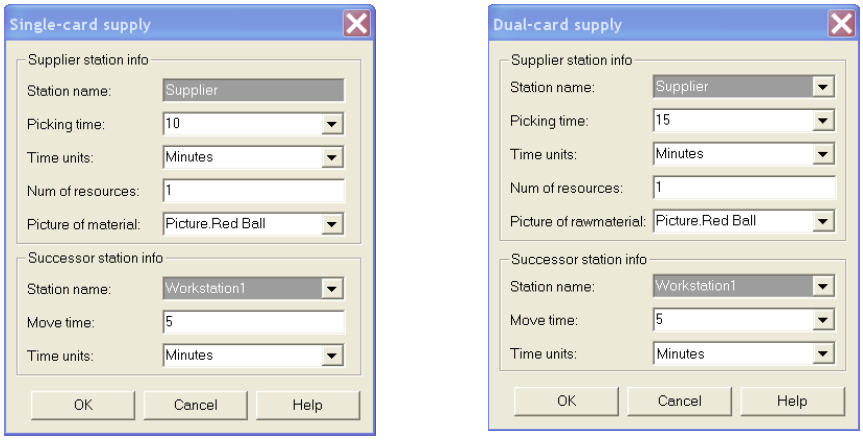


Figure 6 Dialog boxes of *Supplier_S* and *Supplier_D* modules

To the animation we may elect optionally pictures for kanban and product in both modules. To a dual-card module, logically two stations an input and an output buffer are attached which can be animated.

Supplier_S and Supplier_D modules provide the material supply and considering their function and their parameter claim are totally identical. The dialog windows of the modules are shown in the *Figure 6*. The input data of the modules: the name of the station, the picking time, the number of resources (employees or devices) and the data of the next station. We consider the buffers of these stations infinite big one, which means there is always sufficient raw material in these storages, so the processes never starve. The picture of the raw material entity leaving the station is optional.

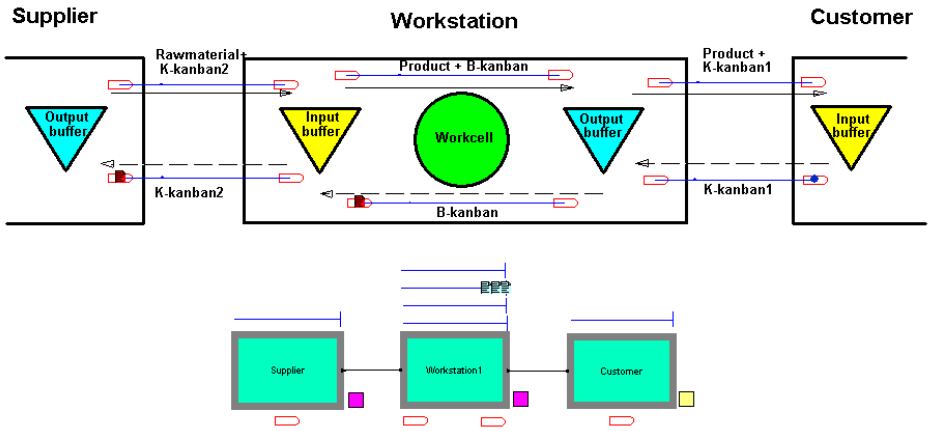


Figure 7 A model of dual-card single-stage Kanban system with animation

Let's consider a demonstration for the application of the modules a simple one dual-card single-stage Kanban process. Customers arrive according to an exponential distribution with 10 minute average arrival intervals into the system. The transportation time of outside kanban (*K-kanban*) and inside kanban (*B-kanban*) are 5 and 2 minutes. The picking time is 15 minutes, the process time is 10 minutes on the workstation and the preparation time is 5 minutes on the supplier station. The length of the simulation time is 100 hours in the example.
Figure 7 depicts the flowchart model and the animation of the example. The flowchart consists of the module instances with queues (half-line above the modules), with resources (squares on right hand side of the modules) and with stations (shapes under the modules). In the animation the buffers, the stations, the routes connecting the stations and the entities moving between the stations (kanban cards and workpieces) appear.

User Specified

Output

Output	Value
Beszallito.Anyag beszallitasok intenzitasa db per min	0.06633333
Beszallito.Anyag beszallitasok szama osszesen	398.00
Beszallito.Beszallitasra varakozo anyagrendelesek atlagos szama	0.3318
Megrendelo.Rendeles teljesites intenzitasa db per min	0.06683333
Megrendelo.Teljesitett rendelesek szama osszesen	401.00
Megrendelo.Term ekrendelesek intenzitasa db per min	0.0987
Megrendelo.Term ekrendelesek szama osszesen	592.00

Figure 8 Extract from the User Specified statistics

The primary objects of the simulation are the examination of the behavior of a given system, the localization of the bottle-neck of a process and the determination of the resource utilization, etc. The answers onto these questions can be found in the standard (*Entity*, *Queue*, *Process*, *Resource*) statistics being made in parallel with the simulation and in the user defined (*User Specified*) statistics. Statistics defined to Kanban modules inform the intensity of the flows between the stations, the length of the queues inside the stations, number of the entities leaving from and coming to the stations. The extract of these statistics belonging to the sample example may be studied in *Figure 8*.

4. Results

The new modules described in the study are applicable for modeling single product, multi-stages, single- and dual-card Kanban systems. Using the modules complex multi-stages simulation models can be built quickly. On one hand the

function of existing systems can be analyzed on the other hand a new systems can be planed with the developed models. The modules of multi-product Kanban systems may be developed similar way with the mentioned concepts and the modification of the described module logic.

References

Arena Professional Reference Guide, Rockwell Software Inc., 2000.

Akturk, M. S., and F. Erhun. (1999): An overview of design and operational issues of kanban systems. International Journal of Production Research 37 (17): 3859-388 1.

Benkő J. (2011): Logisztikai folyamatok szimulációja. Egyetemi jegyzet, Szent István Egyetemi Kiadó, Gödöllő.

Huang, P. Y., L. P. Rees, and B. W. Taylor. (1983): A simulation analysis of the Japanese just-in-time technique (with kanbans) for a multiline, multistage production system, Decision Science 14: 326-344.

Kelton, W. D., R. P. Sadowski, and D. T. Sturrock. (2004): Simulation with Arena. 3rd ed. New York: McGraw-Hill.

Krieg, G. N.(2005): Kanban-Controlled Manufacturing Systems. Springer-Verlag Berlin Heidelberg.

Mitra, D., and I. Mitrani. (1990): Analysis of a kanban discipline for cell coordination in production lines. I. Management Science 36: 1548-1 566.

MANUFACTURING PROCESS DEVELOPMENT WITH 5S AT DIFFERENT TYPES OF PRODUCTION

Adrienn GODA[1], Viktor MEDINA[1], László ZSIDAI[2]

[1] Department of Applied Management, Institute for Engineering Management
[2] Institute for Mechanical Engineering Technology

Abstract

Agricultural machinery manufacturers have to use the specific instructions to fulfill customer requirements. The defective products are the most significant contingency for the factory and the final customer as well. The aim of quality management is to minimize the number of wrong components, products and agricultural equipments, and to maintain good quality with low costs and high efficiency, to conserve and expand markets for their products.
Quality management methods can be classified from different aspects. The professional literature of quality management does not include a widely accepted categorization.
In our present paper we are going to describe an application of a quality management technique for process development at cob cracker adapter manufacturers in Hungary. This analysis is intended to reveal the execution level of the 5S technique in the case of the individual, the series and the mass product manufacturers.

Keywords

Manufacturing, quality management, 5S, lean techniques, process development

1. Introduction

Nowadays the high quality and the good price are more and more appreciated. All companies have to produce with excellent quality, affordable price, flexible and short delivery deadlines. For these reasons it is even more important to increase the efficiency of the manufacturing. Improving the work conditions and developing a productive work environment give good possibilities to produce more efficiently (Parányi, 1999).
To improve the productivity there is a popular and useful technique, called 5S. Hiroyuki Hirano thought that items being organized are a key factor of the productivity. A key aspect of his approach was that the often-needed items must be stored in the most accessible location and they must be returned to the correct

location after use. He used 5 Japanese words (beginning with "S") to define the 5 pillars of the concept (Hiroyuki Hirano, 1995).
Later, this 5S was adopted to the English language, and translated to 5 English words beginning with "S" (see table 1). In the professional literature of quality and lean management there are more translations and version for the English 5S. The most widely used 5 names are: sort, straighten, shine, standardize and sustain. The 5S is a popular lean technique because of its usability.

Table 1. Cross table of correlation between marketing department and innovation performance (Source: own editing on the basis of Kalkowsky, 2004)

„5S" in Japanese	„5S" in Englsih	
Seiri	Sort	Clearly distinguish needed items from unneeded and eliminate the latter. Define exactly the sort of item's use.
Seiton	Straighten Set in order Stabilize Simplify	Keep needed items in the correct place to allow for easy and immediate retrieval. Subtitling the items for the easy identification.
Seiso	Shine Sweep	Keep the workplace neat and clean.
Seiketsu	Standardize Systemize	The activities of the first 3S made habitual. Standardize these activities.
Shitsuke	Sustain	Maintain established procedures and apply them in the corporate culture.

In the daily work of a company, the routines that maintain organization and orderliness are essential to a steady and efficient flow of activities. The 5S method encourages workers to improve their working conditions and helps them to learn to reduce waste, unplanned downtime, and in-process inventory (Productivity Press, 1996).
According to Vollmer's definition, the 5S is a precise and disciplined mean with its strict and critical management approach (Vollmer, 2004). So the 5S technique seems to be a useful tool for improving the productivity, but it should be analysed whether there is any difference among various types of production? So we should examine the 5S technique in the case of individual, serial and mass production types.

2. Methods

The basic objective of the research is to explore and analyse the possibilities of the manufacturing process development with the 5S quality management technique. Our null hypothesis was that there are some differences among various types of productions based on the application of the 5S method in the

corporate practice. To confirm our hypothesis we examined three types of cob cracker adapter manufacturers in Hungary. We made the classification based on their production type. We differentiated individual, serial and mass production.
The production is considered **individual**, when the products are produced individually or in very small quantities with a wide range of, and the repetition of the same products is also irregular (Dudás, 2000). The production type in which a certain amount of product will give to the production in the same time is called **serial production**. It is characterized by larger quantities and same manufacturing conditions in the entire series.
The **mass production** means a constant and uninterrupted production with tight range of production types. The process may be described as mass production, if the production of the structurally identical products is a long time constant, and other products aren't wedged in the process. During mass production in the workplaces are made the same permanent jobs, and allows the extensive application of the flow-system- manufacturing and the automation.
For the examination of the application of the 5S technique in the three production types we made a questionnaire. This questionnaire has 25 statements and was rated by the cob cracker adapter manufacturers, whether the statement is typical of the company or not. The sentences concern to the 5 elements of the 5S. It was declared 5-5 statements about sorting, straightening, shining, standardizing and sustaining. The representatives of the companies could rate them with a scale from 1 to 5. The number 1 means that the declaration is false, number 5 means that it is very representative to the enterprise. We added the 5 rates by elements and every "S" could receive 25 points as maximum. Every company could receive as maximum 125 points.

3. Results

The first step of the 5S is eliminating the unnecessary items, and prioritizing things per requirements, keeping them in easily-accessible places, and everything else storing or discarding. Although this is the initial part of the method this element received the worst rates (see figure 1.). It is interesting that the best rates were given for the sustainability. The total points were 75, which mean 25 multiplied by the 3 types of production. This high evaluation maybe can be explained by that way, that it is easier to maintain a lower level then a higher one. It is supportable by seeing the figure 1 in which we can see that the other cases generally it was given between 3 and 4 points for the statements.

The statements about sorting were:

1. There isn't any unnecessary item at the workplace.
2. The store places and routes are well marked.
3. There aren't any unnecessary working tools at the workplace.
4. The products are stored well ordered and marked
5. The places of the working tools are clearly signed

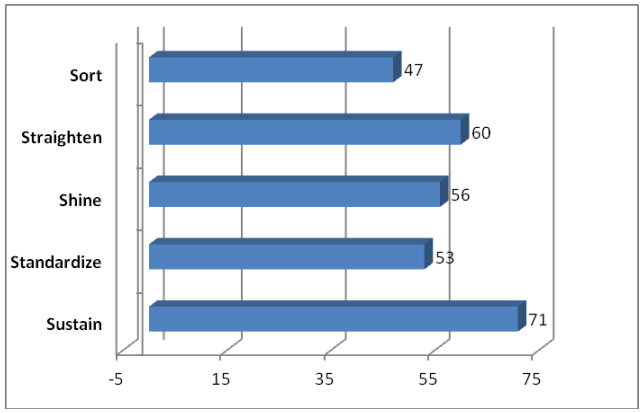


Figure 1. Cumulative points of 5S by elements (Source: own research results)

The sorting elements must be developed in all cases. Even in the case of serial production there are many possibilities to develop (see figure 2.). In the case of mass production sorting gives the best chance to progress.

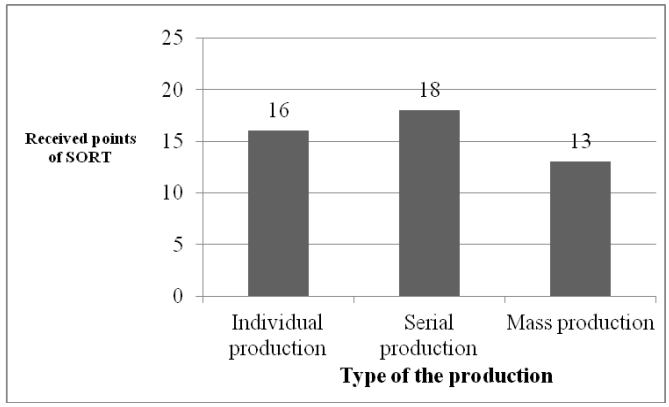


Figure 2. Received points of SORT by types of production (Source: own research results)

Items should be arranged in a manner that promotes efficient work flow, with equipment used most often being the most easily accessible. Workers should not have to bend repetitively to access materials. This 5S element has received the second best points (see figure 1.). In the figure 3 is shown, that this case the individual and serial productions rated more representative the statements:

1. The materials for the production are stored on their marked place.
2. The tools after use are resettled on their place.
3. The workplace makes an ordered impression.
4. The documents are ordered and easily available.
5. The rules of storage are clarified and marked.

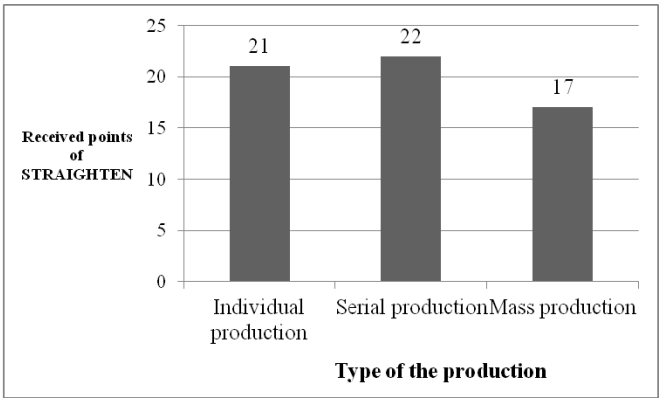


Figure 3. Received points of STRAIGHTEN by types of production (Source: own research results)

Cleaning the workspace and all equipment, and keeping them clean, tidy and organized is quite important. Based on the figure 4 it seems that these requirements are easily fulfilled by the individual manufacturers. But in the case of mass production there are more difficulties. The statements of straightening were in the questionnaire:

1. The equipments are in adequate state.
2. The maintenance of machines is suitable.
3. The workspace it clean and tidy.
4. The tools are clean and tidy.
5. The products waiting for delivery are available for cleaning?

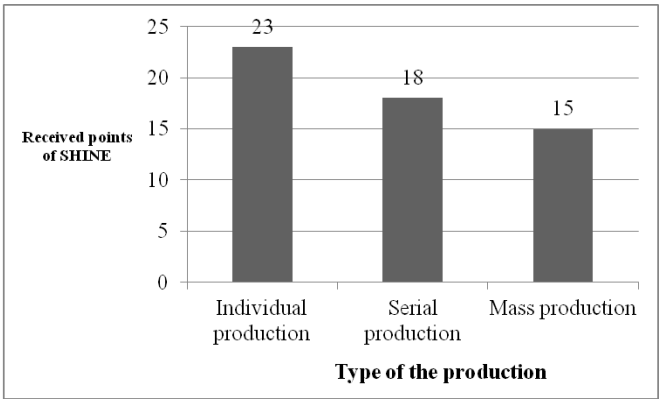


Figure 4. Received points of SHINE by types of production (Source: own research results)

Work practices should be consistent and standardized. All work stations for a particular job should be identical. Everyone should know exactly what his or her responsibilities and tasks are. This 5S element is very important in the case of

mass production, where the standardization has a special meaning, because they work with huge quantities and it is very important to that all employees doing the same job should be able to work in any station with the same tools that are in the same location in every station. In spite of this in the figure 5 we can see, that this element has received the worst rating in the case of mass production. It is also interesting the outstanding performance of the individual production. The statements of this element were:

1. Storage places are clean and ordered
2. The environment of the equipment is clean and transparent.
3. The manufacturing tools are ergonomically ordered.
4. The control of the machines are regular and methodical
5. The operational environment is secured.

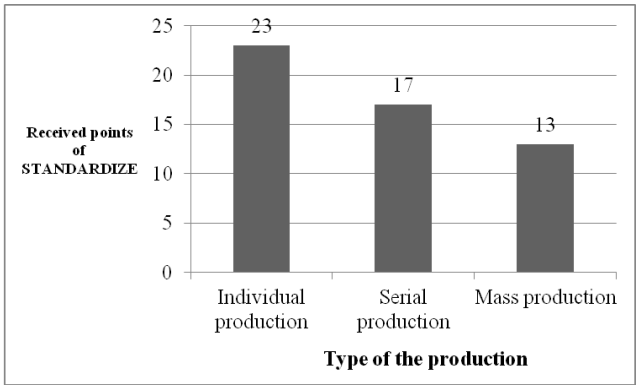


Figure 5. Received points of STANDARDIZE by types of production
(Source: own research results)

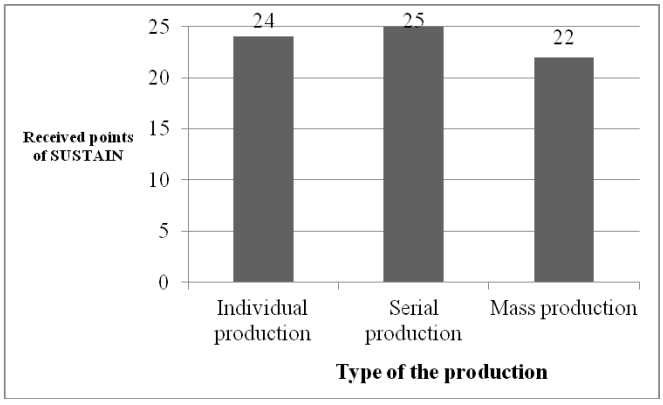


Figure 6: Received points of SUSTAIN by types of production
(Source: own research results)

The main task of the last element of 5S it to maintain and review standards. Once the previous 4 elements have been established, they become the new way to operate. In the figure 6 can be seen that there isn't too much problem with the discipline at the 3 types of production. All of them rated the statements with quite high points:

1. The prescribed coveralls are worn based on the specifications.
2. The smock of the workers is suitable.
3. The technological discipline is appropriate.
4. The instructions and rules are observed during the production.
5. The work discipline of the employees is proper.

We made also a Cross Tabulation Analysis and a Chi-square test to prove statistically if there is any correlation between the type of production and the application of the 5S. The independent variable were the rates of the statements, the dependent variable was the type of production. The result of the analysis rejected a statistically proved correlation between the two variables.

4. Conclusion

After evaluating the results we can conclude that there are differences among the different types of productions according to the application of 5S. The individual production has received the best evaluation. It scored 107 points from the maximum 125. The serial production received the second place with its 100 points. The mass production scored only 80 points, which means that it has to develop the most.

The explanation of these results can be that in the case of individual and serial production it is easier to organize the production. But in the case of mass production there is a huge potential to increase the efficiency of the production with the help of the 5S technique.

To take advantage of this potential and improve the productivity we propose to maintain focus on the new way and do not allow a gradual decline back to the old ways. Other important task of the manufacturers is that while they think about the new way, they also have to think about yet better ways. And when an issue arises such as a suggested improvement, a new way of working, a new tool or a new output requirement, it must be reviewed the first 4 elements of 5S and make changes as appropriate.

In the other two types of productions there is also possibility to elevate the efficacy of the production with a better over thought introduction of the 5S method.

Although there is an empirical difference between the type of production and the application degree of the 5S method, there isn't a statistically approved correlation between them. In spite of this the examination showed us some differences, and we can maintain our null hypothesis, that there are some

differences among various types of productions based on the application of the 5S method in the corporate practice.

5. References

Parányi György (1999): *Quality – economically*, Műszaki Könyvkiadó, 1999. 194. o. (in Hungarian).

Hiroyuki Hirano (1995): 5 Pillars of the Visual Workplace: The Sourcebook for 5S Implementation. Productivity Press p. 353 ISBN 978-1563270475

Kalkowsky, M. (2004): *Lean Operations,* Produktion, 2004. 46. November 11. p. 36.

Productivity Press Development Team (1996): 5S for Operators: 5 Pillars of the Visual Workplace (Portland, Oregon: Productivity Press, 1996)

Vollmer, A. (2004): *Mit Lean Production Zukunft sichem,* =Produktion, 2004. 49sz. dec.2. p.30. (in German).

Dudás I. (2000): Manufacturing I., Basics of manufacturing. Miskolc: Egyetemi Kiadó 205-208 p. (in Hungarian).

THE ROLE OF THE MARKETING IN THE INNOVATION ABILITY OF THE AGRICULTURAL MACHINERY MANUFACTURERS

Árpád BAK, Viktor MEDINA

Department of Applied Management, Institute for Engineering Management

Abstract

The innovation is a basic condition of the competitiveness. The advantage of the innovation could be enforced only, if the new products and technologies are supported by suitable marketing abilities. Therefore the task of the innovation marketing is not only the marketing of innovation, if not supporting the management to satisfy the rational and right market demand, to increase the efficiency of the strategy with the communication of the information, and to increase the social acceptance of the results and diffusion of the innovation. The marketing activity of the SMEs is a weak point in most cases. Based on our examination the situation is similar in the case of the agricultural machinery manufacturers, moreover the weak marketing is one of the most critical elements of their functioning. Our previous results show the manufacturers made efforts for costumer oriented development. The manufacturers told us, that the opening of new markets and the increasing of the market shares were one of the most important goals of their innovation. It was mentioned also the costumer is in the focus of their innovation. In our paper we present the marketing activity of the Hungarian agricultural machinery manufacturers and the examination of the relations between their marketing activity and the innovation efficiency.

Keywords

Innovation, agro-technical development, industrial marketing

1. Introduction

The current production of the Hungarian agricultural machinery manufacturing sector, which used to see better days, lags behind the production of the previous years to a great extent. The organisational structure of the Hungarian agricultural machinery production has totally been transformed, primarily regarding its ownership structure. The general problem of this sector is that they can only spend slight amounts on development an innovation relative to foreign-owned concerns. As a consequence, loss of market is not surprising as a bit more than one-quarter (26-27 per cent) of the current total domestic market turnover derives from domestic manufacturers. The extent of market loss and the general

situation of the Hungarian agricultural machinery manufacturers justify that the present of this sector must be dealt with by searching the ways-out of the crisis and make steps to develop. Before the change of the regime only 27 agricultural machinery plants operated mostly "embedded" in the system of the Hungarian "agri-business". Due to this fact (among others), 60 per cent of the requirements for agricultural machinery in the country were covered by these plants at a more advanced standard than the average of the former Comecon countries. During the past 15-20 years the organisational structure of the Hungarian agricultural machinery production has totally been transformed. Generally, the machine manufacturers operating as small-or medium-sized enterprises appear on the market with "separate" products usually not developed by themselves. Consequently, they are not price-setters, rather price takers. The product line of the companies that are successful in the international competition primarily consists of mass-produced and highly automated products. The Hungarian agricultural machinery manufacturers-partly due to their size- are not able to mass-produce in such an extent that they could compete with the West-European, American and Asian companies of huge capital power either in productivity, price or product range. A drastic innovation wave could mean a break out of this situation. Regarding innovation, the Hungarian agricultural machinery manufacturers also significantly lag behind as they can only spend slight amounts on development relative to foreign-owned concerns. As a consequence, loss of market is not surprising as a bit more than one-quarter (26-27 percent) of the current total domestic market turnover comes from domestic manufacturers.

The conclusions of our paper are based on the examination results of questionnaires and in-depth interviews that were carried out at 35 Hungarian agricultural machinery manufacturing companies. The characteristic features of the companies that were involved in the examination reflect the situation of the entire sector in Hungary properly. In our paper first of all the method of the empirical research is presented where the structure of the questionnaire used in the research and the process of data recording and processing are shown in details.

2. Methods

The basic objective of the research is to explore and analyse the innovation activity of the Hungarian agricultural machinery manufacturers, its results and influencing factors. Finally our objective is to have a general idea of the innovation activity of the organisations involved, the special features of innovations, the partners taking part in the processes and the impact of innovation on the general situation of the companies through our examinations. Besides the brand-new or significantly developed products and technological procedure innovations, organisational features, marketing activity and the environment of the innovation are also considered. The questionnaire serving as

the basis of primary research embraces three years, from 2007 to 2009. According to the estimations of experts the number of agricultural machinery manufacturing companies is between 160 and 170 in Hungary. (A great part of the enterprises are involved in more than one activity: a lot of predominantly small enterprises are also engaged in other activities besides machinery production so that is why it is difficult to define the actual number of 'agricultural machinery manufacturers' exactly). Most of the organisations that are subject to our analysis are small enterprises whose annual revenue does not reach one billion HUF. As there was not an available list on all the companies on the basis of which a pattern of probability could have been compiled, the companies that could be drawn into the research had to be defined in another way. To find the companies necessary for carrying out the questionnaire, the address list of MEGOSZ (National Association of Agricultural Machinery Manufacturers) served as a basis and the heads of this professional organisation were also consulted.

Sample-taking cannot be regarded representative. However, during the research it was not our objective to draw conclusions that can be generalised for the basic population. Our basic objective was to give a thorough examination of innovation activity and to achieve it. We tried to select the organisations regarded to be suitable on the basis of preliminary professional considerations. As such a thorough examination dealing with the innovation activity of agricultural machinery manufacturers was not carried out in the past 25 years on a national level. We consider our research is to resolve discrepancies in the professional field.

In compliance with the general methodological requirements first of all some pilot questions were asked on the basis of which the questionnaire was finalised. Data recording took place between March 2010 and September 2010. The duration of in-depth interviews was various, typically 90-100 minutes per interview. A positive feature of them was that data providers mainly come from the senior management (chief executive officer, head of production or technical manager). In this way first-hand information on the general situation, actual projects and strategic plans of the organisation involved was gained besides the reliability of data. The atmosphere of the interviews was typically of honesty and intimacy. Some of our interviewees have already expressed their enquiry in our results. The questionnaires compiled on the basis of the interviews and sent out by post were also accompanied by a guide to filling in. A kind of evaluation of our preliminary work is that all the responding organisations gave answers that could be assessed. The statistical processing of data recorded by the questionnaires was carried out by using SPSS 13.0 programme.

3. Results

The success of the innovation is decided on the market, therefore it is important that the product development how it is prepared and followed by market

research and marketing. It is possible that the earlier phases' deficiency can be corrected by marketing activity, but it can also happen that the previous good results can be destroyed by a wrong market activity. Therefore, the harmonization of agro-technological innovation processes is essential.
The technical advantage of the innovation can only be realized if adequate marketing skills can support and complement the new products and technologies.
The innovation processes have been described by the first linear models that product ideas are born, based on these new product is planned, produced, and sold. However is more effective when the process starts out from the market needs, the new products are planned, manufactured based on these, and during the selling process the satisfaction of the needs is controlled. Nowadays it is essential, that the marketing has to link all the processes, including the innovation too.
The task of the innovation marketing is not just selling the novelty, but to acquaint the expectations of the relevant stakeholders and with this to help for the management to increase the support of the strategy by satisfying the market demands on a reasonable and legitimate way, and on the other hand, to promote the acceptance (diffusion) of planned and implemented results of the innovations as well.
According to our previous analyzes the investigated companies have recognized the need for development. The expenses of innovation and R&D show us an upward trend, even though the global economic crisis has broken most of the company's career growth. The companies expend for marketing just over one percent of their sales value (Figure 1).
It is well known that the marketing of Hungarian SMEs working poor in most of the cases, and there is no exception either at the agricultural machinery sector. Moreover, the results so far show that this field is perhaps one of the most critical elements of the manufacturer's functioning.

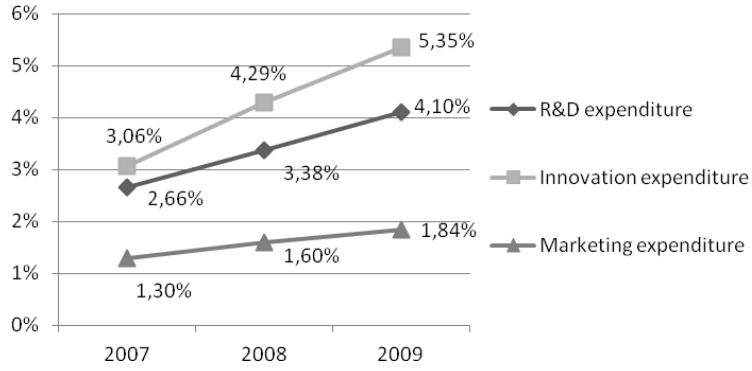


Figure 1. Expenditures on innovation, R&D and marketing (Source: own research results)

In spite of the low marketing expenditure the domestic agricultural machinery manufacturers declared that they endeavour to develop user-oriented. They say the one of the most important goal of their innovations is to satisfy customer needs, open new markets and increase the market share. The major partners of R&D cooperation are their customers, and among the innovation-related information are also the customer's needs the most important.

According to our study, **only 40 percent of the respondents conduct regular customer satisfaction study and market research**. Which, considering to the objectives and priorities is a poor performance. An innovative company has to develop not only technically, but also in respect of knowledge about the costumers. The SMEs have to also recognize that the relationship among corporate functions and their development play an increasing role in the shaping of competitive advantage. Namely the technological change and product innovation process is not only qualitative transformation or quantitative development of the technical elements, but it means the development of the interacting organizational features/elements too. The statements that "We are unsuccessful in the market because our marketing colleagues cannot sell the product" are out of date. Although the Hungarian agricultural machinery manufacturers have not to handle the organizational innovation in strictly part of the strategy, but they should think over a change of their attitude in certain areas of leadership and functioning.

Based on our examination the **domestic agricultural machinery manufacturers use basically the traditional communication channels**. In our survey we examined what are the most important success factors of the innovation, according to the companies' opinions. Unequivocally **they see the participation at professional exhibitions, as the most significant success factor of innovations**. It was an interesting matching that the exhibitions were also highly evaluated by the companies as an important information source for the innovation. The appreciation of shows can be justified by two reasons. On the one hand, **companies and their leaders can obtain direct information about competitors, their developments and current market trends**. On the other hand, the manufacturers expect the exhibition **as a primary marketing tool**. They can organize presentations of their new products in one place, where the potential buyers are concentrated.

The Internet, the online databases and the professional portals play a growing role among the marketing tools. In spite of this we found only few companies trying to collect data on the internet about the changes of customer's habits, characteristics. The practice shows that nowadays from web questionnaires, satisfaction surveys can be obtained cheaply a wide range of information about the trends of the customers demand and behaviour, etc. We think it would be beneficial to take advantage of this opportunity, but it would be possible only through the development of a comprehensive customer relationship management. During the interviews it was found out that the companies use only elementary customer databases (paper or Excel spreadsheet).

From the possibilities of the internet the Hungarian manufacturers use mostly only the webpage. **About 80 percent of the studied companies have own website**, which couldn't considered a bad ratio. The level of the web pages has wide range. Usually the websites are multilingual, and we can find product descriptions, images, price lists and other more information. We think there is still great possibilities in the development of this sites, mainly for those companies is highly recommended to develop them more, which would like to enter into foreign countries. A well designed website, which contains many actual information certainly make a good impression on the potential buyers.
It should also pay attention to the new and modern internet tools, which give good opportunities for targeting and new way of contacting with the costumers (e.g. various direct marketing channels: newsletters or Facebook etc.).
We made a Cross Tab Analysis to examine the relationship between innovation performance and marketing department. The independent variable was whether the company has a marketing department. It was possible to answer yes or no. We introduced a so-called innovative index for the dependent variable. The innovative index can have 3 values:

- Laggards: In the examined period there hasn't been introduced a new product or technology.
- Adopters: New products or manufacturing technology has been introduced in the examined period.
- Innovators: New products or manufacturing technology has been introduced in the last 3 years.

Our null hypothesis was that there is no correlation between the two variables. If we reject the null hypothesis, it means that there is a coherency between the two variables. To determine the statistical significance between the two variables we used the Pearson's Chi-square test. The strength of the connection was characterized by the Cramer's V index. Based our research experiences we defined our null hypothesis, which means that Hungarian agricultural machinery manufacturers subordinate their marketing activities to short-term market objectives and for this reason their marketing has a weak effect on their innovation and R&D activities. So the two variables mustn't have a strong correlation.
After finishing the calculations we received for the Chi-square value 0.254. The two-sided significance level is 0.223, which exceeds the chosen 5% significance level, so **we keep the null hypothesis of the Chi-square test, that there is no correlation between the two variables**. This means we could not detect a coherency between the innovation success and marketing activity.
The first table clearly shows the internal relations as well. We can see that 75 percent of laggards, 63.6 percent of adopters and 71.4 percent of innovators haven't a marketing department. If we compare the results to the entire sample (69.0 per cent), it is clear that all three dependent variable approaches to the average. This also means that this relation depends weakly together. Our hypothesis is confirmed, the marketing activity of domestic agricultural

machinery manufacturers has weak effect on the R & D. This also means that the marketing activity of the most innovative companies of the sector only plays an unremarkable role in the market success. These results are not so surprising according to another fact of our study, that the companies usually employ only a person (moreover a part-time employer) for marketing.

Table 1. Cross table of correlation between marketing department and innovation performance (Source: own research results)

			Innovation index			Total
			Laggards	Adopters	Innovators	
Has the company a marketing department?	No	Partition coefficient based on rows (%)	15.0	35.0	50.0	100
		Partition coefficient based on columns (%)	75.0	63.6	71.4	69.0
		Total (%)	10.3	24.1	34.5	69.0
	Yes	Partition coefficient based on rows (%)	11.1	44.4	44.4	100
		Partition coefficient based on columns (%)	25.0	36.4	28.6	31.0
		Total (%)	3.4	13.8	13.8	31.0
Total		Partition coefficient based on rows (%)	13.8	37.9	48.3	100
		Partition coefficient based on columns (%)	100	100	100	100

4. Conclusion

In this paper we examined the influencing factors of the relationships between the agricultural machinery innovations (R&D) and marketing activities.
The characteristics of the Hungarian agricultural machinery manufacturers drawn in the examination illustrate the situation of the sector in the country. A decisive part of the organisations (83%) are small-and medium-sized enterprises. All in all, only about 26-27% of the national need for machinery derives from national manufacturers. Their attitude in development is reflected by the fact that more than 70 percent of their products are sold on the domestic market. As a result of their intention to increase export, one-quarter of their products are launched on foreign markets. Regarding the factors that influence their sales results we concluded that almost one-third of the examined organisations had a separate R&D department and the proportion of organisations that have a separate marketing department is similar.

In a previous study (Bak-Husti, 2011) we found that the domestic agricultural machinery manufacturers are behind the West European competitors with the efficiency, so in the fight for improving the productivity indicators they prefer short-term actions, which are immediate confirmed by the markets.

Our Cross Tab Analysis showed that the domestic agricultural machinery manufacturers subordinate their marketing activities under short-term market objectives and their marketing has a weak effect on the innovation and R&D activities. So in the practice of the enterprises the long-term corporate (marketing) considerations take a back-seat because of the underplaying short-term revenue requirements.

From the results it can conjecture that the marketing plays role only in short-term revenue targets and it is linked to immediate decisions, and on this way it doesn't influence really the long-term market perceptions of R&D development. It is justified by the low proportion of the marketing expenditures. While the ratio of innovation and R & D expenditures are 4-5 percent of the revenue, the marketing spending is hardly more than one percent.

Based on the summary of the research we can see an interesting paradox. Although the manufacturers think, that it is very important the satisfaction of the costumer's demand, they don't allocate more resources to the marketing activities. So we cannot find any strong relation between the innovation and marketing, because of the uniformly low marketing budgets.

The other important final conclusion of the study is the **weak utilization of the wide range of marketing tools** by the enterprises. But this low marketing activity is not only because of the low marketing budget, but also it is based on the lack of marketing knowledge and an appropriate market approach. We propose for the manufacturers to expand the use of marketing tools, which is not only a question of money.

5. References

Árpád Bak, István Husti (2011): The Examination of the Innovation Activity of the National Agricultural Machinery Manufacturers, Mechanical Engineering Letters, Gödöllő.

Edquist, Ch. (1997): Systems of Innovation. Technologies, Institutions and Organisations. Pinter, London & Washington.

Halpen László-Muraközy Balázs (2010): Innovation and efficiency of enterprises in Hungary. Közgazdasági szemle, LVII évf.,4. pp. 293-317 (in Hungarian).

Husti István (2010): Elements of innovation management. Gödöllő (in Hungarian).

KSH(2010): R&D 2009. Statisztikai tükör, 2010 IV. évfolyam 89. szám (in Hungarian).

Losoncz Miklós (2008): The EU accession and the strategy of R&D and technology. Közgazdasági Szemle LV: évf, 2. pp. 169-182. (in Hungarian).

6. INSTITUTE FOR MECHANICS AND MACHINERY

PROFESSOR DR. ISTVÁN SZABÓ
DIRECTOR OF THE INSTITUTE

Dear Reader,

The Institute of Mechanics and Machinery has been known for integrating fundamental engineering disciplines in the fields of classical mechanics, design, computer aided technologies, and machine constructions since 2005, when the Institute was established by merging two formerly independent departments. Today it consists of three scientific and educational units:

- Department of Mechanics and Technical Drawing
- Department of Machinery
- Department of Agricultural and Food-industrial Machines

Courses provided by our faculty members cover a wide range from basic engineering areas (eg. computer aided drawing, statics, strength of materials etc.) to more specialized domains (biomechanics, computer aided analyses, machine design, bio-systems engineering, digital image processing, etc.) at all levels of studies (Post secondary, BSc, MSc, PhD). Also special post graduate trainings are being offered by the Institute including the one in the area of Township operation engineering, or Spa Engineering and Management.

The Institute of Mechanics and Machinery focuses on several advanced research areas. Running projects financed by the European Union, the Hungarian National Office for Research and Technology and also by industrial partners. Activities are concentrated around the following key topics:

- Biomechanics
- Applied engineering numerical methods
- Computer based modeling and analyses
- Advanced methods in engineering design
- Energy purpose biomass production

Selected papers published in the following section – with some limitations due to the physical boundaries of the Letters given - may give some impression of the scientific activity and interest of the researchers and PhD students working at the Institute for Mechanics and Machinery.

SPECIFIC SURFACE AREA OF CEREAL GRINDS AS A FINENESS PARAMETER

- Study on Hammer Mill and Grinding -

Lajos FOGARASI, László BENSE, István FÜLÖP

Department of Farm and Food Machinery, Institute for Mechanics and Machinery

Abstract

From the data base of a many-year hammer-mill test series on grinding different grains and other vegetable materials in different conditions, data points of the conjugate particle-size and calculated specific surface-area values were plotted in. The close fit of the curves indicates an apparently simple characteristic relationship (model law) between the two fineness parameters of grinds – the specific surface area vs. nominal particle. The gained regression function is a simple general hyperbola – the "fineness characteristic curve".
This research project are supported by the European Union and co-financed by the European Social Fund (TÁMOP-4.2.1.B-11/2/KMR-2011-0003).

Keywords

Comminution, grinding feed materials, particle fineness, particle size, particle size distribution,

Introduction

In many modern, high production plants where the equipment runs at least two shifts per day, the cost of energy during one year can easily exceed the cost of a new hammer mill. In other words, the energy to operate a hammer mill (or roller mill) during its normal expected life will be 10 to 20 times more expensive that the machine cost alone.
Similarly, the energy requirement and the cost of grinding exponentially increase with the fineness of the finished product (grits) (Figure 1).

Determining and expressing of grits fineness

Determining and expressing fineness of grind has been the subject of study as long as feed ingredients have been prepared. While appearances or feel may allow an operator to control a process, subjective evaluation is inaccurate at best and makes objective measurement and control virtually impossible. Descriptive

terms such as coarse, medium and fine are simply not adequate. Describing the process or equipment is also subject to wide differences in terms of finished particle size(s) produced.

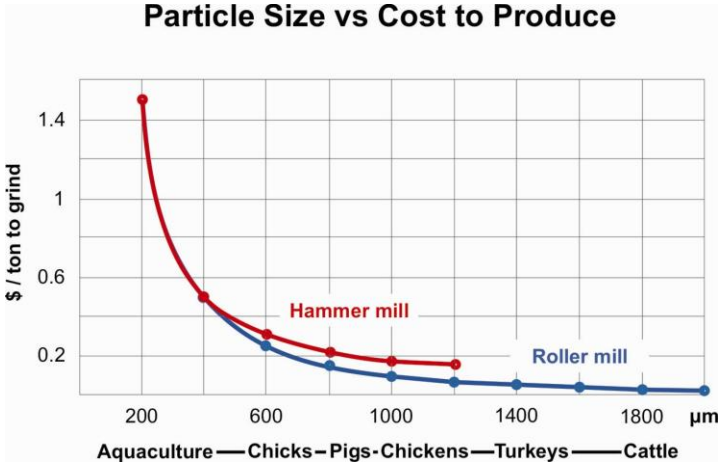
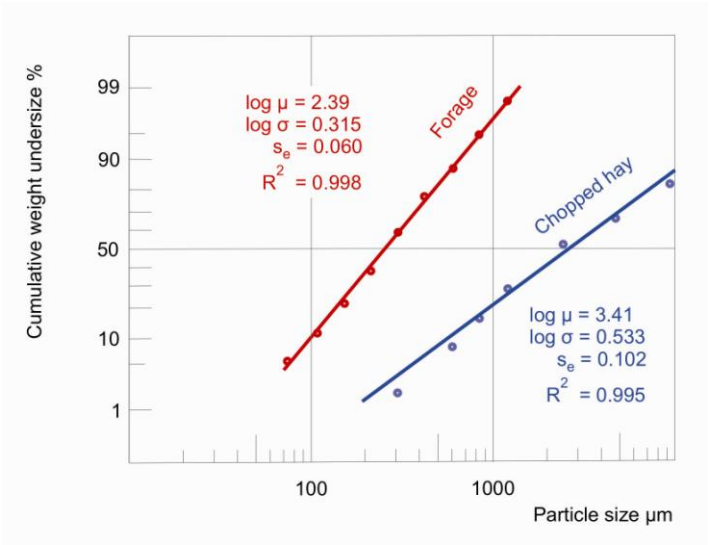


Figure 1. Cost of production as a function of grits fineness (mean particle size) (Heimann, 2001) with typical fineness ranges for different farm animals and aquaculture

Factors such as moisture content of the grain, condition of the hammers and/or screens (hammer mill) or the condition of the corrugations (roller mills) can produce widely varying results.
In addition, the quality of the grain or other materials being processed can have a dramatic impact on the fineness and quality of the finished ground products.
The best measurement of finished particle sizing will be some form of *sieve analysis*, expressed in terms of mean particle size or percentage (ranges) on or passing various test sieves. A complete sieve analysis will not only describe the average particle size but will also indicate peculiarities in the distribution, such as excessive levels of fine or coarse particles, etc. Typical descriptions that lend themselves to objective measurement and control might be "corn ground to 750 microns" or "75% ≤ 14 meshes".
The widespread methods of determining the fineness parameters i.e. particle size and (specific) surface area of grits give well reproducible results, however, those (apart from the rare exceptions) cannot generally be transformed into the other analytic processes. As to the measurement error, according to Heywood, the results of sieving analysis scatter in the range of ± 17 % as compared to the average value at a 95-% confidence level, in all particle-size ranges.
To express the fineness of ground feed materials, the *log-normal approach* has been approved as a standard for grains, concentrate size distribution, amongst others all over the world, by the American Society of Agricultural Engineers

as well (*1969*). An interesting and important extension of this representation was when the method had been approved for use and describing such extreme materials as forage particles (*chopped or pelleted grass hay*), cow rumen and faecal materials as well (Figure 2).

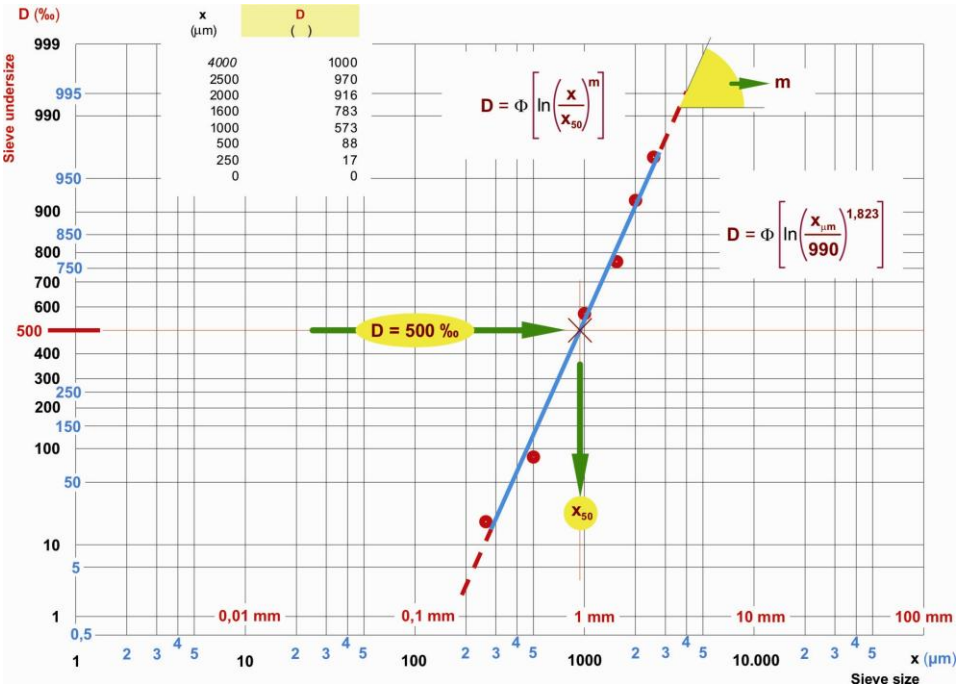


Regression parameters
log μ – $\log_{10}$ mean; log σ – $\log_{10}$ standard deviation; s_e – pooled residual standard deviation for sample set (Waldo, 1971)

Figure 2. Log-normal probability plots of forage sieve data on the left and a randomly selected sieving of chopped hay on the right

For different kinds of forage materials, a size class, as characterized by a given sieve size interval, does not necessarily mean the same thing with respect to actual size (volume, surface area, or weight) of individual particles. This problem can only be overcome by particle characterization methods other than simple sieving – such methods would have to separate by shape (and maybe density) as well as by size.

This, however, is a question independent of whether the log-normal distribution characterizes simple sieving results. It can be stated: *if simple sieving* of feed materials or forages is *acceptable*, then the work here reported *justifies* use of the *log-normal description* (Figure 3). The particle-size distribution curve "straightens" when sieve data plotted on the logarithmic Gaussian paper (diagram net) and the regression becomes simplified to fitting a line (linear regression).



(Korzenszky, 2009)

Figure 3. The general lognormal representation of particle size distribution of grinds plotted on logarithmic probability (Gaussian) paper (*an arbitrarily selected sample*) x_{50} – particle size to D = 50 % (here 500 ‰); m – slope of the regression line in the diagram net

The known expressing methods of fineness of ground feed materials use very different (so-called nominal or characteristic) particle sizes. These all can be determined from a well constructed distribution diagram and it is a very useful tool when different (especially old or conservative) data bases must be interpreted, analysed or compared (Figure 4). The characteristic particle sizes scribed in the linear-scale particle-size distribution diagram shows quite a crowded picture.

Without knowing the full distribution (or at least the model law and deviation parameter), neither the specific surface area of grinds can be determined nor the nominal particle sizes can be transformed into each other.

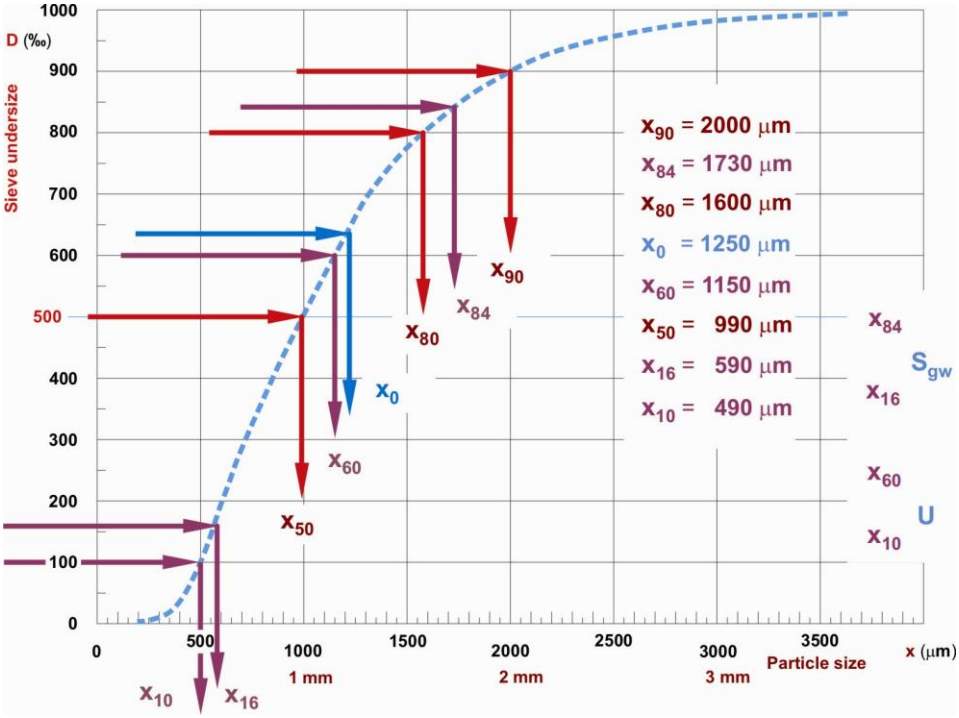
Relationship between specific surface area and particle-size distribution

Theoretically, a measured and expressed particle-size distribution can be converted into the parameter of specific surface area of the particles but, contrarily, there is no way to reconvert the already known specific surface area value of a ground material into its particle sizes (especially its particle-size distribution). According to Kihlstedt's hypothesis, the Bond's nominal particle

size ***Z*** – an earlier used industrial fineness index, d_{80} or x_{80}; the size of the sieve which the 80 % of particles falls through – is in a direct relationship with the specific surface area (Figure 5):

$$A_{vol\,spec}\sqrt{d_{80}} \cong const$$

where: $A_{vol\ spec}$ is volumetric specific surface area (cm^2/cm^3), d_{80} is Bond's index Z – the nominal particle size at 80 % cumulative mass of undersize



(Fogarasi 1996)

Figure 4. Particle size distribution of grinds – nominal particle sizes by the old and existing standards (*the ordinary average or mean particle size* x_{mean} *by an old Hungarian standard is not shown in the diagram*)
E.g. x_{90} – particle size to D = 90 % (here 900 ‰) etc.; x_{90}, x_{80} – parameters from earlier industrial or scientific tests; x_0 – nominal particle size of RR(S)B model (DIN standard); x_{50} – median particle size by mass (e.g. Hungarian, ISO, ASAE standards); x_{60}, x_{10} – auxiliary parameters for calculating the non-uniformity factor U by Hungarian standard; x_{84}, x_{16} – auxiliary parameters for calculating the geometric standard deviation of particle size S_{gw} by ASAE; m – slope of the regression line in the diagram net

Kihlstedt's relationship is a general hyperbolic function, however in his calculations, the constant value of about 750 altered according to the sieve sizes but it proved statistically correct in the domain of usual sieve analysis between 75 and 50,000 μm. Anyway, the hypothesis is based on the traditional Gaudin-Andreyev-Schumann logarithmic distribution model which is a quite good

approach (linear when plotted on a log-log paper) up to D = 80 to 90% undersize values.

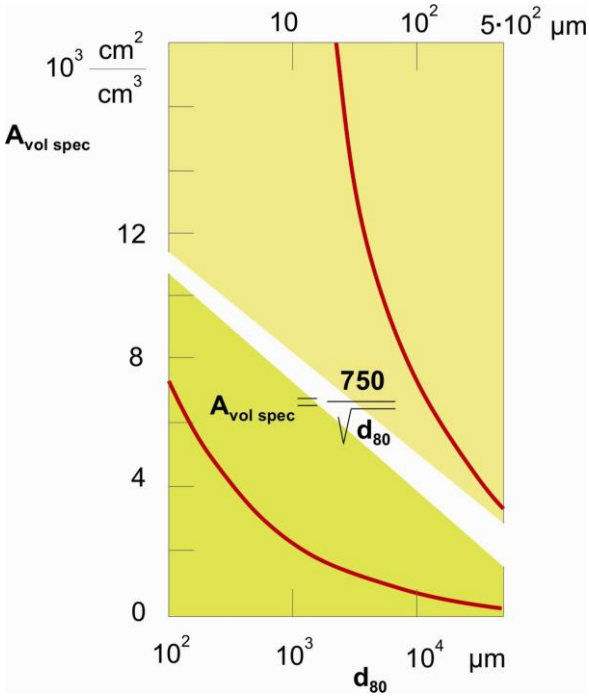
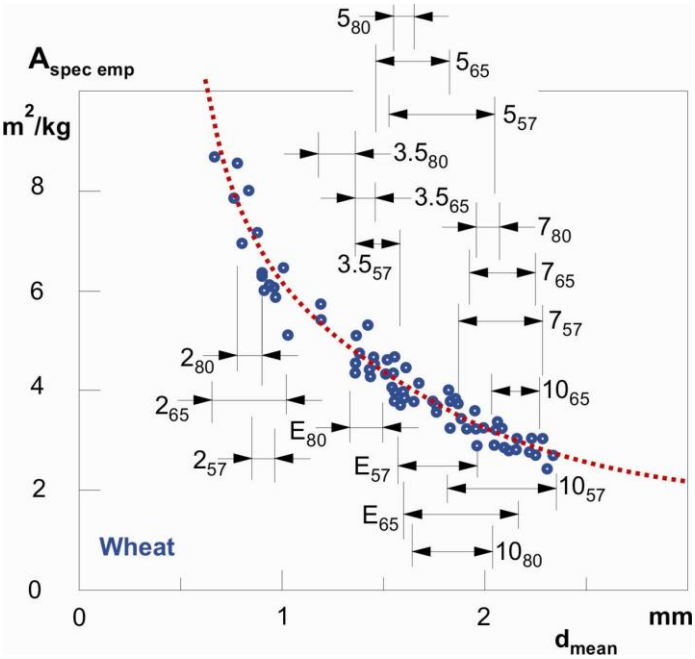


Figure 5. Kihlstedt's relationship: the volumetric specific surface area of the ground material as a function of the nominal particle size d_{80}
An empirical curve fitted on a multi-element measured-computed point set of quite different industrial (mineral) materials, scattering in a wide range

The specific surface area of the ground material A_{spec} plays a significant role in the different comminution-energetic theories and investigations; the input grinding (crushing) work is in a certain proportion to the newly produced particle-surface area. In fact, the value of the specific surface area of a granular (powder) bulk is a fineness parameter – and, in addition, it is only a single figure of characterizing the particle fineness, however, not a simply "sensible or tactile" property in comparison with the particle size.
The (specific) particle-surface area data directly measured by one of the introduced methods are not reliable at all; the different techniques yield very different values even of the same sample. The difference could be of two times in order of magnitude so it is expedient to use the calculated surface area. However, due to the simplifications, the computed parameter is not the real surface area of particles but a number as a good surface area character.
From the data base of a many-year hammer-mill test series on grinding different grains and other vegetable materials in different conditions, using Kihlstedt's

investigation as a model, the data points of the conjugate particle-size and calculated specific surface-area values were plotted in diagrams and regression functions (curves) were fitted on the "measured" points (Figure 6) (Fogarasi, 1990). The close fit of the curves is clearly shown in the figures, indicating an apparently simple characteristic relationship (model law) between two fineness parameters of the samples of grinds – the specific surface area vs. nominal particle. Here the selected parameters are the ordinary average (mean) particle size d_{mean} (or x_{mean}) and the ordinary (so-called "empirical") specific surface area ($A_{spec\ emp}$, assuming that the particles of grinds are cubic); both can directly be calculated from the data sheet of sieving analysis, however, with full knowledge of the size distribution, the calculation is refined and more accurate. The diagrams plotted with other nominal particle sizes e.g. x_0 or x_{50} (d_{50}) and specific surface areas A_{spec} computed by arbitrarily chosen techniques (e.g. according to the RR(S)B model or ASAE standard) result in very similar function type.



Fogarasi 1990

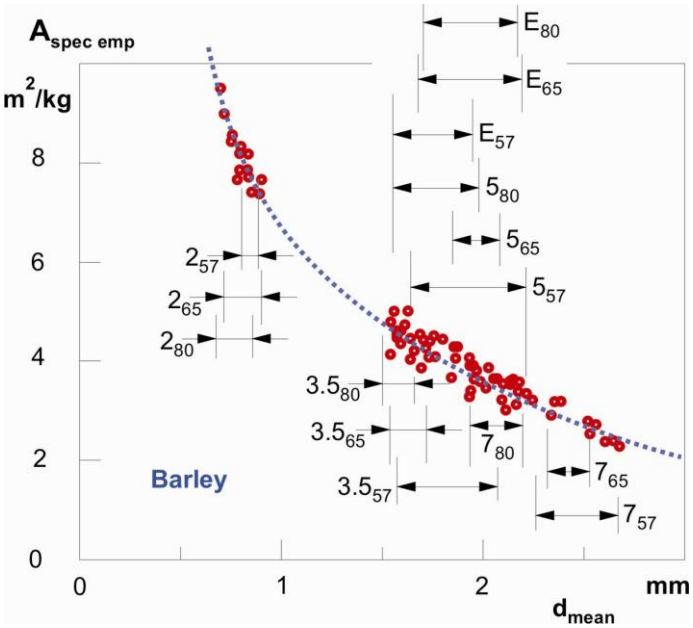
Figure 6. Storage-dry wheat grits – "Empirical" specific surface area $A_{spec\ emp}$ vs. average particle size d_{mean} by mass
Conventional hammer-mill screens: 2, 3.5, 5, 7 and 10 – diameters of screen aperture (punched round holes) in mm; Screen of expanded steel plate E – similar to the conventional 5-mm screen as to its size; Lower indices – 57, 65 and 80 m/s measured hammer-tip (peripheral) speed

In this case, the regression curve is the simplest general hyperbola; for *wheat* samples its equation is

$$\overline{A}_{specemp} = \frac{7.80}{d_{mean}} \text{ (m}^2\text{/kg)}$$

where: d_{mean} is the ordinary average particle size (otherwise x_{mean} or x_{ave}), mm

Very similar regression functions were gained in the case of barley and maize kernels ground by the test hammer mill with 12 hammers (Figures 7 and 8).



Fogarasi 1996

Figure 7. Storage-dry barley grits – "Empirical" specific surface area $A_{spec\ emp}$ vs. average particle size d_{mean} by mass
Definition of lettering: See Figure 6

It is worth mentioning that, according to independent measurements grinding barley by a similar hammer mill with 24 hammers, this function is

$$\overline{A}_{specemp} = \frac{8.15}{d_{mean}} \text{ (m}^2\text{/kg)}$$

where: d_{mean} is the empirical average particle size, mm

If the above equations are transformed to the solid-matter density of 1000 kg/m^3 (or actually volume base), the following formulae are gained:

$$\bar{A}'_{specemp} = \frac{10.64}{d_{mean}} \text{ (m}^2\text{/kg) for wheat (12-hammer mill)}$$

$$\bar{A}'_{specemp} = \frac{10.61}{d_{mean}} \text{ (m}^2\text{/kg) for barley (24-hammer mill)}$$

where the body-density of kernels ρ_{solid} = 1000 kg/m^3 – an imaginary density value
The shape of these formulae requires an extraordinary attention.

This essential relationship can be considered as the fineness characteristic curve of an actual material to be ground in a hammer mill which is invariant independently of the operating parameters such as mill screen size, peripheral speed, and mass flow.
The above correlations can be proved in an analytical way as well. Apart from the detailed discussion, the ASAE standard process is taken here for demonstration as it follows.
The equation for estimating the total surface area of particles in a sample charge is:

$$A_{st} = \frac{\beta_s W_t}{\beta_v \rho} \exp\left(4.5\sigma_{\ln}^2 - \ln \mu_{gw}\right)$$

where:
A_{st} is estimated total surface area of a charge, cm^2
β_s is shape factor for calculating surface area of particles. Cubical, $\beta_s = 6$; Spherical, $\beta_s = \pi$
β_v is shape factor for calculating volume of particles. Cubical, $\beta_v = 1$; Spherical, $\beta_v = \pi/6$
ρ is particle density of the material, g/cm^3
σ_{ln} is log-normal geometric standard deviation of parent population by mass in natural logarithm, use S_{ln} as an estimate
μ_{gw} is geometric mean particle diameter of parent population by mass, cm, use d_{gw} as an estimate
W_t is mass of a charge, g

If the constant properties are included in a single parameter C_{ASAE}, the above equation, substituting the measured-calculated variables S_{ln} and d_{gw} (or d_{50}) for σ_{ln} and μ_{gw}, respectively, can be written down with the following formula:

$$A_{st} = C_{ASAE} \frac{e^{S_{\ln}^2}}{d_{gw}} \text{ (cm}^2\text{)}$$

During an earlier research (*Fogarasi et al. 1990*) it was derived that the deviation-property of the particle-size distribution of hammer mill products (in the forms of standard deviation, n, U, S_{gw} or S_{log} as well as m) behaves as a probability variable itself, and has an expectable value (or average value) in fixed conditions – in the case of the same material and the same mill. With this, the expectable value of S_{ln} can be substituted in the numerator of the above fraction:

$$A_{st} = C_{ASAE} \frac{e^{\bar{S}_{\ln}^2}}{d_{gw}} \text{ (cm}^2\text{)}$$

Accordingly, after determining the expectable value of S_{log} or S_{gw} (more exactly its mean value calculated from a quite large series of ground samples) the equation can be written down as –

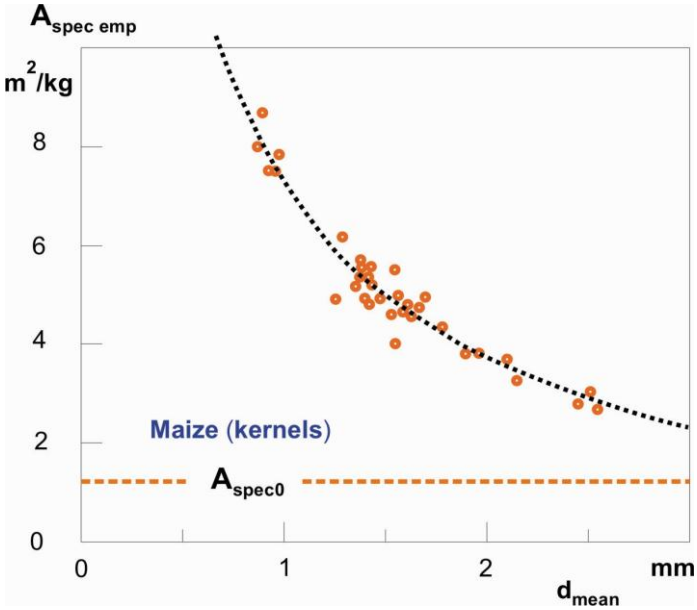
$$A_{st} = \frac{C'\left(\bar{S}_{gw}\right)}{d_{gw}} \text{ (cm}^2\text{)}$$

And this is a similitude of the above simple hyperbola – a material grinding and machine characteristic function or otherwise "fineness characteristic curve" which can be determined by tests for each hammer mills (probably for other mill types as well) and materials to be ground. Knowing the actual characteristic curve, the grinding calculations and the analysis of the mill operation can significantly be simplified. In a certain sense, the constant of the characteristic function can be considered as a special grinding material property.

Specific surface area of whole grains

When grinding, one of the most important character of the grinding performance is the produced new surface area (so-called surface-area increase) which is the difference between the surface area of the fine grits particles and that of the input material (Figure 8). In the actual cases, the input material frequently is fed in the form of whole grains. The behaviour of the grains is different from the comminution of the already crushed particles and the acceptable determination of its specific surface area is difficult as well. However, using the fineness characteristic curves with a certain extrapolation, the specific surface area of a

kernel can be defined and estimated as that of a big particle just started to be ground.



Fogarasi 1996

Figure 8. Storage-dry maize grits – "Empirical" specific surface area $A_{spec\ emp}$ vs. average particle size d_{mean} by mass; A_{spec0} – specific surface area of the kernels fed in

Conclusions

The best measurement of finished particle sizing is some form of *sieve analysis* but a complete sieve analysis will not only describe the average particle size but will also indicate peculiarities in the distribution, such as excessive levels of fine or coarse particles, etc. It can be stated: if simple sieving of feed materials or forages is acceptable, then the work here reported justifies use of the log-normal description.

The specific surface area of the ground material A_{spec} plays a significant role in the different comminution-energetic theories and investigations; the value of the specific surface area of a granular (powder) bulk is a fineness parameter – and, in addition, it is only a single figure of characterizing the particle fineness, however, not a simply "sensible or tactile" property in comparison with the particle size.

An apparently simple characteristic relationship (model law) exists between the two fineness parameters of grinds – specific surface area vs. any arbitrarily

chosen or defined nominal particle size; however, with full knowledge of the size distribution, the calculation is refined and more accurate. The regression curve is a general hyperbola.
The deviation-property of the particle-size distribution of hammer mill products (in the forms of standard deviation, n, U, S_{gw} or S_{log} as well as m) behaves as a probability variable itself, and has an expectable value (or average value) in fixed conditions – with the same material and the same mill.
Accordingly, using the standard ASAE, after determining the expectable value of S_{log} or S_{gw} (more exactly its mean value calculated from a quite large series of ground samples) the equation can be written down as –

$$A_{st} = \frac{C'\left(\bar{S}_{gw}\right)}{d_{gw}} = \frac{Const}{d_{gw}} \text{ (cm}^2\text{)}$$

And this is a similitude of other fineness characteristics with different standard particle-size distribution parameters.

References

Fogarasi L., Mikecz I., Akdeniz, C. (1990): Influence of Screen Construction on Fineness of Grits made by Hammer Mill. *Hungarian Agricultural Engineering 3/1990*

Heimann, M. (2001): Economics of Grinding for Pelleted Feeds. *Roskamp Champion.*

Waldo, D. R. et al. (1971): Logarithmic Normal Distribution for Description of Sieved Forage Materials. *Journal of Dairy Science Vol. 54, No. 10*

Korzenszky P., Fogarasi L. (2009): Comminution of cereal feed components – new technological facilities. *International Conferences on Agricultural Engineering, Synergy and Technical development in Agricultural Engineering, Gödöllő.*

CHAFF SEPARATION IN COMBINE HARVESTERS

István BOGNÁR[1], Béla M. CSIZMADIA[2], Péter SZENDRŐ[1]

[1]Department of Machinery, Institute for Mechanics and Machinery
[2]Department of Mechanics and Technical Drawing, Institute for Mechanics and Machinery

Abstract

Due to the negative effects of climate change the time interval for grain harvesting under optimal conditions is becoming increasingly narrower. The dynamic growth of the Earth's population as well as the use of species of increasingly greater yield necessitate the development of machinery capable of harvesting greater quantities over a shorter period of time.
As a result of these developments, combine harvesters operating by new principles, first those with the axial and later those using the hybrid system appeared, characterised by a significantly increased throughput compared to their conventional counterparts.
While operating hybrid combine harvesters, i.e. those equipped with an axial flow residual grain separation rotor unit mounted along the longitudinal axle of the machine, users experienced inability to fully utilise the full potential of the machines under dry harvesting conditions. The reason of reduced capacity turned out to be the excessive chaff separation, overloading the machine's cleaning unit, thereby limiting its throughput capacity.
The precise description of the grain and chaff separation processes in rotary residual grain separator systems as well as the clarification of the correlations between their influencing factors therefore emerged as important challenges in order to evaluate the operation of the equipment to assess its efficiency and to forecast the effects of possible design amendments on the work process.
The present article reports on the examinations conducted in the field of chaff separation processes in combine harvesters and the factors influencing chaff separation.

Keywords

Combine harvester, chaff separation, residual grain separation, hybrid system

Introduction

The world's dominant agricultural machinery plant is always improving the grain harvest for their planes, in order to increase throughput, improve work quality, and reduce the power requirements (Bognár, Szendrő, 2007). The difference

between the hybrid system and conventional combine harvesters in terms of construction and operation is the rotating axial flow residual grain separation unit mounted along the longitudinal axis of the machine instead of the straw walkers. This is made up of two parts: the rotating rotor and the stationary concave. Due to the centrifugal force in the equipment, residual grain separation is significantly more efficient than in conventional machines. The forces the straw and the grains are subjected to are about 7-10 fold greater than in conventional systems, allowing increased throughput compared to straw walker separators at identical levels of grain loss (Bognár, Szendrő, 2008).
Resulting from the operation of the machine, the straw gradually disintegrates into small particles due to the intensified mechanical interaction between the straw and the pieces of the rotor (Bognár, Szendrő, 2006). This is the very chaff generation, which under dry continental harvesting conditions may be of such significant extent that it would drastically reduce the utilisation rate of the combine harvester, thereby its overall capacity by overloading the cleaning unit of the machine (Bognár, 2010). Chaff separation therefore influences the machine's operation negatively (Hunyor, Szendrő, 2006); the present goal is thus the implementation of a structural optimisation using the results of the examinations which would achieve reduced chaff separation without deteriorating the present level of grain separation. A further negative phenomenon is that the process may render the harvested straw completely unsuitable for any further processing at low levels of straw humidity.
These observations substantiate the necessity to examine and analyse the chaff separation process. The task of primary importance is the precise description of the grain and chaff separation processes in rotary residual grain separator systems as well as the clarification of the correlations between their influencing factors (Kutzbach, 2000). Once in possession of the results, the operation and the efficiency of the systems can be assessed and the effects of possible design amendments on the work process can be forecast.

Materials and methods

Winter wheat harvested from a previously indicated and carefully selected homogeneous parcel under dry conditions, tied into sheaves and packed into containers was used for the experiments. The characteristic straw to grain ratio of the grain stocked for laboratory measurements was around 1:1 according to preliminary measurements with the grain's moisture content being at 15-16%, and the straw's moisture content at 18-20% respectively.

The measurement method used was the complex process developed previously at the central research and development laboratory. A converted hybrid combine harvester was used for the experiments (fitted with tangential flow threshing equipment with accelerator drum mounted across and axial flow residual grain separator rotors mounted lengthwise). The key principle of the experiment was

the collection of the quantity of separated grain and chaff under the rotor separator unit with the help of measuring dishes placed along the longitudinal axis of the equipment. Chaff can be removed from the grain-chaff mixture using appropriate cleansing procedures. The quantities of grain and chaff obtained in this manner were recorded.

During the tests, the throughput capacity of the machine was varied as well as the tangential velocities of the residual grain separator's rotors with a number of experimental separator devices. Further to this, experiments were conducted using concaves of lengths 2600, 2080, and 1560 mm as well. Experiments have also been run using straws of various moisture contents.

Results

Significant differences can be observed between the histograms displaying the processes of grain and chaff separation in the perspective that the histogram representing grain separation decreases considerably along the longitudinal axis of the equipment (Fig.1.) while the histogram representing chaff separation seems to set at a relatively high and constant value especially at the last one third of the concave. Significant chaff separation can therefore be observed at minimal levels of grain separation. The longer the equipment, the more chaff falling through it.

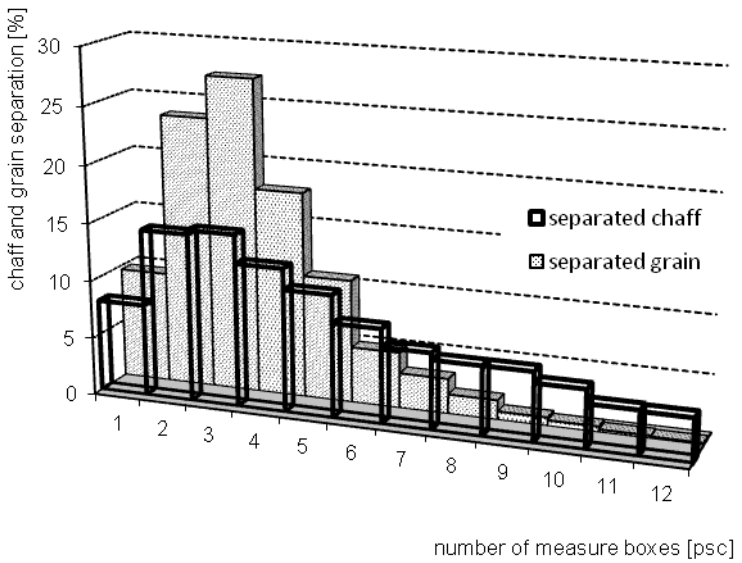


Figure1. Chaff and grain separation

It can be seen that the various structural solutions at similar settings of parameters result in a similar nature of chaff separation but with varying extents along the longitudinal axis of the equipment. The different chaff separation values are the results of the various intensities of mechanical interaction in the various structural solutions. Solutions with more intense mechanical interaction result in greater levels of chaff generation. According to the Table Nr.1., the most favourable solution in terms of chaff separation was equipment no. 2 with an average value of 94%, the separator unit producing the highest level of chaff generation was equipment no. 1; its chaff separation in comparison to the best being 110%.

Table 1.

Effect of the rotor structural design to the chaff separation		
QNKB=25t/h, vR=22,3m/s		
	chaff [kg]	difference [%]
Equipment 1	14,08	103,68
Equipment 2	12,75	93,88
Equipment 3	13,91	102,44
average:	13,58	100

By increasing the throughput capacity, the quantity of material inside the system is increased. The mechanical interaction with this increased quantity results in a specific increase of chaff separation.

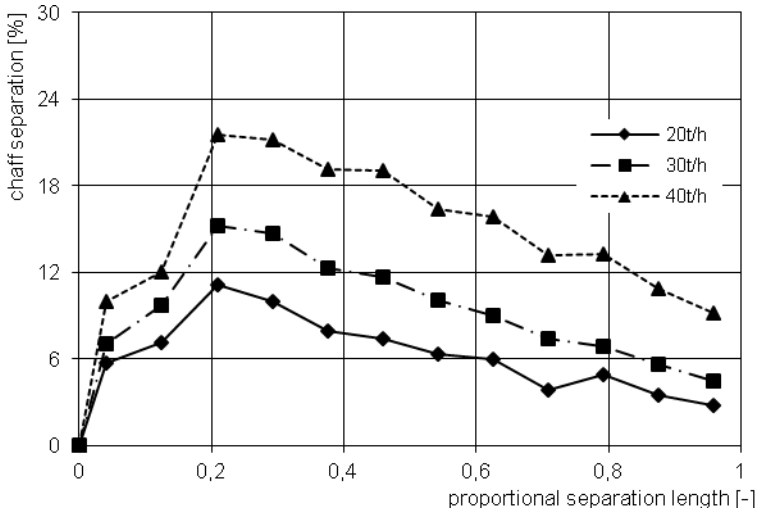


Figure 2. Chaff separation over the throughput

The graph (Fig.2.) clearly shows the process at various levels of load and allows to establish that the process is of a similar nature in each case but with varying intensity. According to the Table Nr.2., if the value of chaff separation corresponding to a throughput of 20t/h is taken as a base value, i.e. 100%, then at loads of 30t/h 150% of the base value could be measured and at maximal throughput – the double of the base load – the quantity of separated chaff was 240% of the base quantity! This phenomenon is explained by the larger mass of material in the system, the thicker layer of straw can only move with a proportionally greater level of resistance. Internal and external friction values increase, the tangential velocity component of the material decreases slightly, the relative velocity difference between the material and the rotor parts, resulting in more intense mechanical interaction lasting for a longer period of time.

Table 2.

Effect of the throughput to the chaff separation		
throughput	chaff [kg]	difference [%]
20t/h	10,806	100
30t/h	16,072	148,73
40t/h	25,594	236,85

It can be established that the effect of the tangential velocity of the rotor of the separator unit is less intense than what can be observed when varying the throughput capacity. The difference in chaff generation at minimal and maximal rotor tangential velocities is only 7% with the greater value corresponding to the greater tangential velocity. It can therefore be established that larger quantities of chaff are separated at larger tangential velocities; still, due to the factors of grain loss and grain separation as well as energy requirement, operation at greater tangential velocities is more advantageous.

Variation of the effective transmission surfaces of the concaves is done with the explicit aim of – as theory suggests – reducing the extent of chaff separation. The measurements conducted in the course of the present study were aimed at proving this hypothesis.
At identical settings, the total quantity of chaff separated along the full separation length of the concave was taken as the base quantity, i.e. 100%. The chaff separation measured after the closure of one concave segment – this representing 20% of the transmission surface – turned out to be 93% of the base value, while the simultaneous closure of the first two segments, i.e. 40% of the total transmission surface reduced the chaff separation to 75% of the base value! This means a load reduction of 25% on the cleaning unit of the combine harvester meaning an approximate 25% room for augmenting the machine's utilisation rate. It must not be forgotten however, that in this case a 40%

reduction of the separation length may result in an increase in the value of grain loss.

The properties of the vegetable matters have a significant influence on the grain separation process and chaff separation, too. Experiments were conducted with straws of various moisture contents to establish the straw's behaviour during mechanical interactions. The experiments were run in two moisture content intervals: one part of the measurements at a moisture content of 16.8% and the other part at 12.7%. The results clearly show the correlation between chaff separation and straw moisture content. (Fig.3.) According to the Table Nr.3., if the chaff separation value corresponding to 17% straw moisture content is taken as base value, i.e. 100%, then chaff generation at a straw moisture content level of 13% is 155%!

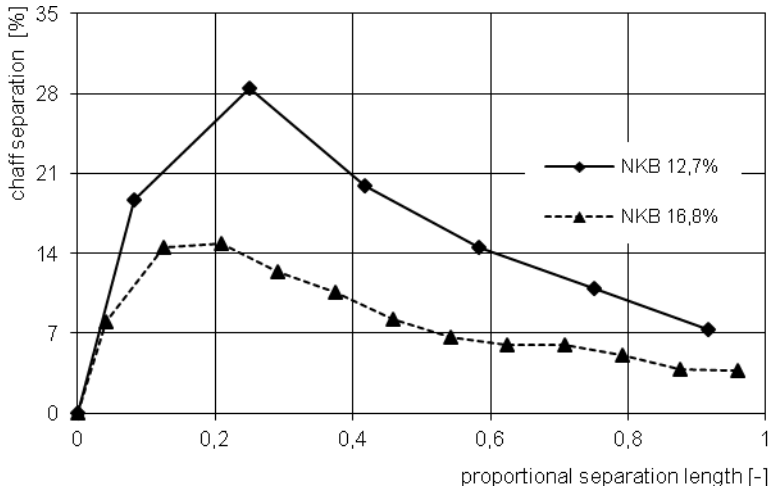


Figure 3. Chaff separation over the straw moisture content

Table 3.

Effect of the straw moisture content tothe chaff separation		
QNKB=25t/h, vR=22,3m/s		
moisture content	chaff [kg]	difference [%]
NKB 16,8%	12,752	100
NKB 12,7%	19,681	154,34

Conclusions

The results seem to suggest that due to the different nature of chaff and grain separation a solution to the problems described above could be the application of concaves with transmission slits differentiated in terms of size along the longitudinal axis of the equipment. This means that at places where grain separation is significant, concaves with normal sized openings should be used and where grain separation is low but chaff separation is still significant, concaves with smaller transmission openings should be applied. This would still ensure passage for the grains, but would restrict chaff from passing through in significant quantities.

Relying on the results of the experiments, an alternative way of reducing chaff separation could be a differentiation by length of the axial velocity component of the throughput while keeping the operational RPM constant. This would have a particularly significant effect under dry harvesting conditions, as the larger axial velocity would put the material through the system faster, resulting in less chaff being separated. The point of intervention into the chaff separation process would be where grain separation is no longer significant, but that of chaff still is.

My own measurements have shown that straw loses its flexibility at about 15-16% moisture content and becomes fragile. Based on this, dry harvesting conditions could be defined as below 16% of straw moisture content. It can be established that harvesting conditions on their own may contribute to an increase of up to 50% of chaff separation, overloading certain parts of the equipment, significantly reducing total capacity of the machine.

Acknowledgements

I would like to acknowledge the support of the TÁMOP – 4.2.1/B-11/2/KMR-2011-0003 project.

References

Bognár I, Szendrő P. (2006): Arató-cséplő gépek rotoros szem-szalma leválasztó berendezésének vizsgálata, Mezőgazdasági Technika XLVII évfolyam, 5. szám, 2-4. o.

Bognár I, Szendrő P. (2007): Az arató-cséplő gépek szalmarázó berendezésében lejátszódó szemleválasztási folyamat elméleti leírásának újszerű megközelítése. GÉP LVIII: 3-6. o.

Bognár I, Szendrő P. (2008): Arató-cséplő gépek szemleválasztása. Mezőgazdasági Technika XLIX: 2-4. o.

Bognár I. (2010): The efficiency increasment of combine harvesters based on new principles. Danubia-Adria 27th Symososium on Advances in Experimental Mechanics, Wroclaw, Poland, 2010.09.22-2010.09.25. SIMP Polish Society of Mechanical Engineers and Technicans, pp. 17-18.(ISBN:978-83-87982-59-1)

Hunyor Cs, Szendrő P. (2006) Az arató-cséplő gép anyagtranszport folyamatának dinamikai modellje az anyagáram mérése alapján, MTA AMB XXX. Kutatási és Fejlesztési Tanácskozás, Szent István Egyetem Gépészmérnöki Kar – FVM Mezőgazdasági Gépesítési Intézet, Gödöllő. 295-298. o.

Kutzbach H.D. (2000) Ansatze zur Simulation der Dresch- und Trennprozesse im Mahdrescher, VDI Tagung Landtechnik, Münster, p.17-22.

PILED UP WOOD CHIPS INVESTIGATION WITH 3D IMAGING

Attila LÁGYMÁNYOSI, István SZABÓ

Department of Machinery, Institute for Mechanics and Machinery

Abstract

Visual appearance is a key factor for many industrial and agricultural products in terms of quality. The most important attributes are colours and shape. Digital image processing is a popular and important application to investigate visual appearance. Conventional 2D-imaging provides a good implementation of colour identification, but is not a perfect solution for surface monitoring [Koschan, 2008]. 3D laser imaging as a new technology provides some additional data on the investigated objects, beyond conventional imaging. Three dimensional representation looses colour information of the object, but provides a point to point surface mapping. This article presents results of a research which applied 3D image analysis in order to demonstrate and utilise advantages in terms of additional data over conventional images. It is concluded that the additional information gained can be used to describe object surface features in a more thorough manner.

Keywords

Image processing, surface analysis

Introduction

In the field of agriculture post and pre harvest and in many other cases digital image processing is a popular and important application [Molto, 1996]. For example the visual appearance is a key factor in quality assessment and sorting plants. Therefore applied procedure, are based on visual inspection. In this paper in scragginess analyses of applicable 3 dimensional imaging procedures are described.

Surface unevenness is not only a potential characteristic data in the case of homogenous materials, but also in bulk material sets [Gonzalez, 2008]. Elements on the surface of the material sets causes irregularities on the whole set according to its size. Assuming complete mixing we can characterise the size of elements in the whole set based on the knowledge of the surface element sizes. In the case of multiple element material sets fundamentally not the respective sizes of the elements themselves but the distribution of the elemental multitudes (size classes) will be a characterising data.

Thus, if we can give a good description of the elements of the set with a statistical methodology then we can assign this characteristics to the set and compare sets based on this information. A typical such material set can be firewood chopping utilised as a renewable energy source for heat production. The chopping size introduced in the combustion area has a definite influence on combustion quality [Bense, 2006]. It is therefore straightforward that knowledge of firewood chopping provides additional information for configuration of combustion parameters.

THE APPLIED SYSTEM AND MATERIALS

The applied apparatus

Throughout the experiments a 3D laser scanner of the type Zscanner 700 was used. The main technical parameters were: sampling rate 18000 sample / sec., 2 built in cameras, improved resolution of 0,1 mm, maximal accuracy of XY positioning is 50 μm if the investigated volume is 100 mm x 100 mm. The applied computer was a PC with the following features: I7 quad processor, 6Gb memory, graphical subsystem with 1Gb memory. The connection between the scanner and computer was an IEE1394 interface.

The investigated materials

The examined material is wood chopping. The filtering method was as follows: 3 classes were applied, one under 4mm, one between 4 and 8 mm, and one over 8 mm. The fractional selection resulted in three size classes, but the size filtering only applied to two out of the three physical dimensions; it was of course possible that in the 4 mm fraction a chopping of 4mm diameter but 40mm length could penetrate. This resulted in a given probability distribution of size inside the given size groups, and the somewhat blurred distinction of the fractions.
Nevertheless it is understood that the sets formulated in the above manner carry distinctive size characteristics, and can only be assessed statistically.

The developed investigation method

During the experiment lighting was applied with no extra requirements in mind. In the preparation of the scanned recordings exclusively a filtered artificial lighting was used to ensure minimal bias over the scanner's own lighting apparatus [Szalay, 2011]. Fractions of chopping was placed in boxes with open tops. Size of the boxes was selected to ensure the largest possible cover for chopping with multiple layers of cover. Multiple layers was necessary to provide an environment which is possibly the closest to the surface exhibited by real material sets. Surface was always equalised to ensure that its unevenness only relates to size irregularities of respective elements.

Figure. 1. One case of the investigated material

The investigations

During the analysis all size classes were 3D-scanned after applying 3 different mixings to each of them.

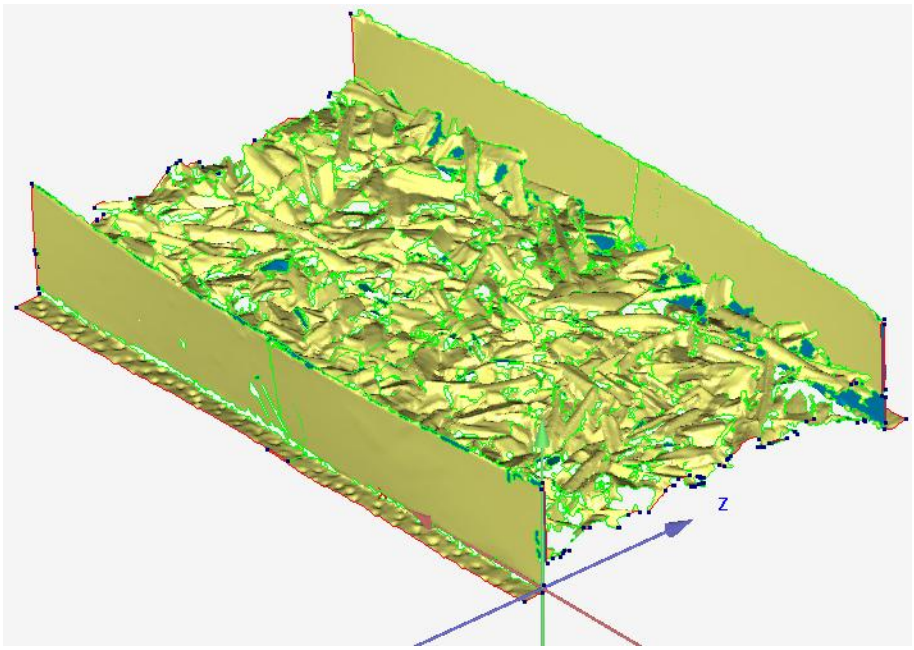


Figure 2. One example of the scanned material

Scanned images represent the allocation of firewood choppings distribution on the examined material surface. The respective points of the image can be stored in a 3D matrix. Examining a 2D slice matrix of the 3D matrix allows us to get a cross-section of the scanned material set. If the cross section is selected perpendicularly to the surface of the set, we get a cross cut image of the original set.

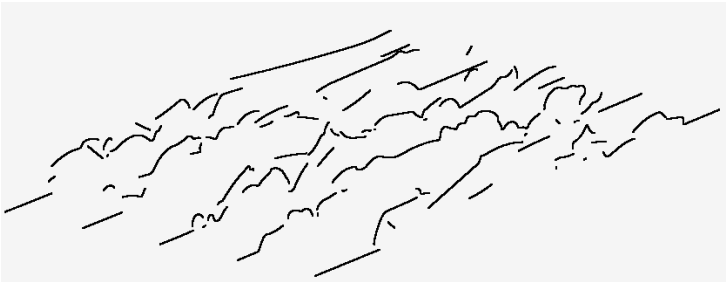


Figure 3. The selected sections form 3D matrix

For the purpose of our analysis always 5-5 slices were cut from the surface descriptor matrices. Considering the triple repetition of the three size classes this altogether resulted in 45 cross cut images and data sets.
Putting the data of the cross cut samples in a two dimensional vectorspace it can be stated that the cross cuts from the same fractions are very similar, while interfraction crosscuts show significant difference.

Results and measured data

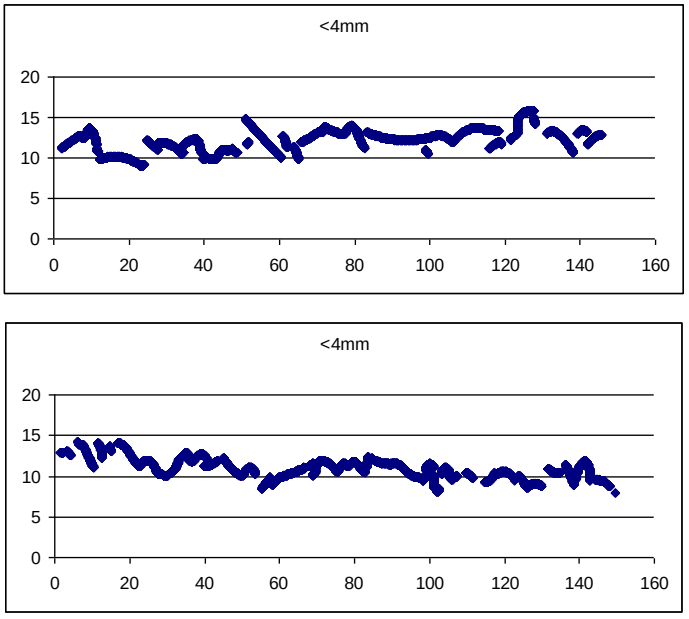


Figure 4. Comparison of the sections from same fractions (<4 mm)
x and y axes scale are [mm]

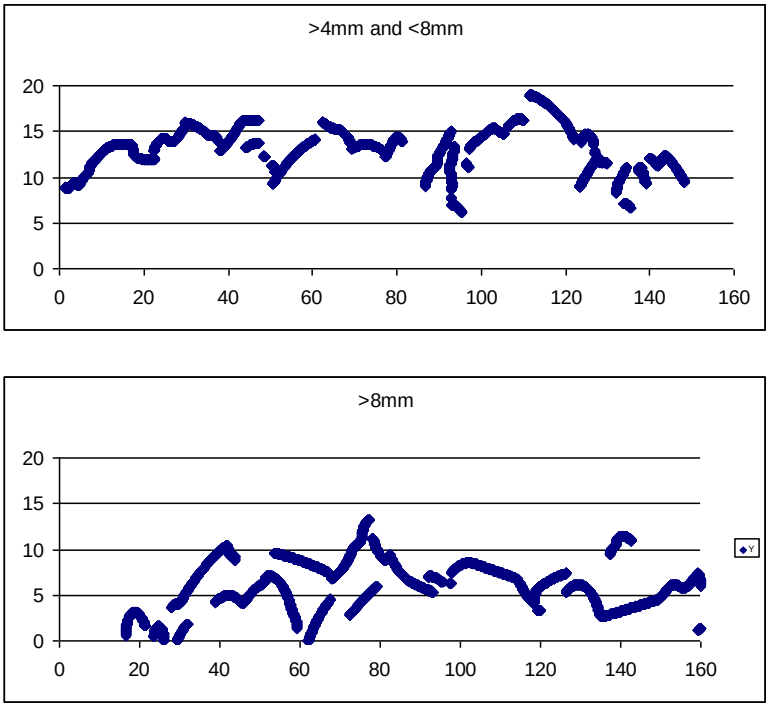


Figure 5. Comparison of the sections from different fractions (<4 mm and 4-8 mm) x and y axes scale are [mm]

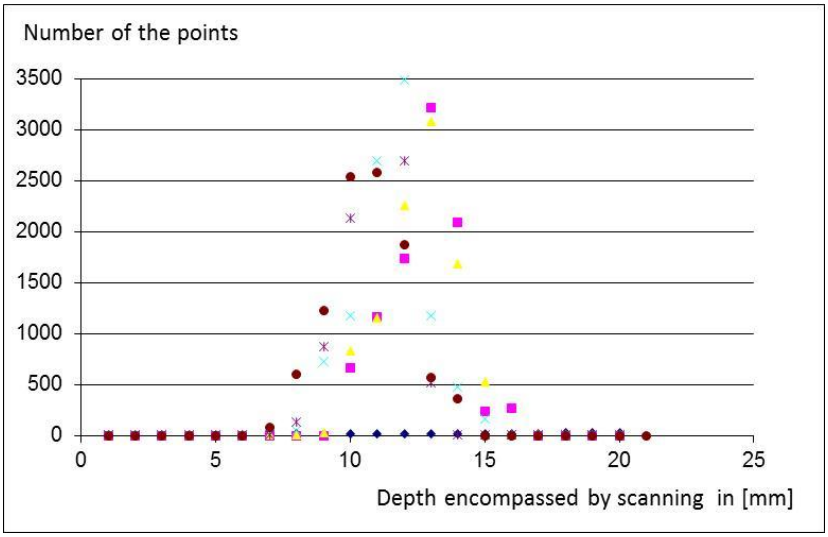


Figure 6. Number of the measured poins

The above figures well demonstrate that depth encompassed by scanning differs radically. This is the consequence of the difference in depth of compression of the chopping, which is a function of the fill factor of the chopping. Space filling depends on the size of the elements in the set. Therefore it can be stated that the deeper insight we gain into the set the larger the size of the elements we can find on the top layer.
Enumerating the points composing the cross-section we get the frequency.The following distribution is shown by the set of points in the cross section of the <4 mm classes after 5-times sampling.

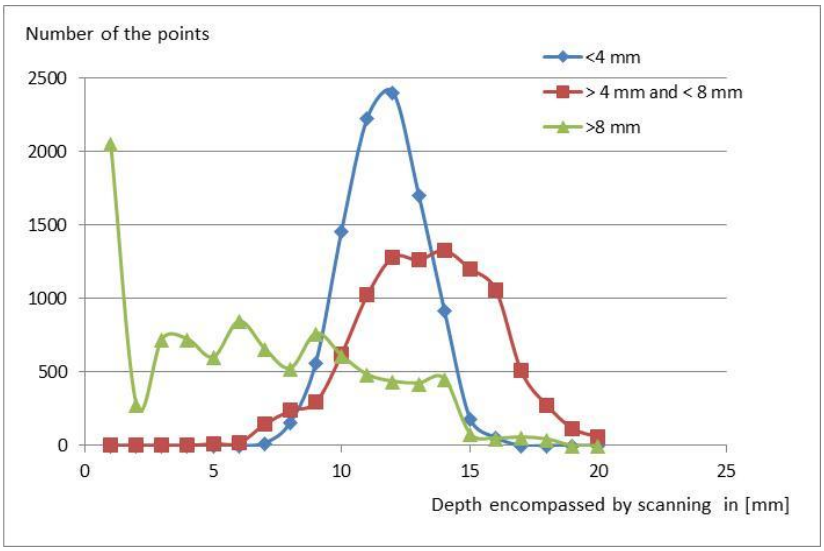


Figure 7. The shape of distribution

The figure above distribution is shown by the set of number of points averages in the cross section of the of the three size classes after 5-times sampling
The images well depict that in the case of the smallest size class the frequency values surpass element numbers of 3000, and more than 90% of the points are positioned between 7 and 16 mm. The next size class already spreads out to the interval of 5-20mm, and the largest values fall under 2000 elements. The third image shows the frequency value of the largest size class, with an even more outspread range of values.
Examining the standard deviation of frequency we get the following values (with 90 % from the number of points):
Case of <4 mm = 4
Case of >4 mm and <8 mm = 6
Case of >8 mm = unknown

This shows that spread of image points is larger as the elements composing the set examined become smaller.

Conclusions

Summarising the results, we can state that the images of wood-chopping recorded in 3D can provide a valuable starting point to draw conclusions on the size of the elements in the examined material set. This relationship can be demonstrated in the under 8mm size class with strong reliability. The larger size classes could be assessed to be less coherent with the applied method and equipment. Applicability can be improved further with larger element image generation. The ongoing research is expected to provide applicability in a broader size spectrum by further refining and modifying the applied methodology. Additional objective is the crosscut image generation without 3D imaging as this would result in a real-time industrial applicability by making evaluation faster and simpler.

Acknowledgements

This paper was sponsored by OTKA K68103 project.

References

Molto, E. et al (1996): An artificial vision system for fruit quality assessment Conference of European Agricultural Engineering, Madrid 1996, 96F-078

Bense, L. - Szendrő, P. – Vincze, Gy. (2006): Theoretical Questions of Burning-Purpose Use of Wood-Chips Congeries. Hungarian Agricultural Engineering, 2006, 19. 20-22.

Szalay, D. K. (2011): A terepi és laboratóriumi spektroradiometria szerepe a légi hiperspektrális technológiában. XVI Fiatal Műszakiak Tudományos Ülésszaka, Kolozsvár, Románia 2011.03.24-25. 305-308 pp. (In Hungarian)

Koschan, A. and Abidi, M. (2008): Digital Color Image Processing, Wiley-Interscience, New Jersey, USA, ISBN 978-0-470-14708-5

Gonzalez. R. C. and Woods R. E. (2008): Digital Image Processing, Pearson International Edition, New Jersey, USA, ISBN 978-0-13-505267-9

7. INVITED PAPERS

István RÉTI
Technical Background of Very High Efficiency Gaas Based Solar Cell Preparation
The author is research partner of the Faculty in the field of energetics, especially of solar energy utilization.

Péter TÓVÁRI, István SZABÓ, Mihály HERDOVICS, Tibor VOJTELA, László FENYVESI
Drying and Energetic Analysis of Wood Chip Pile
The authors are research partner of the Faculty in the field of agricultural engineering.

Moira MIRANDA, András EDELMAYER, Sándor MOLNÁR
Performance Verification of Advanced Filtering Alternatives for Robust Fault Tolerant State Estimation in Nonlinear Processes
The authors are research partner of the Faculty in the field of informatics.

Jozef RÉDL, Veronika VÁLIKOVÁ, Juraj MAGA
Application of Rotor Dynamics to Assessment of Liquid Lubricant's Properties
The authors are research partners of the Faculty in the field of tribology and agricultural machinery.

TECHNICAL BACKGROUND OF VERY HIGH EFFICIENCY GAAS BASED SOLAR CELL PREPARATION

István RÉTI

Research Institute for Technical Physics and Materials Science; Budapest, Hungary

Abstract

This work represents the electronical and mechatronical control-system for molecular source of our molecular beam epitaxial equipment. The molecular beam epitaxy system is an equipment for crystal growth, which machine opens the door to preparation of nanosystems under controlled circumstances. These nanostructures serve as building elements for the very high efficiency solar cells. The molecular beam epitaxial machine has different molecular sources. The regulation of these four molecular sources of the cells carried out partly mechanically and partly electronically. The mechanical control is realized with the help of an DC electro engine throgh a transmission, which moves the shutters over the sources. This mechatronical control is extended with a further possibility of the source regulation, wich is carried out with the help of the temperature control of the Knudsen cells. The control of motors and temperature of molecular sources happens with a PLC-based control system.

Keywords:

Molecular beam epitaxy, control, ultra-high vacuum

1. Introduction

A molecular beam epitaxy is the most shophisticated method for producing nanostructures. The grow of the nanostructures are made with the help of molecular beam in ultra high vacuum. The use of the nanostructures revolutionized the material science and the electronics - for instance solar cells (PV panel) [Farkas, I 2010]. The efficiency can substantially raised by the help of the nanostructures. The GaAs based solar cells can reach high efficiency by using nanostructures. The efficiency of the multi-quantum-well solar cell MQW-SC exceeds 40%, the intermediate-band-quantum-dot solar cell IBQD-SC exceeds 60% [Nemcsics, A. 2001, Nemcsics, Á. 2009]. The condition of the fabrication of the above mentioned GaAs-based solar cell structures is the growth of controlled semiconductor crystal-layers and nanostructures. These few atom-line layers and other nano-sized objects can be realized particularly with molecular-beam-epitaxy. One of the classical methods to prepare thin films with

clean surface is the evaporation in ultra-high-vacuum (UHV). Depending on the substrate surface and on the evaporation conditions, the deposited films can be crystalline. In this case they are epitaxial films and the preparation method (and the equipment) is then called MBE [Herman, M. A. Sitter, H. 1989]. Surfaces of such epitaxial films grown under UHV conditions are ideal for layer and surface studies [Hanlon, J. F. O. 1989]. MBE is the most versatile technique for preparing clean and well-defined surfaces, interfaces, layers and nanostructures of different semiconductors [Nemcsics, A. Heyn, Ch. Stemmann, A. Schramm, A. Welsch, H. Hansen, W. 2009]. We present in this paper the mechatronic system of the molecular beam regulation connecting to this technology.

2. Description of the MBE Equipment

The experimental setup described here is the only one presently in whole Hungary. Figure 1. shows our UHV system with facilities for MBE. Recently, some developments were carried out on the setup [Nemcsics, A. Réti, I. Serényi, M. Tényi, G. V. Hodován, R. Gábor, J. Taar, I. G. Pántos, J. Bozsik, J. Molnár, S. Jankóné Rózsa, M. 2009, Nemcsics, A. Réti, I. Tényi, G. V. Kucsera, P. Tóth, L. Harmat, P. Amadou, M. Csutorás, M. Kupás-Deák, B. Sándor, T. Bozsik, J. 2010]. This equipment consists of three chambers. One of them is the reactor for sample preparation. Another one is the loading chamber, while the third one is an additional chamber for vacuum pumps. Vacuum valves are between the chambers. Knudsen-type crucibles are used as effusion cells for the evaporation.

Figure 1. View of the molecular-beam epitaxial laboratory

Our molecular sources are As, Ga, In and Al. These cells are tubular crucibles made from pyrolytic boron nitride. The crucibles are mounted within spiral Ta heater windings which are themselves enclosed within Ta-foil radiation shields. A requirement for the source oven and the whole unit is a very low production of impurities in the molecular beam. The oven set-up is therefore surrounded by a water cooled panel on which shutters are mounted which can close and open one of the other effusion cell. Also the space between sources and the sample is shielded by a cooling shield at liquid N_2. The sample is mounted on the Mo holder by liquid In. Our equipment has a 12 keV reflection high-energy electron diffraction (RHEED) for in-situ investigation. A fluorescent screen is located in opposite of the RHEED electron gun. The sample can be moved by magnetic rod.

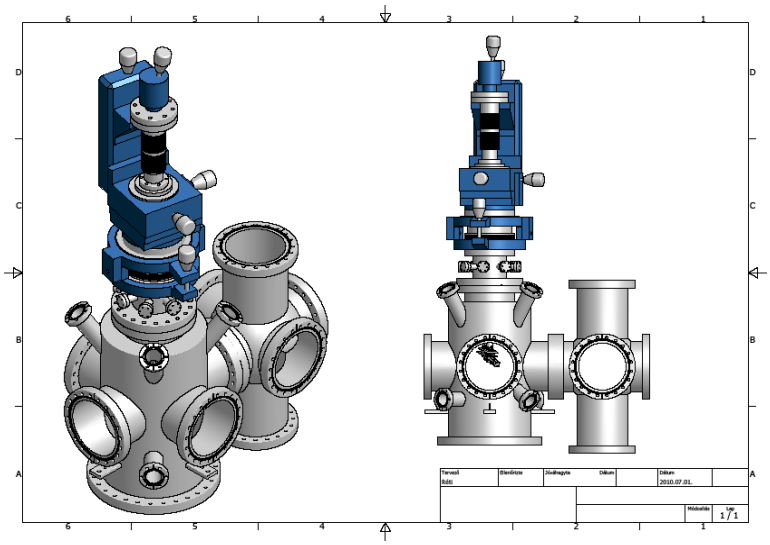


Figure 2. Sketch of the main chamber of the molecular beam epitaxial equipment

Structure growing takes place in an UHV chamber. To reach the vacuum of 10^{-10} and 10^{-11} mbar we have to heat our system for which we used a PLC- controlled regulation [Description of the Siemens Simatic S7PLC system]. To raise the efficiency of heating we have to fasten the heating wires to the UHV chamber. We used a special fixing mechanics for this [Bugyjás, J. 2010]. To make heating consistent the system has to be encased.Crystal growing takes place in the main chamber (Fig.2). An auxiliary chamber is connected to the main chamber, where the outlets for a turbomolecular (Balzers TPU 520) and an ion getter vacuum pump (Varian Diode) are located. The ion getter pump can be isolated from the main chamber by a valve (Fig. 3.). The third chamber is a load lock, which enables the simultaneous loading of several samples into the system. Thus several growing processes could be made without opening the vacuum system.

The load lock is pumped by a small turbomolecular pump (Pfeiffer HiPace 80), and can be separated from the main chamber with a valve. Samples are moved inside the system with the help of magnetic rods. The chamber pressure is monitored with cold-cathode ionization gauges, temperature control is realized with thermal sensors and electric heaters. In situ growth monitoring is done with a 12 keV RHEED [Description of the Siemens Simatic S7PLC system] device (Riber CER 606 – ACE 1010).

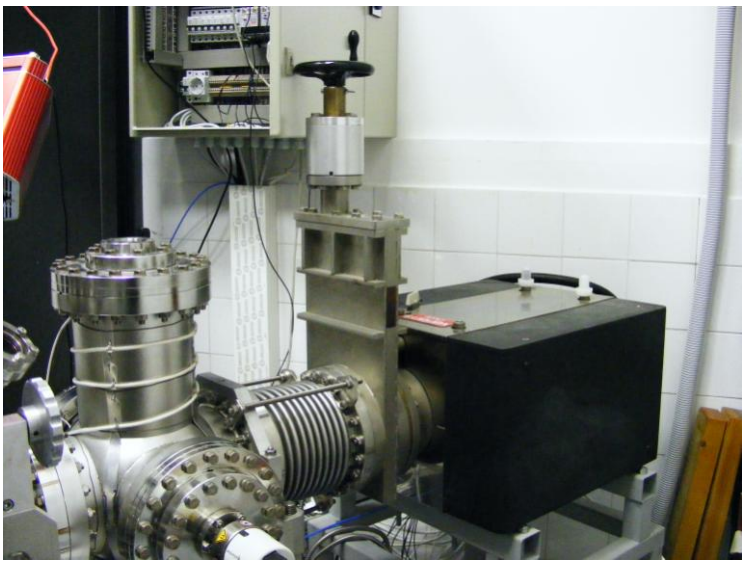


Figure 3. The valved Varian diode

3. Description of the Heating System

Development of the system meant providing solutions for controlling problems (such as heating of the chamber, substrates and Knudsen cells, or moving the shutters) as well as image processing (RHEED evaluation) and micromechanical problems (fixing the substrate to the holder, moving the substrate holder, connecting stepping motors and rotating bushings). Image evaluation and process control creates a feedback system, enabling e.g. the fabrication of super lattices, or creation of structures with subcritical layer thickness. To reach ultra high vacuum the chamber must be heated. For the free path and small residual impurity in the molecular beam epitaxy, 10^{-10} –10^{-11} mbar vacuum should be attained. The chambers must be fumigated by controlled heating. For the steady heating, the chambers must be whipped round with heating wire, measuring points must be developed and must be packed. The numbers and locations of the heating wires and measuring points are really very important. For the packing of the very complicated developed vacuum chamber, for the producing the

emission of several components and introspection windows, a full computational documentation has to be created. The temperaure measurement is done with the observance of the background's partial pressure. The control happens with PLC. The heating system can be seen in Fig. 4.

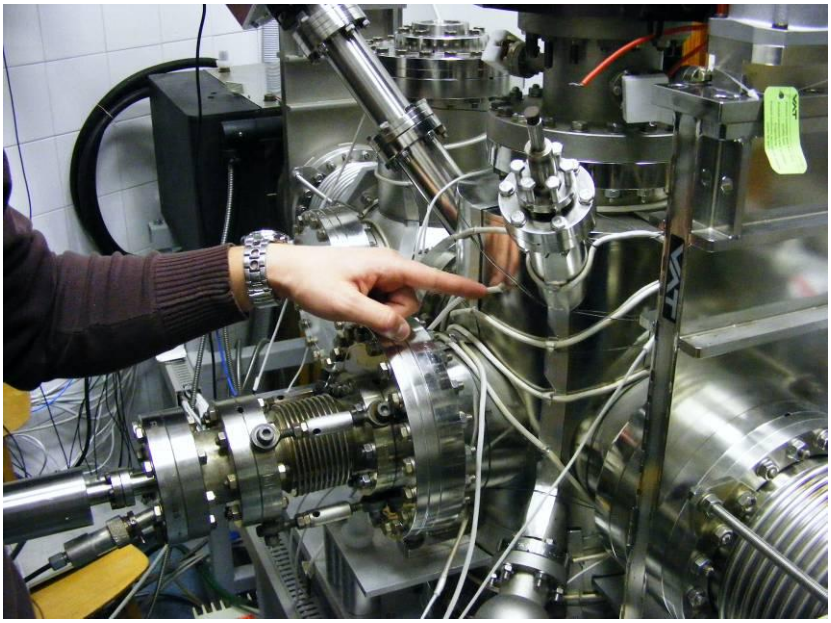


Figure 4. Mounting of the heating-fibres

The MBE growth needs very high requirements with the equipment. So we have to use a reliable and precise control [Nemcsics, A. Réti, I. Serényi, M. Tényi, G. V. Hodován, R. Gábor, J. Taar, I. G. Pántos, J. Bozsik, J. Molnár, S. Jankóné Rózsa, M. 2009]. From the view-point of electronics, the control is slow. The growth rate ranges between 0.1 – 1 ML/sec (mono layer, ML). In our case, the UHV condition means 10^{-10} – 10^{-11} mbar pressure. After the closing of the system, we need the outheating of the chambers to achieve this vacuum level. The outheating needs slow control without overshoot. In the case of the shutter movement, we need slower control, because the shutter has to close before the next ML begins to grow. The heating of the effusion cells and the sample holder are also relatively slow processes because we need processes free from overshoot. Because of the reliability, we use PLC system for the control of the MBE.

A PLC (Programmable Logical Controller) is an industrial computer used in automation of wide range of processes. They are used in many industries and machines. Unlike general-purpose computers, it is designed for multiple inputs and outputs (I/O) arrangements, extended temperature ranges, immunity to

electrical noise, and has very reliable hardware structure. Programs to control machine operation are typically stored in non-volatile memory. A PLC is an example of a realtime system since output must be produced in response to input conditions within a bounded time, otherwise unintended operation will result. The main difference from other computers is that PLCs are armored for severe conditions, and have the facility for extensive I/O arrangement. These connect the PLC to sensors and actuators.

As we mentioned, the PLC gives reliable control, therefore we use it in the control of our MBE equipment. The type of the used system is Siemens Simatic S7. The system consists of the central rack with the CPU, 24V power supply and the I/O modules. In the same cabinet the protecting equipments, fuses and the relays are placed. The supervision of the system is possible by using an operator computer which is in this case a touch display with function keys. The detailed parameters of this PLC system are described elsewhere [Description of the Siemens Simatic S7PLC system] (Fig. 5.).

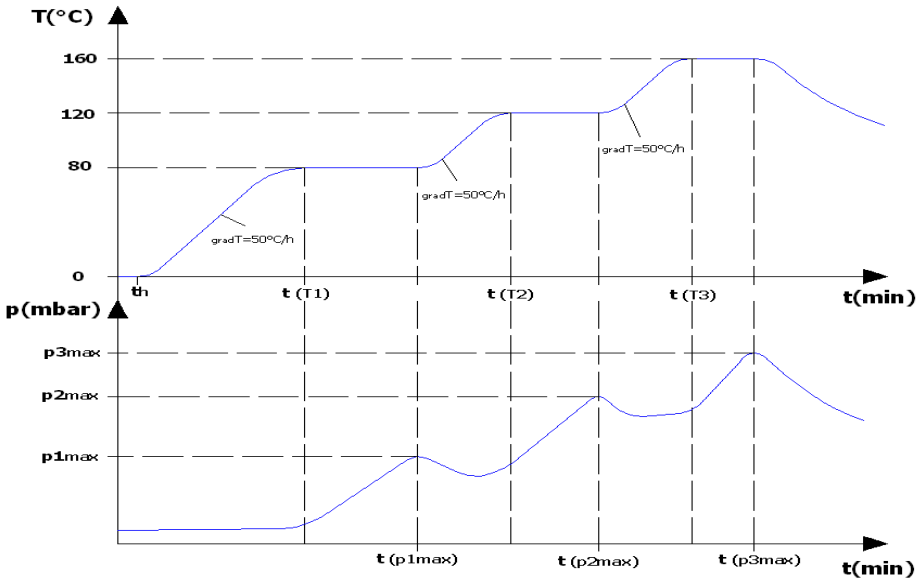


Figure 5. Temporal course of the out-heating process

Before the crystal growing procedure starts, the equipment needs to be heated up, following a specific temperature characteristics to evaporate different contaminating materials and clean the vacuum chamber. The temperature is measured at 6 different points and 6 heating wires are controlled. The closed loop heat control is done by the above mentioned Simatic S7 PLC which runs PI control algorithms to ensure error free tracking of the temperature setpoint. The whole system is operated from a PC operator station (OS). The OS reads and

writes the process variables of the PLC. The temperature characteristics and other functions such as the sequential control of the molecular sources can be parameterized on the OS (and controlled by the PLC). The system was built with the contribution of students.
Heating cables had to be placed on the chamber evenly. The stainless steel chamber walls must not be welded, nor holes drilled into it, thus a special solution for a looser fixture became necessary. The fixing elements were folded from aluminium and are presented on Fig. 4.
In positioning the heating cables we had to take into consideration that parts of the system must not be heated above 200°C, such as the chamber window or the manipulator. The number and distribution of the Isopad heating cables for each chamber were determined according to the weight and surface area of the modules.
The heating process consists of three steps. First, the chambers are heated up to 80°C. Simultaneously, contaminants peel off from the chamber walls, increasing system pressure, thus it is important to operate the vacuum pumps during this step. After reaching T_1=80°C, the temperature is kept constant. This is necessary as contaminants take time to detach from the walls. Rapid heating up of the system could cause contamination to burn to the walls, preventing reaching required vacuum pressure. The second step of the heating profile starts after reaching $3x10^{-6}$ mbar system pressure. Maximum temperature in this step is T_2=120°C. Further desorption of contaminants from the walls will increase vacuum pressure again, thus at the end of this step another constant temperature heating period takes place, until system pressure returns to $3x10^{-6}$ mbar. The third step consists of heating the system up to T_3=160°C and keeping this temperature. This step can last for days, depending on how long the chambers had been vented. Program controlled gradual heating of the system is of great significance, as a sudden rise in temperature could result in contamination burning to the chamber walls, which inturn could prevent the system from reaching the required chamber pressure.

4. The mechanic and electronics of the shutter-moving

The strict, according to plan dosage of molecular beam has an important role in the programming layer-grow. It happens with the heating of the current source (Knudsen cell) (Fig. 6.), then conducting the leaving molecules through a cryogenic pipe section. This pipe (like a cannon-conduit) points to the wafer.

With the opening and cover of this cannon-conduits mouth, we can handle the amount of the molecules, that come to the model. It means, that we fixed a cover-disc to the end of a stick which enters to a vacuum-space, that parallels with cannon-conduits. It turns to the wanted direction through a transcription by a DC motor. These turning-mechanisms were earlier manual controlled and were build-in not to restrict each others movings. Because of this, according to a

normal base area, we had to turn four different directional and length stick, that ends are thickened. We solved this by putting the four actuation motor to the common base area and each turns a fork. We fixed the cogs - which are running to the forks - to the mechanisms that turns the blende. Hereby, even the angle difference of the motor - and blende-axis, it managed to realize the motor turn with a minimal backlash. We restrained the slew of the forks turning with the common base area switch, to let the blendes do their allowed moving.

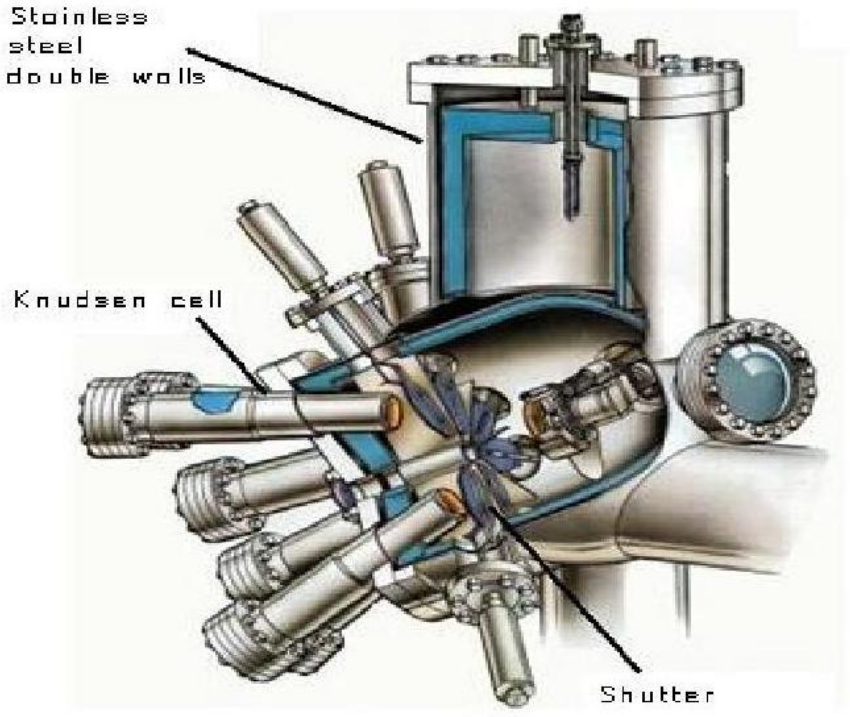


Figure 6. View of Knudsen cells with shutter

Thus the motor-driven moving of the aperture has come true, the possibility of programmable electronical controll has been provided. The principle of controlling is a Siemens S7-300 PLC. We've chosen the blanket statement of the Knudsen cells for starting point. The program does this with each motor in every inrush and only after this marks back its control capability. Thus the motor works under 12V DC we solved the rotation direction problem with polarity change. The micro-switches fixed to the last-state got two functions (Fig. 7.).

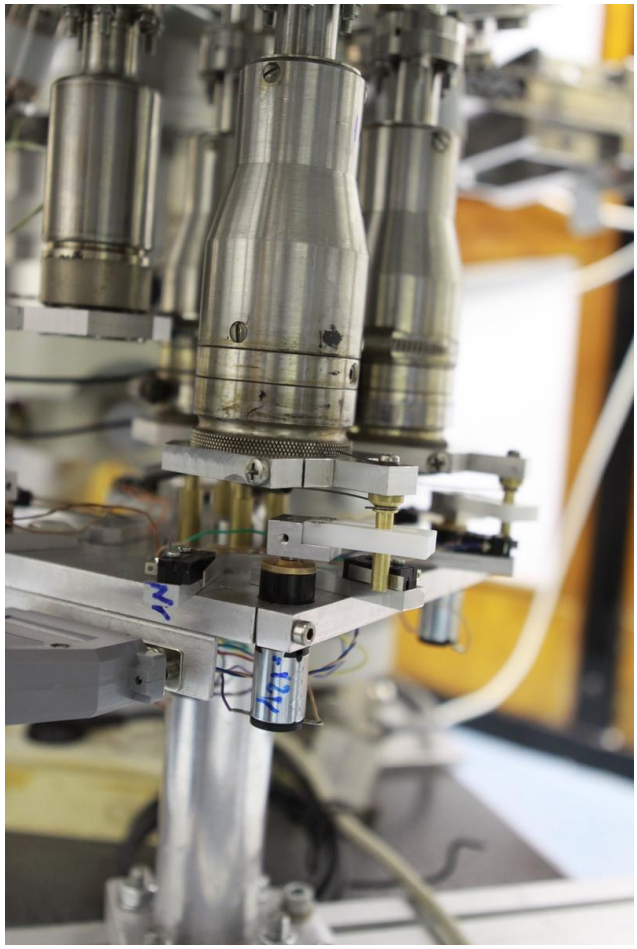


Figure 7. Realization of shutter-mover of our MBE

On one hand they give a signal if the motor's rotator switches them to opened or closed state, and on the other hand their (switching) on/off states is a signal for a motor-observer, which ensures the safety from random overloads. So if either driver-motor stops running for some reason, it only gets current for 4 seconds at most. We use two relays for controlling every motor. With the circuit diagram below it's easy to follow the controlling.

We can follow the all-time status of the Knudsen cells from a perspicuous window of the control program. Here we can see the momentary temperature and the open or close status of the cells. This operator surface completion is made with the help of WinCC flexible program. (Fig.8.)

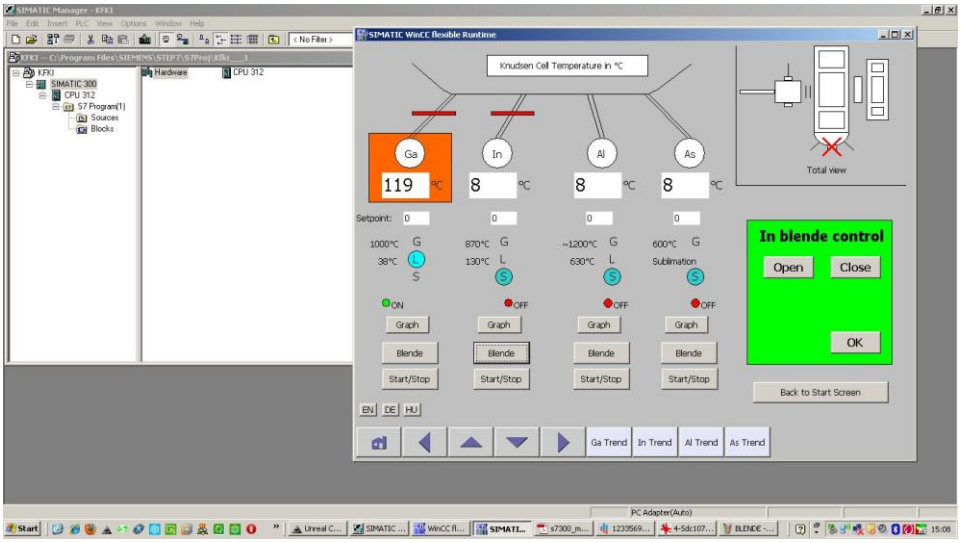


Figure 8. Operator surface of our MBE

5. Conclusion

In this paper, precision manipulator was carried out with the help of composition of a simple linear and simple rotational mover. This self-made solution is more economically than an originally precisions manipulator. We realized a electromechanical equipment for the shutter movement without slip. We developed a new sample holder. An outheating system was realized for the MBE mashine. Optimal positions of the temperature sensors were choosen for very homogeneous chamber outheating.

Acknowledgement

I want to thank Ákos Nemcsics and István Farkas for the mental support and Gergő Bátori, Antal Ürmös and András Straszner for the technical help.

References

Farkas, I. (2010): Possibilities of the solar energy utilization in Hungary, Magyar Tudomány, Vol. 8, pp 937-946 (in Hungarian).

Nemcsics, Á. (2001): The solar cell and its development perspectives, Akadémiai Kiadó, Budapest, (in Hungarian).

Nemcsics, Á. (2009): New directions in solar cell technology, Elektronikai Technológia, Mikrotechnika 48/1, p 39-45 (in Hungarian).
Herman, M. A. Sitter, H. (1989): Molecular Beam Epitaxy – Fundamentals and Current Status, Springer-Verlag, Berlin.
Hanlon, J. F. O. (1989): User Guide to Vacuum Technology, Wiley-Interscience, New York
Nemcsics, Á. Heyn, Ch. Stemmann, A. Schramm, A. Welsch, H. Hansen, W. (2009): The RHEED tracking of the droplet epitaxial grown quantum dot and ring structures, Mat. Sci. Eng. B **165** , pp. 118-121.
Nemcsics, Á. Réti, I. Serényi, M. Tényi, G. V. Hodován, R. Gábor, J. Taar, I. G. Pántos, J. Bozsik, J. Molnár, S. Jankóné Rózsa, M. (2009): Summary of MBE installation in BMF-MIT and MTA-MFA corporate laboratory, Elektronikai Technológia, Mikrotechnika 48,, pp. 33-35 (in Hungarian).
Nemcsics, Á. Réti, I. Tényi, G. V. Kucsera, P. Tóth, L. Harmat, P. Amadou, M. Csutorás, M. Kupás-Deák, B. Sándor, T. Bozsik, J. (2010): Technological conditions of nanostructures preparations with MBE, Gép LXI, pp. 29-32 (in Hungarian).
Description of the Siemens Simatic S7PLC system
Bugyjás, J. (2010): Elements of electromechanical construction, ÓE-LVL-2019, Budapest, (in Hungarian).

DRYING AND ENERGETIC ANALYSIS OF WOOD CHIP PILE

Péter TÓVÁRI[1], István SZABÓ[2], Mihály HERDOVICS[1], Tibor VOJTELA[1], László FENYVESI[1]

[1] Hungarian Institute of Agricultural Engineering (MGI)
[2] Department of Machinery, Institute for Mechanics and Machinery

Abstract

The purpose of this research work was to determine the change of the pile characteristics of drying process. Air ventilation was used for the investigation by which we could determine the drying process of the wood-chip pile. Summer statistic values of several years were utilized for the choice of appropriate drying temperature. Air-drying system of wood-chip pile was investigated under different designs and circumstances. During the air-drying process we can avoid the most expensive part of the drying: air heating. In this case we achieve a more economical solution and we can use a drying, which causes less loading of environment.

Keywords

Biomass, drying, wood chip, moisture content, sustainable

Introduction

One of the most important energetic and technological properties of the wood-chip is the moisture content, because over 35% biodegradation processes are take place in the pile. The result demonstrates that organic matter content and caloric value are reduced, so the available energetic efficiency is decreased under the utilization. Hot air is commonly used in drying procedures,. It is important to consider that the energy price rises and the decrement of energy carriers so decreasing of energy consumption is also required as well as the development of new procedures. The air-drying of wood-chip pile is carried out by the drying effect of suitable ventilation.. The purpose of our research was to investigate the different drying procedures – under laboratory conditions – which have lower power consumption making a more cost- effectiveness drying possible (Beke et al., 1985;Fekete and Menyhárt, 1975).

Material and methods

Investigations of wood-chip pile drying were in a I. measuring circuit – in the laboratory
In the 1st measuring circuit we placed a drying cabinet on a scale wherein the temperature could be regulated. An exhaust fan was responsible for the ventilation in the drying cabinet which is connected to an outgoing chunk. 25 ^{o}C temperature was set in the drying cabinet and the samples were dried on constant weight. Fig. 1 shows the design of the I. measuring circuit.

Fig. 1. I. measuring circuit – Laboratory investigation

Fig. 2. II. measuring circuit – Pilot scale investigation

Weight loss was recorded continuously, which is originated from the decreasing of moisture content. Different intensities of exhaustions were used while repeating the investigations, - namely variable ventilation number. As a result for the modification of ventilation number different drying curves were obtained with the help of these drying intensity became evaluable and comparable (Hrenkó, 1988).

II. Measuring circuit – Pilot scale investigations
In this measuring circuit a lower ventilation number (compared with the value of 1st measuring circuit) was utilized, because of the wider ranging investigations of drying. The temperate measuring place was prepared, where the low ventilation was provided continuously. The previously defined 25 °C temperature was used in the measuring place, too. In this measuring circuit, where we placed a platform – made of PVC pipe – on a scale, whereat we put the previously used close- leaky stock. The ambient air exhaustion through the wood-chip pile was occurred through the platform by the same axial fan which was previously used. Fig. 2 shows the design of the II. measuring circuit (Komka, 2002).

Results and discussion

The results of the drying investigations were evaluated with widely used procedures which can be found in professional literature. Table 1. demonstrates the most significant results

Table 1. The results of calculations

	I. measuring circuit		II. Measuring circuit	
	Q= 2.62 [m^3/h]	Q = 5.54 [m^3/h]	Q= 2.62 [m^3/h]	Q = 5.54 [m^3/h]
Stock volume: V [m^3]	0.01702		-	
Drying cabinet volume: $V_{drying\ cabinet.}$ [m^3]	0.02464		0.02464	
Drying time: T [h]	286	211	191	211,5
Total exhausted air: $V_{air.}$ [m^3]	749.32	1168.94	500.42	1171.71
Weight of the exhausted air: $m_{air.}$ [kg]	899.184	1402.728	600.5	1406.052
Quantity of the carried water m_{water} [g]	1020	1290	1045	1115
Specific dehydration: [g/$kg_{air.}$]	1.13	0.91	1.74	0.79
Efficiency of dehydration: [g/h]	3.5	6.11	5.47	5.27
Originated ventilation number: [m^3/h/m^3]	106.3	224.8	n/a	n/a

As a result of the 1st. measuring circuit it can be determined that the higher air volume utilizes the vapour capacity of air less. The amount of water carried by 1 kg air (specific) is less, because the air has less time for regaining moisture. However, this difference is not significant, thus the higher air volume - on the whole - is able to carry more water which can be seen clearly in desiccation efficiency. Fig.3. shows the drying curve which contains the different results of investigations regarding on air speed (repeated three times).
As shown in the results of Fig 3, during the drying process, the drying speed is decreased until it reaches the point where there is no moisture loss. The drying efficiency has the same results like the 1st. measuring circuit. In case of the utilization of higher air volume it can be determined that the amount of carried water is significantly less, thus the air quickly pass through the wood-chip pile and it has not enough time for water absorption. The calculated desiccation efficiency also shows that the higher amount of air is able to carry less water hourly, which causes higher energy consumption. Fig.4. shows the drying curve which contains the different results of investigations regarding on air speed (repeated three times).

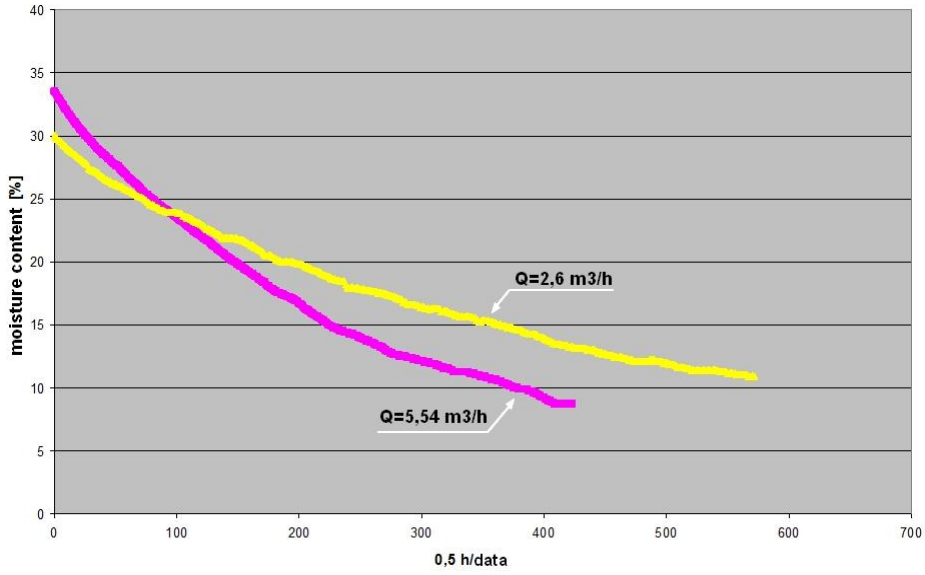


Fig.3. Drying diagram of the laboratory measuring circuit

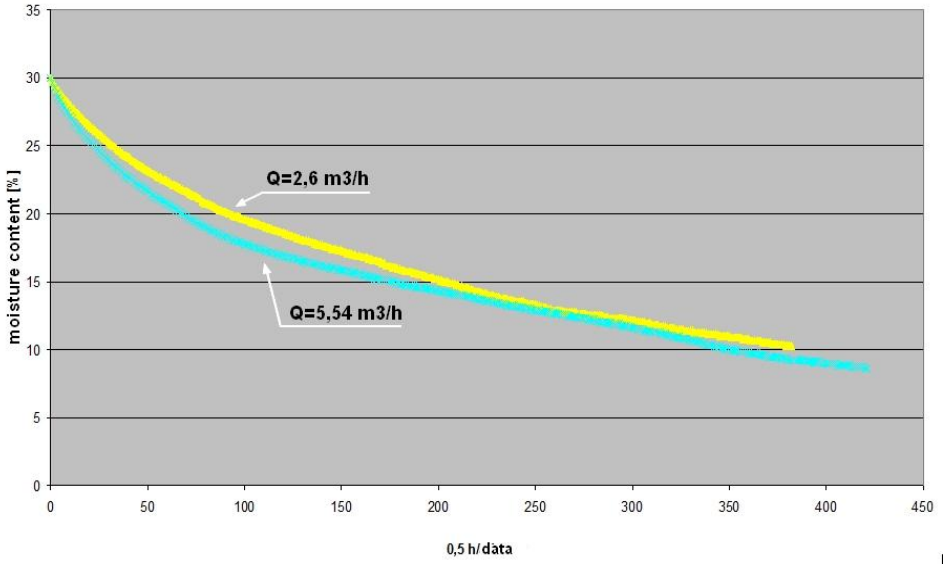


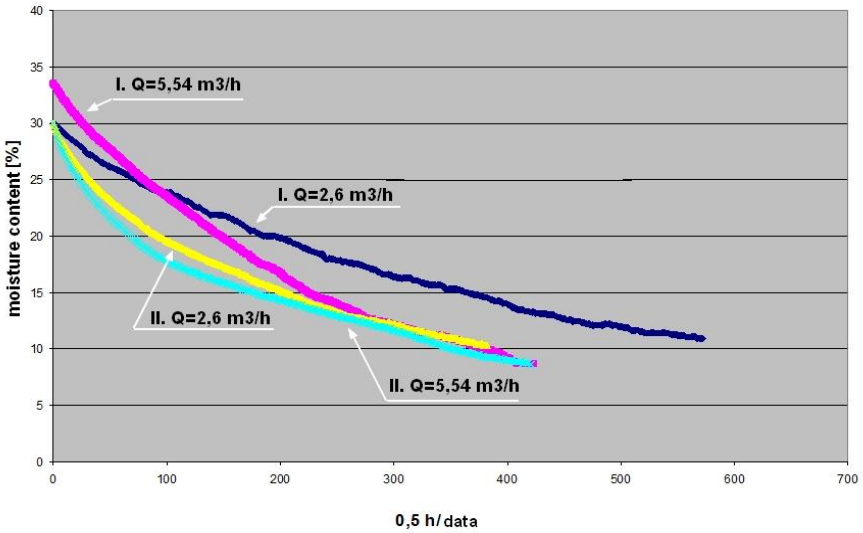
Fig. 4. Drying diagram of the pilot scale measuring circuit

After the drying investigations (in both measuring circuit), moisture content distribution of wood-chip was analyzed within the stock. In the 1st. measuring circuit the moisture content is lower in the external parts of the stock than inside after drying. 25 °C temperature was adjusted in the drying cabinet which has direct contact with outside layer. Inside the drying cabinet due to the evaporation lower temperature was obtained, so the balance of moisture content was set in higher value. The results of the 2nd. measuring circuit demonstrate that the moisture content distribution is balanced. There are no significant differences. The evenness can be explained by the fact that unlike the previous ones the ambient air was exhausted through the wood-chip, so close similar temperature was obtained inside the pile. Due to this that balance of moisture content was set in the same result.

Conclusions

In the interest of comparison, a so-called cumulated drying diagram was compiled. This graph provides graphical comparison.
The graph demonstrates that the higher air volume allows faster drying in the beginning of the process, but after a certain limit is no longer economic. Passed through air has not enough time to absorb sufficient water vapor, so this is the cause of the dissipation of energy. In the case of 1st. measuring circuit this effect is not observed, because the speed of the exhaustion causes a faster finish of the investigation series. This can be explained due to the relatively small humidity volume is higher in the drying cabinet than in the ambient air. In accordance

with these things the higher air can be saturated. Because of the higher volume of the measuring place, the humidity is lower, therefore the drying air needs more time to achieve saturation point. From an economic point of view, lower exhaustion of air is more sufficient, which is more economical. In case of drying in measuring place drying curves demonstrate that, the first part of the drying is more abrupt, because the ambient air passes through the whole pile. This is the reason why that higher temperature arose inside than in the other measuring circuit. Higher air temperature can absorb more vapour, thus at the same time more water passes from the pile. In the case of 1st measuring circuit the exhaustion of air as compared with the pile is implemented in point mode. For this reason the drying zone is formed in a concentrated way. The statement is confirmed by the measurement of the moisture content distribution. In case of the 2nd measuring circuit the exhaustion of air is realized on the whole surface, therefore the moisture content distribution is balanced. On the grounds of investigations it can be declared that the utilization of cold-air drying provides a suitable procedure to achieve the appropriate moisture content value. It is important to choose the sufficient drying intensity according to the characteristics and the volume of the material. The necessary air volume for a unit value of the drying material is easily definable with calculations considering profitability as well as drying time.



5. Fig. Cumulated drying diagram

References

Beke J., Várkonyi J., Vas A. (1985): Agricultural crop drying, Mezőgazdasági Kiadó (in Hungarian)
Fekete I. – Menyhárt J.(1975): Theoretical basis of air engineering, Műszaki Könyvkiadó, (in Hungarian)
Hrenkó J.(1988): Final report: Storage experiments of briquetting straw with the aim of quality preservation FM Műszaki Intézet, (in Hungarian)
Imre L.(1974): Drying handbook, Műszaki Könyvkiadó, (in Hungarian)
Komka Gy.(2002): Report: Development test of LM-Type ramping dryer ventilation system, FVM Műszaki Intézet, (in Hungarian)

PERFORMANCE VERIFICATION OF ADVANCED FILTERING ALTERNATIVES FOR ROBUST FAULT TOLERANT STATE ESTIMATION IN NONLINEAR PROCESSES

Moira MIRANDA[1], András EDELMAYER[2], Sándor MOLNÁR[3]

[1] Laboratory of Unit Operations, Department of Chemical Engineering
Faculty of Engineering, University of Los Andes, 5101 Mérida, Venezuela.
[2] Systems and Control Laboratory, Computer and Automation Research Institute
Hungarian Academy of Sciences, H-1111, Budapest, XI, Kende u. 13-17. Hungary.
[3]Institute of Mathematics and Informatics, Faculty of Mechanical Engineering
Szent István University, H-2103, Gödöllő, Páter Károly u. 1, Hungary.

Abstract

In this paper a comparative study of alternative filtering solutions for robust state estimation in nonlinear systems is presented. The main issue the paper addresses is to find solution to make the state estimation robust against external disturbances, sensor failures and nuisance variation of the quality of the measurement signals. In the first part of the work the performance properties of the extended and unscented Kalman filters are discussed to reveal the differences, and congruencies of the two different approaches of filters' synthesis. Design features, such as the complexity of implementation and the sensitivity of the state initialization process on uncertain covariance information at regular filters startup are evaluated in a variety of real plant scenarios. The second part of the paper presents a dual layer approach for robust fault tolerant estimation of nonlinear processes using a combined filtering strategy. In this solution, from the one hand, the filter is made robust in face of environment uncertainty using filter adaptation. To this end, the filter identifies the measurement covariance by means of recursive estimation, upon which the adaptation relies, to suppress the effect of sporadic variations in the quality of measurements as well as compensates for incipient sensor faults. From the other hand, fault monitoring is continuously applied to the filter's innovation in an attempt to initiate filter reconfiguration when the adaptation mechanism alone is not able to overcome the failure. The discussion of the results is embedded in the case study of state estimation of a batch distillation process.

Keywords

Adaptive Kalman filter, Dependable systems, Enhanced state estimation, Extended Kalman filter, Sensor fault tolerance, Filter reconfiguration, Unscented filtering.

1. Introduction

An impressive number of references in the literature has been devoted to state and parameter estimation techniques for linear and nonlinear systems in the past decades. When selecting from the range of estimation methods for the solution of a particular estimation problem, at least two fundamental properties of the dynamical system at hand have to be properly appreciated. They are the phenomena of modeling uncertainty and the nature of the exogenous disturbances. These properties, in the most cases, may determine whether one formulate the estimation problem in linear or nonlinear and in deterministic or stochastic settings. Different approaches provide different answers to the treatment of the problems.

Based on the assumptions that (i) modeling uncertainty can be safely disregarded and (ii) the disturbance statistics of the system is reliably known, i.e., it can be modeled as a Gaussian random process with known mean and covariance, for state estimation in linear systems the traditional Kalman filter (KF) has been unrivaled for decades. This is because of its optimality and robustness and the relatively simple model of implementation that makes it a viable solution on many platforms.

When it came to state estimation for nonlinear systems, the corresponding nonlinear filter is obtained by straightforward extension of the linear solution, resulting in the formulation of the extended Kalman filter (EKF) (A. Gelb, 1974; P. S. Maybeck 1982). Beyond doubts, in spite of its varying performance in terms of estimation accuracy, ease of implementation, robustness, and computational burden, the EKF still dominates the nonlinear state and parameter estimation techniques (H. W. Sorenson, 1985). There are two fields of applications where EKF attracts distinguished attention: process control and chemical engineering are motivated by the simple implementation and the favoring behavior of EKF in the presence of smooth nonlinear dynamics. Localization and mapping, including dynamic map building technologies, to estimate the state vector containing both the pose (position and orientation) of a mobile unit and the landmark locations of the infrastructure, is another area where EKF has become ultimately popular in the past years, see (S. J. Julier and J. K. Uhlmann, 1985; U. Frese, 2006; T. Bailey at all, 2006; A. Edelmayer at all, 2010).

Quite interestingly, just the most successful applications of the target fields characterized above reveal the EKF solution unsustainable in certain situations.

EKF approximates the state distribution as Gaussian random variable (GRV) and can handle only a limited amount of nonlinearity. This is one of the well-known limitations of the traditional EKF that it linearizes the nonlinear model in a way that the traditional linear KF can be applied. The linearization is based on first-order Taylor method, which approximate, in every time step, the nonlinear state transition and observation equations about the estimated state trajectory, with a linear time varying (LTV) system. This approach is problematic from not a

single point of view. First, the linearization tends to introduce large errors and even cause instability of the filter. Moreover, not paying attention to rudimentary system variations, stability issues may potentially arise when the noise covariances increase and/or change, as compared to the dynamics of the system.
Sporadically changing intensity of measurement noise due to e.g., communication nonperformance in the sensor network, however, is an issue both in advanced process control and localization applications (A. Edelmayer at all, 2010) that are increasingly implemented as distributed operations. Moreover, EKF necessitates the use of the derivatives of the state and measurement variables, which calls for the calculation of the Jacobian and its matrix exponential, which is a computationally sensitive and expensive operation (C. Moler and C. Van Loan, 1978). One of the reasons for inconsistency of EKF-based localization solutions is due to the violation of some fundamental constraints governing the relationship between various Jacobians when they are evaluated at the current state estimate (S. Huang and G. Dissanayake, 2007).
Another distinctive attribute of the filter is that it is required to converge to the actual state even when it is initialized with partially known initial conditions. This is an important feature (both in process control and localization), because the filter's operation, more or less frequently, is subject to reset and resume, regularly. Reset mode, sometimes, is an inherent feature of the operational strategy of the filters (A. Edelmayer at all, 2010). This is to be done in a process environment when the initial conditions are rarely known. In these situations traditional EKF is difficult to tune, its convergence time may be varied, the resetting conditions and the Jacobian can be hard to derive.
There have been a few innovations of the EKF in the past decade. Though particle filters (PFs), (i.e., another nonlinear derivative of the KF), can handle arbitrary distributions and nonlinearities reportedly providing better performance than EKF in general, they are computationally very complex to derive and as such, there is no chance to apply them in resource constrained embedded applications. Unscented filtering (UF) solutions (S. J. Julier at all, 1995), which approximate Gaussian type distributions over a fixed number of deterministic parameters, are based on the expectation that it gives a performance tradeoff between PF and EKF.
Though particle filters (PFs), (i.e., another derivative of the KF), can handle arbitrary distributions and nonlinearities reportedly providing better performance than EKF in general, they are computationally very complex to derive. Correspondingly, there is no chance to apply them in resource constrained embedded computation. In the meanwhile, (S. J. Julier at all, 1995) and (S. Julier at all, 2000) introduced a new generalization of the Kalman filter, the unscented filter (UF) that approximate Gaussian distributions over a fixed number of parameters. A relatively new development in nonlinear state and parameter estimation is the use of the UF idea in an ever-increasing scale; see (A. Romanenko and J. A. A. M. Castro, 2004; M. Nørgaard, N. K. Poulsen, and O. Ravn, 2000; R. M. Oisiovici and S. L. Cruz, 2000; R. M. Oisiovici and S. L.

Cruz, 2001). This is based on the expectation that unscented filtering might give a performance tradeoff between PF and EKF.

In a current laboratory work program we aim to collect experience with the design, implementation and operation of alternative filtering methods in an attempt to eliminate the shortcomings of the traditional EKF solution especially in distributed filtering applications, resembling to (A. Edelmayer at all, 2010). The objective is to select implementation candidates amenable to embedded system applications. In a series of laboratory experiments, therefore, we investigated a number of applications that pose challenging problems substantially similar to the ones occurring in high-accuracy state estimation in nonlinear systems. The work with real, non-simulated process data and access to real plant measurements got priority in this project. Thus, among others, we arrived to the investigation of some popular chemical engineering applications, i.e., batch and reactive distillations, whose pilot plants could be rendered at our disposal for experimentation.

In this paper the characteristics of advanced filtering solutions in view of the above problem formulation are presented and their performance is characterized in a case study of a batch distillation application. Solution alternatives in two different filtering approaches are investigated. The objective in both approaches is to explore and verify the capabilities of the particular solutions, i.e., unscented and adaptive filtering, in the enhancement of estimation accuracy and robustness of state estimation in process environments where the application of the EKF scheme is prevented, or at least severely restricted, by one of the above mentioned circumstances.

In the first part of this paper the EKF and UF solutions in a system state estimation framework applied to the distillation process are investigated in a comparative study. The objective is to contrast the performance of the particular filter implementations in terms of estimation accuracy and robustness.

In the second part, the concept of adaptive estimation is investigated in a framework where multiple redundant measurements are available and the sensor fault tolerance is an issue. In this approach, in order to develop robustness to changing noise statistics and occasionally disrupting measurement signals, an adaptive EKF is derived. This is done in an application to a nonlinear system estimation problem in an attempt to robustly handle uncertainty in the sensor noise statistic that identifies the value of the covariance in the filter's equations. It will be shown how the instrumentation specificity of batch distillation processes, such as the availability of multiple redundant measurements, together with the application of adaptive filtering create the potential for enhanced fault tolerant process estimation in a dual layer solution.

It is interesting to note that neither unscented nor adaptive fault tolerant estimation for chemical distillation processes have been referenced in the literature earlier. Therefore, this investigation validates itself as a reasonable contribution to the designated area.

2. Estimation for batch distillation

Batch distillation is a popular unit operation widely used in fine chemistry, pharmaceutical, cosmetics and biochemical industries to process small amounts of materials with high added value. Batch distillation is a precious technology, capable of separating components of a multi-component mixture in a single operation, with great accuracy. The need for preciseness calls for control methods capable to ensure quality requirements of products and processes. Because of the ever increasing demand for low volume specialty chemicals, batch distillation is becoming an increasingly important technology (M. Barolo and P. D. Cengio, 2001; J. K. Kim and D. P. Ju, 1999).

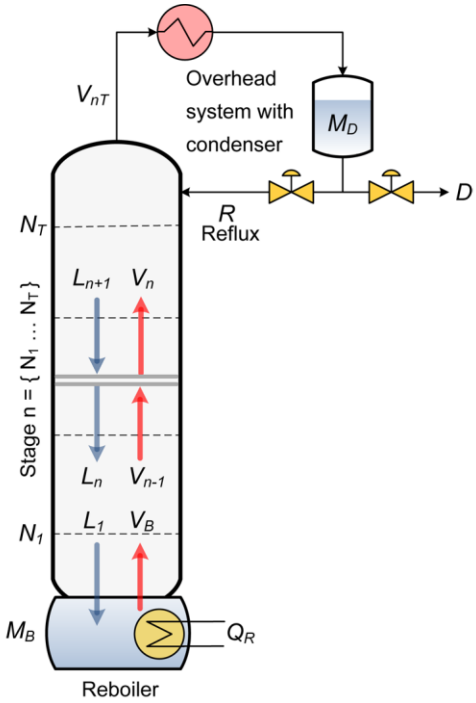


Figure 1. Distillation column schematic.

Operation management and control of batch distillation require knowledge of products compositions during the entire duration of the process. Some of these systems, for instance, are controlled by Model Predictive Controls (MPCs) that necessitate the availability of accurate data obtained from state estimation.
A traditional way of obtaining this knowledge is by means of the use of composition analyzers. Though very accurate, they are expensive instrumentation, and the measurement process is fully manual. This, besides introducing additional delays in the control loop, necessitates direct supervision

from the operational personnel. For the state-of-the-art of batch distillation technologies, see (T. Mejdell and S. Skogestad, 1991; R. M. Oisiovici and S. L. Cruz, 2000) and (C. Venkateswarlu and S. Avantika, 2001), this type of manual interaction is highly undesirable.

Temperature measurements available at individual column trays provide analytically redundant set of measurements. The most popular alternative to composition controllers is the utilization of standard temperature feedback controllers. Temperature measurements, however, are not accurate indicators of composition variation (T. Mejdell and S. Skogestad, 1991). Another alternative is the use of state estimators, which are based on secondary temperature measurements as observations.

Batch distillation, as well as continuous distillation, is a complex, high-order nonlinear process, whose dynamics varies over time. There have been deterministic (E. Quintero-Marmol at all, 1991) and probabilistic modeling approaches (R. M. Oisiovici and S. L. Cruz, 2000; R. M. Oisiovici and S. L. Cruz, 2001) used to state estimator design for batch columns, which are typically based on the piece-wise linearization of the nonlinear dynamics. The probabilistic approaches are ultimately based on the EKF solution of the filtering problem. The EKF applications, however, even if they are successful, report numerical problems and computation issues regarding the calculation of the Jacobian, other difficulties in filter start-up and tuning of the covariance matrices (C. Venkateswarlu and S. Avantika, 2001).

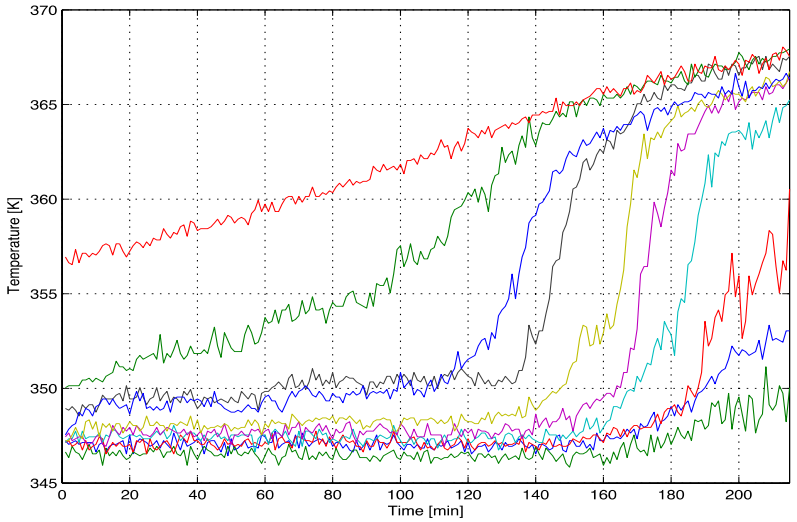


Figure 2. Temperature *measurements* available at individual column trays provide analytically redundant set of measurements.

2.1 Formulation of the filtering problem

The batch distillation process and its sensory system, in their nominal representations, are modeled as a nonlinear dynamical system described by ordinary differential equations

$$\begin{aligned} \dot{x}(t) &= \phi(x(t),\, u(t)), \\ \varphi(t) &= h(x(t)). \end{aligned} \tag{1)-(2)}$$

For a special case this can be written in state space form by means of the set of state and observation equations affected by additive noise

$$\begin{aligned} \dot{x} &= \sum_{i=1}^{m} g_i(x)\, u_i + w_i, \\ \varphi_i &= h_i(x) + v_i, \qquad 1 \le i \le p, \end{aligned} \tag{3)-(4)}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $\varphi \in \mathbb{R}^p$ denote the state, the input and the output of the system. The noise $w(t)$ and $v(t)$, which affect the system and measurement equations, respectively, are considered zero-mean Gaussian white noise independent of each other as well as of $x(0)$.

Let, moreover, Q_i and R_i denote the covariance matrices of v_i and w_i for all i, respectively, as

$$Q_i = \mathbf{E}\{v_k v_k^T\}, \quad R_i = \mathbf{E}\{w_k w_k^T\}. \tag{5}$$

Our objective is to construct a statistical filter in the following sections, which is robust to the nuissance variations and changes of $v_i(t)$, in an attempt to give an estimate $\hat{x}$ of the state x, which is a GRV of (3), assuming the process is observable by all, or by some combination of the available measurements.

2.2 Column models and model validation

High definition nonlinear models for batch distillation processes (such as the one depicted by Fig. 1) have been available in the literature from many sources for years, see e.g., (I. M. Mujtaba, 2004). They typically consist of a large number of coupled nonlinear differential equations describing the variation of compositions along individual column trays based on the tray's thermodynamics and hydraulics, and models of the condenser and the reboiler behaviour by taking energy and material balances, liquid hold-ups, vapor and liquid flow-rates etc. into consideration. High definition models are indispensable tools for the analysis, however, they are overly complicated for controller and filter design because of the implied computational complexity of the implementation.

Therefore, filter synthesis necessitates a simpler representation, which still captures the essential parts of the dynamics of (3).
Though, a high definition model, similar to the one published in (I. M. Mujtaba, 2004) was used for validation purposes (*cf.* model validation later), a simple representation of (3) for two components mixtures of batch distillation processes was developed for filter synthesis. In the following part the simplified model is given, the properties of the more complex high definition model are not detailed. The interested reader is directed to (I. M. Mujtaba, 2004) for more information.
This model and the basic assumptions of the modeling process can be summarized as follows. The state variables are the liquid compositions of the individual stages (stills and trays), i.e., $x_i = c^{\ell}{}_i$. Since $\sum_i^{n_c} c^{\ell}{}_i = 1$, for a system with n_c components, it is sufficient to consider $n_c - 1$ state variables at each stage only, because the composition of the n_c^{th} component can be obtained by simple subtraction. The process is driven by the reflux *R* as input, which is given as

$$u = R = \left[\frac{L}{V}\right],$$

where *L* and *V* are the liquid and the vapor flow rates, respectively. For simplicity, both *L* and *V* are assumed constant. Then, the simplified state space representation of (3) with $x \in \mathrm{R}^n$ and $n_c = 2$ subject to noise is written as

$$\begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_j \\ \vdots \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} \frac{V}{M_1}(y_2 - x_1) \\ \vdots \\ \frac{V}{M_j}(y_{j+1} - y_j) + \frac{RV}{M_j}(x_{j-1} - x_j) \\ \vdots \\ \frac{RV}{M_n}(x_{n-1} - x_n) + \frac{V}{M_n}(y_n - x_n) \end{bmatrix} + \begin{bmatrix} w_1 \\ \vdots \\ w_j \\ \vdots \\ w_n \end{bmatrix}, \tag{6}$$

where the section index $j = 1$ stands for the *condenser*, $j = n$ for the *reboiler* and ($j = 2 \ldots n-1$), for the *column trays* dynamics. The mole fraction (composition) of component c_i in the liquid and vapor of stage j is denoted by x_j and y_j, respectively. The system state is not directly observable. The observed process variables are the tray temperatures ($\varphi_i = T_i$) and the

measurement system, similarly to (7), is modeled as the function $h(x)$ of the state vector corrupted by measurement noise, i.e.,

$$T = h(x) + v. \tag{7}$$

The nonlinear relationship $h(x)$ between temperature and composition can be obtained from the vapor-liquid equilibrium (VLE) equations. Since the system pressure and the a priori composition estimates are known, this can be given by solving the bubble-point temperature problem for each stage in the following way. The mole fractions x_j, y_j of component c_i in stage j, in equilibrium, can be represented by the thermodynamic model described by the modified Rault's law

$$y_j = \gamma_j x_j \left(\frac{p_j^V}{p} \right), \tag{8}$$

where γ_j are the activity coefficients calculated from the Wilson equation, see (N. P. Chopey (Ed.), 2004), and p is the column pressure, which is assumed constant in our case (i.e., atmospheric distillation).

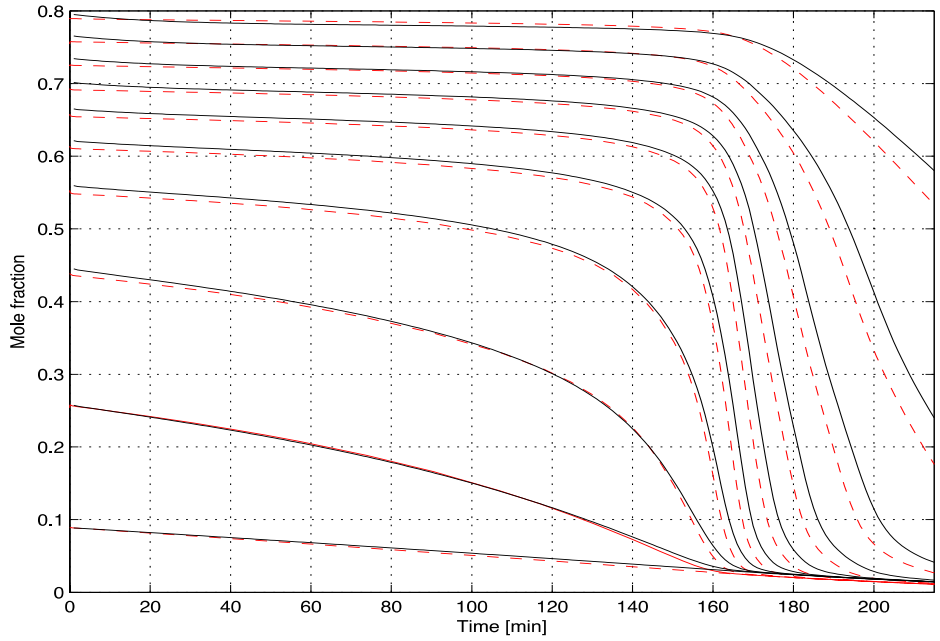


Figure 3. Results of process simulation based on a high definition process model (black-continuous line) and the simplistic model (red-dashed lines), respectively.

The tray vapor pressure p_j^V is defined by the Antoine vapor-pressure correlation equation $T_j = \psi p_j^v$, see (N. P. Chopey (Ed.), 2004). Realize that the thermodynamics of (8), together with the VLE condition $\sum_{i=1}^{n_c} y_i = 1$, give the equilibrium temperatures in (7) as the function $h(x)$ for every tray.

The accuracy of the process model (6) was validated against the high definition model. The comparative result of process simulations, showing the variation of fluid compositions along eight column trays, can be seen in Fig. 3. The results validates the simplistic formulation (6) with respect to the more complex high definition description.

2.3 Preliminaries to filter synthesis

The process, depicted by Fig. 1, consists of 10 separate distillation stages, i.e., 8 column trays plus the reboiler and a total condenser as usual. Trays are identified by the parameter NT. The control-input is the reflux of liquid flow rate which acts on the plant. The sensor outputs used for control and estimation purposes are the temperature data of the column trays. There is one temperature measurement available for each tray.

The instrumentation includes temperature sensors of varying quality. The reliability of individual sensors is frequently inadequate to satisfy the reliability requirements of this type of industrial processes. Due to the coupled nature of the nonlinear dynamics, however, these sensors provide a redundant set of measurements upon which the estimation and control of the process can be based. It is known from the literature that the n-component distillation process is considered observable if at least n temperature measurements along column trays are available for filter synthesis. For the proof of this condition, see (C. C. Yu and W. L. Luyben, 1987). In cases when more than n measurements may be used, the performance of the estimator may improve. In this paper the case of using the minimum necessary number of measurements is taken into consideration, i.e., two of the 10 available temperature measurements are used for filter synthesis. Due to the availability of redundant sensors, the estimation of a particular state of the global state vector of the system can be accomplished relying on various different measurements.

In the following part the core filtering algorithm is considered as the extension of the standard linear (Kalman) filter, which can be characterized briefly as follows. Let the estimate and its covariance provided by the filter be represented with $\hat{x}$ and P. Since the system is not linear, the Riccati matrices that attempt to approximate the *a priori* and the *a posteriori* covariances are defined, respectively, as

$$P_{k|k-1} \approx \mathbf{E}\{e_{k|k-1} e_{k|k-1}^T\} \quad \text{and} \quad P_{k|k} \approx \mathbf{E}\{e_{k|k} e_{k|k}^T\}. \tag{9}$$

The filter is initialized with $x_{o|o} = x_o$ and $P_{o|o} = P_o$, and then operated recursively performing a single cycle each time a new set of measurements becomes available. Each iteration propagates the estimate from the time the last measurement was obtained to the current time. The propagation process consists of two stages: update and prediction as usual. The update equations are responsible for the feedback, i.e., for incorporating a new measurement set into the *a priori* estimate to obtain an improved *a posteriori* estimate. The *a posteriori* state estimate $\hat{x}_{k|k}$ is computed as a linear combination of an *a priori* estimate $\hat{x}_{k|k-1}$ and a weighted difference between an actual measurement φ_k and a measurement prediction:

$$\begin{aligned} \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k\left[\varphi_k - \bar{h}(\hat{x}_{k|k-1})\right], \\ K_k &= P_{k|k-1}\bar{H}_k^T\left(\bar{H}_k P_{k|k-1}\bar{H}_k^T + R_k\right)^{-1}, \end{aligned} \tag{10)-(11)}$$

where $\bar{H}_k$ is the Jacobian matrix of partial derivatives of $h(x)$ with respect to x. The covariance matrix is updated by

$$P_{k|k} = (I - K_k\bar{H}_k)P_{k|k-1.} \tag{11}$$

The prediction equations are responsible for projecting the current state and error covariance estimates forward to obtain *a priori* estimates for the next time step. The state and covariance matrix in the next sampling instant are estimated by

$$\begin{aligned} \hat{x}_{k+1|k} &= \bar{f}(\hat{x}_{k|k}, u_k), \\ P_{k|k+1} &= \bar{F}_k P_{k|k}\bar{F}_k^T + Q_k, \end{aligned} \tag{12}$$

where $\bar{F}_k$ is the Jacobian matrix of partial derivatives of $\bar{f}(x)$ *w.r.t.* x.

3. The unscented filter

To develop the idea further consider the unscented solution, which is based on the intuitive expectation that with a fixed number of random parameters it should be easier to approximate a Gaussian random distribution than it is to approximate an arbitrary nonlinear function (S. Julier at all, 2000). Therefore, the state distribution of (7) is approximated by a GRV as usual, but is now represented with a minimal set of sample points (also called sigma points), which are selected around the mean $\bar{x}$ and covariance P and captured in the vectors

$$\mathrm{X}_i = \bar{x} + \left[(n+\lambda)P\right]_i^{1/2}, \qquad i = 1, \ \ldots \ n,$$
$$\mathrm{X}_i = \bar{x} - \left[(n+\lambda)P\right]_{i-n}^{1/2}, \qquad i = n+1, \ \ldots \ 2n,$$

with $\mathrm{X}_i \in \mathbb{R}^{(2n+1)}$ and scaling parametrization λ. For details on the choice of λ, see e.g., (E. A. Wan and R. van der Merwe, 2001). The sigma matrix at time k and specific scaling λ can be calculated by

$$\mathrm{X}_k = \left[\bar{x}_k, \ \bar{x}_k + \gamma P_k^{1/2}, \ \bar{x}_k - \gamma P_k^{1/2}\right],$$

with $\gamma = (n-\lambda)^{1/2}$. As the sigma states are propagated through the nonlinear dynamics normally, i.e., $f = f(\mathrm{C})$, the mean and covariance for the observations $\varphi(t)$ are approximated using a weighted sample mean and covariance of the posterior sigma points. This principle, with straightforward modification of (12) and (9), leads to the time and measurement update equations of the unscented filter, for more details see (E. A. Wan and R. van der Merwe, 2001).

4. The fault tolerant adaptive filter

Fault tolerant filtering is approximated in a dual layer approach. From the one hand, the filter is attempted to make robust against measurement drop-outs and excessive variations in the input signals by means of recursive estimation of the noise covariance. To this effect, real-time correction of the filter gain, based on the actual characteristics of the noise, suppresses the effect of sporadic variations in the quality of the measurements as well as compensates for slowly worsening sensory conditions and incipient sensor faults. From the other hand, fault monitoring is continuously applied to the filter's innovation in an attempt to initiate a filter reconfiguration action when the adaptation mechanism alone is not able to overcome the failure situation.

4.1 Estimation of the noise covariance

The idea of using adaptively tuned EKF to increase estimation accuracy and robustness against system modeling errors and variations in the driving noise by means of real-time estimation of the covariances R and Q is not new. There have been many different adaptation schemes developed in the past years. The idea in a classical approach was presented in (R. Mehra, 1970) and (R. Mehra, 1971), which is known as the Maximum Likelihood Estimation (MLE) that have been used in various forms in many applications. Another variants are the multiple model and the covariance matching method that try to make the elements of the online estimates of the innovation (or residual) covariances

consistent with their reference values as calculated by the state estimator. The idea is that the innovations (residual) covariance should correspond to its reference form by modifying Q and R until they matches. For further details of the adaptation techniques, see (P. S. Maybeck, 1982).

The basic idea of the adopted approach is to adaptively estimate the covariance of the measurement processes based on the filter's residual and using the innovation for sensor fault detection. Then, based on the decision of the existence of the fault, filter reconfiguration is initiated in an attempt to restore the normal operational conditions of the filter.

The method relies on the evaluation of the *residual* sequence $r_k = j_k - H_k \hat{x}_{k|k}$ that can be derived from (9) in a straightforward way. One difficulty of the residual-based approach is that positive definitiveness of R is to be additionally ensured. One method to achieve it is through the application of the side condition presented in (J. Wang, 2000). Application of a similar approach for the enhancement of GPS-signal based navigation was reported in (R. Campana and L. Marradi, 2000) and (F. D. Busse at all, 2002). Using the above considerations the estimated value of the covariance R at time k is given as

$$\hat{R}_k = \hat{C}_{\bar{v}} + H_k \hat{P}_k H_k^T, \tag{13}$$

where the estimated variance of the residual covariance in a moving window of size m is calculated as

$$\hat{C}_{\bar{v}} = \frac{1}{m} \sum_{i=1}^{m} r_{k-i} r_{k-i}^T. \tag{14}$$

Then, the estimated value of R is used in the measurement update (9) at time $k+1$ by adaptively recalculating the filter gain K according to the varying noise conditions.

4.2 Fault detection and filter reconfiguration

The idea of fault tolerance is based on the detection of the abnormal signal behavior and the proper selection of the measurement signals. In case a particular measurement gets faulty and the fault is detected the faulty sensor is replaced with a healthy one in real-time in an attempt to restore the unaffected operational mode of the filter, instantaneously.

In contrast to filter adaptation, which is a residual-based method, as can be seen in (13)-(14), the idea of fault detection relies on the filter's innovation as follows. The innovation $\rho_k = \varphi_k - H_k \hat{x}_{k|k-1}$ is known to be a white sequence if the linear Gaussian condition for the measurement noise is reasonably true. Detection of the abnormal (faulty) behavior of the measurements, therefore, can be based on standard whiteness tests carried out on the innovation $\rho(t)$. In this paper the

Normalized Innovation Squared (NIS) statistic is applied for this purpose that was defined in (M. Scalzo at all, 2009) in the form

$$\Psi = \rho_k^T S_k^{-1} \rho_k, \tag{15}$$

where the innovations covariance matrix is calculated as

$$S_k = H_k P_{k|k-1} K_k^T + R. \tag{16}$$

Since the statistic of (15) generated by the white nose driven filter's innovation follows a c^2 distribution the test if Ψ falls outside a confidence region for a c^2 random variable provides an indication of a fault in the sensory system when its value is compared to a threshold. In order to minimize false alarm rate of the detector, on account of the overly sensitivity to outliers and non-consistent measurement data, the test is evaluated along a sliding window over the values of (15) that results in the test statistic calculated in the particular time k as

$$J_k = \sum_{i=k-\ell}^{k} \frac{\Psi_i}{\omega} \quad \text{with} \quad \ell = \omega - 1, \tag{17}$$

where ω is the length of the moving evaluation window.

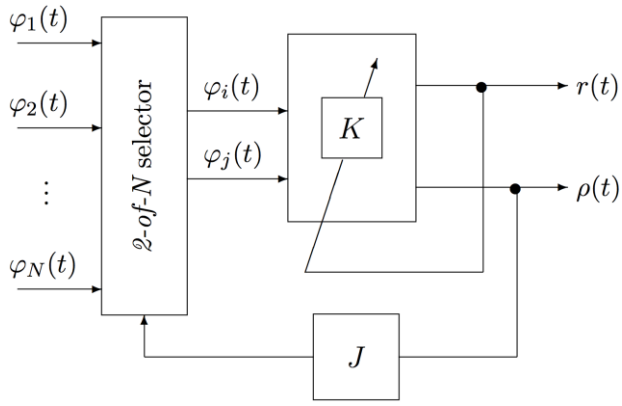


Figure 4. Dual layer fault tolerance scheme based on (i) adaptive tuning of the filter gain and (ii) sensor fault detection and filter reconfiguration.

Then, the windowed variable J_k is used as a means for triggering filter reconfiguration, i.e., for deciding whether the current filter applied during time k should be further used for time $k+1$, or to choose an alternative filter that can be obtained by the replacement of the erroneous measurement.

During measurement switching, the information, i.e., the current state and covariance estimates contained by the filter in the last time step is forwarded to the new filtering setup. For measurement replacement a 2-of-the-N selection algorithm is applied, which, quite intelligibly, prefers opting the least neighboring healthy signal from the set of N redundant available measurements.

5. Experimental results

5.1 Pilot plant

The laboratory equipment, used for engineering education, G.U.N.T. CE-600 distillation column unit served for the purposes of the pilot. The column, which is made of steel, has 8 trays and it is of 750 mm height with 55 mm inner diameter. The column has a reboiler (10 litres), with a maximum applicable heat duty of 2.400 W, and a total condenser. The heat duty is supplied by an electric blanket. A power divisor installed permits to work at different heat duties. The actually applied heat duty was calculated for the vaporization of water which is of a high heat of vaporization, in comparison with other components.

Temperature measurements are available at ten different locations of the column, see Fig. 2, eight at the trays and two at the condenser, which are provided by temperature transducers Pt-100. The column is also equipped with a pressure transducer to measure the relative pressure between the top and the bottom of the column. The reflux ratio is controlled by means of the actuation of an electromagnetic valve. The plant has two additional valves, one for controlling the water flow-rate at the condenser and the other for controlling the absolute pressure. A flow meter gives access to the water flow-rate at the condenser. The plant instrumentation is connected with a computer and controlled with a process supervision and data acquisition software.

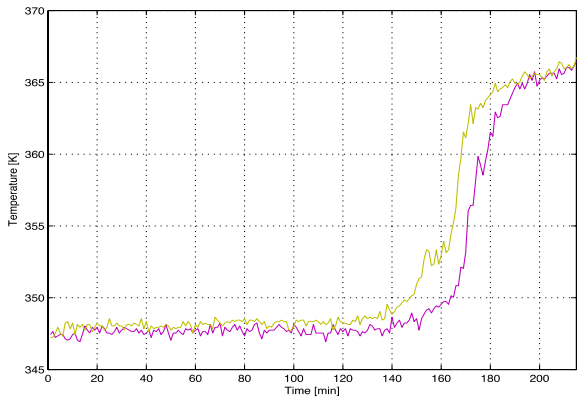


Figure 5. Temperature measurements of trays #6 and #7 used for filtering.

5.1 Experimental scenario

Distillation experiments with ethanol and water mixtures have been performed. Heat duty of 360 W was applied to the distillation process that worked at a constant reflux ratio of 2.33. The initial volume of the mixture was 7 litres. The initial composition of the mixture was 15 to 85 mole fraction ethanol and water, respectively.

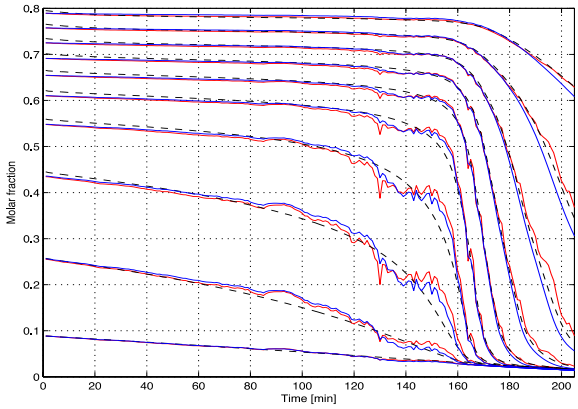


Figure 6. Comparison of filter performance: EKF estimations (blue lines) and UF estimations (red lines) compare favorably with the simulated nominal process behavior (black dashed lines) along column trays.

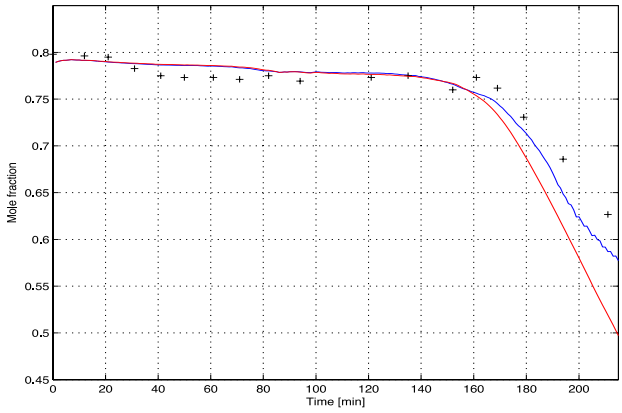


Figure 7. Composition variation of the end product: EKF estimate (red line), UF estimate (blue line) along with composition measurements achieved by the hand-held densitometer (marker line).

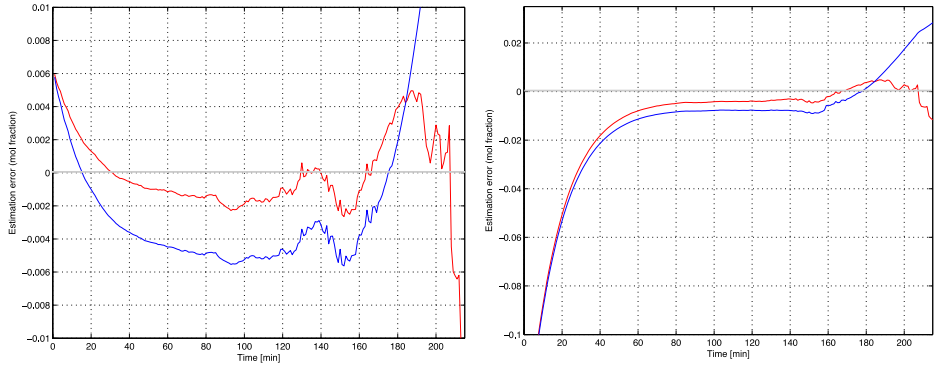
The column was charged and heated up under total reflux until steady-state conditions were achieved. Then, the reflux was turned from maximum to the steady-state value 2.33 and the distillate product was withdrawn at the top of the column.

For the verification of the distillate liquid phase compositions, samples of 5 ml volumes were taken out of the column, and analyzed by means of the hand-held portable density meter (Densito 30 PX) by ± 0.001 g/cm^3 nominal accuracy. Relying on (C. C. Yu and W. L. Luyben, 1987), which concluded that the composition variables of the distillation, from the measurement system (7), are observable if the number of temperature measurements is at least equal to $n_c - 1$, the two component observation matrix was composed of the measurement vectors T_6 and T_7, see Fig. 5. This was validated by our experiments that using three or more measurements as filter input gives no significantly better result than that of the filter relying on a two component observation matrix.

5.2 Simulation results

Process simulations were performed in Matlab. In the first place EKF and UF algorithms were implemented and the filters were applied to the real process data acquired from the pilot plant. The simulation results have been compared with the real experimental data.

Based on the raw temperature measurements (see Fig. 5), the performance of the individual filters *w.r.t.* simulated process behavior obtained from the high definition model is shown, tray-by-tray, in Fig. 6.



Filter 8. Filter convergence in presence of a) moderate and b) large initial condition mismatch. EKF estimation error (blue line), UF *estimation* error (red line).

Estimation accuracy is characterized by the mean square error (MSE) statistics of the estimation for EKF and UF, respectively, see Table I. For the characterization of the most important performance measure, i.e., estimation accuracy of end product composition, see Fig. 7, which gives the estimated mole fraction values *w.r.t.* the accurate densitometer measurements in the condenser.

In an investigation of sensitivity on initial conditions, two experiments are presented here: Fig. 8 and Fig. 9 show the convergence behaviour in the presence of moderate (5%) and severe (25%) initial conditions mismatch concerning the deviation from the nominal state value.

Table I. MSE indices for EKF and UF estimations

	MSE	
	EKF	**UF**
Condenser	0.0079	0.0026
Plate 1	0.0139	0.0120
Plate 2	0.0104	0.0114
Plate 3	0.0076	0.0074
Plate 4	0.1050	0.0080
Plate 5	0.0089	0.0116
Plate 6	0.0131	0.0184
Plate 7	0.0135	0.0177
Plate 8	0.0098	0.0081
Reboiler	0.0024	0.0018

In the second place the performance of the adaptive fault tolerant filter was contrasted with the standard EKF solution. The filter behavior in the incidents of abnormal input signals was investigated in various experimental scenarios. The performance of three different filtering setup namely, (i) the normal EKF, (ii) the adaptive EKF and (iii) the fault tolerant adaptive EKF in response to two typical sensor flaw scenarios are illustrated by the simulation results in a comparative way. In Scenario 1 the effect of nuisance input signal variations and incipient sensor faults is studied. To this end, moderate increase in the measurement noise covariance was applied to one of the filter's input at time 120 in the simulations. In Scenario 2 a continuous bias, as another typical abnormality of the sensory system, was formed by adding a constant term to the measurement, which entered the system at time 150. Simulation results comparing the performance characteristics of the three filtering solutions are shown in Fig. 10 and 11.
It can be seen how the particular filtering approaches affect estimation accuracy in the presence of the two fault scenarios. While the standard EKF estimate fails to follow the real state almost inmediately subsequent to the fault injection time the adaptive filter is capable to compensate the fault effect very nicely and the adaptive filter extended with the fault tolerant mechanism become nearly insensitive to the introduced abnormalities.

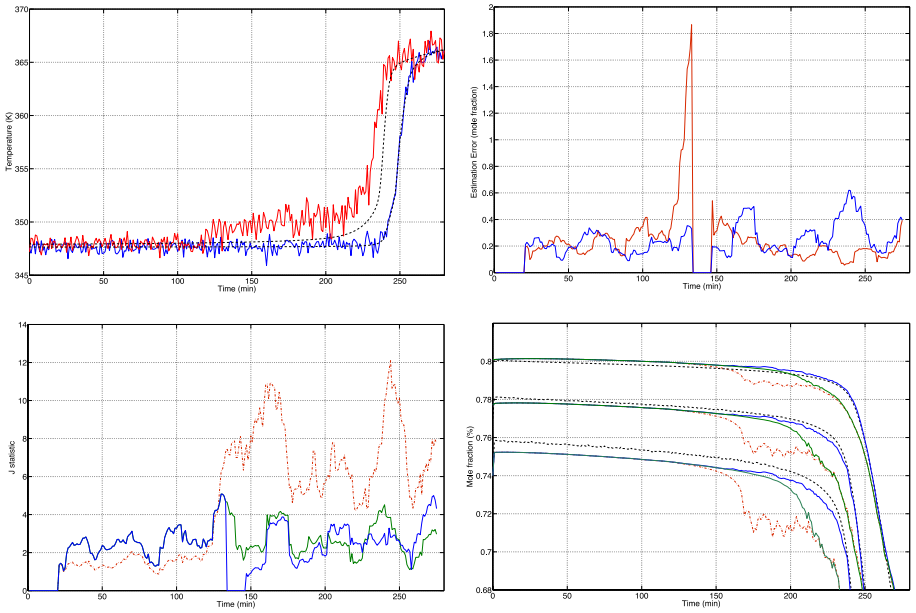


Figure 10. From top left to bottom right: a) Filter temperature measurements subject to nuisance variation of the input noise cropping up at time 120. b) Residual errors of the EKF. c) J statistic of individual filters when triggered by the violated threshold of the residual error. d) Accuracy of estimation. (red: EKF, green: adaptive EKF, blue: adaptive EKF with filter reconfiguration, dotted black: real process behavior).

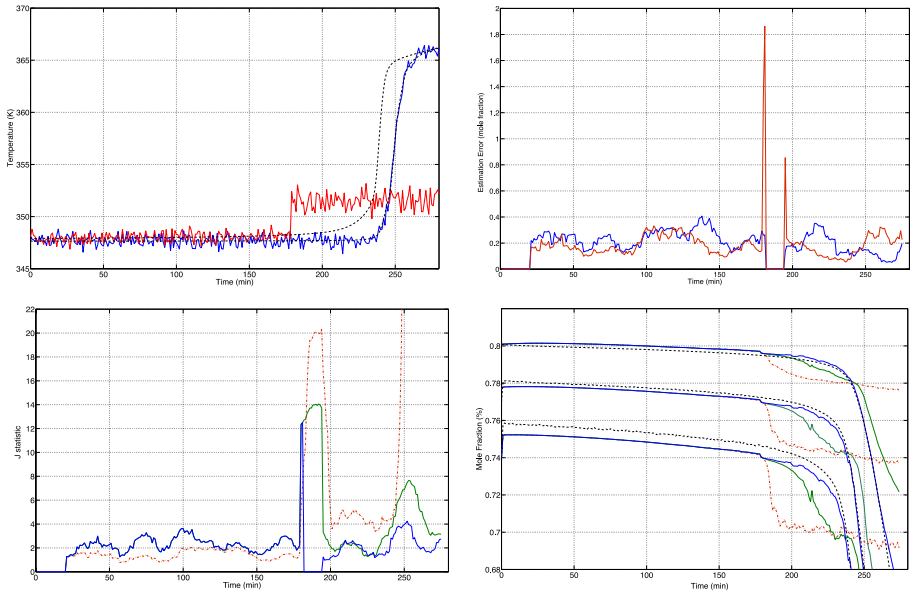


Figure 11. From top left to bottom right: a) Temperature measurements of the filter, which is subject to a measurement bias appearing at time 150. b) Residual error of the EKF. c) J statistic of individual filters when triggered by the violated threshold of the residual error. d) Accuracy of estimation. (red: EKF, green: adaptive EKF, blue: adaptive EKF with filter reconfiguration, dotted black: real process behavior).

5. Conclusion

In this paper the implementation and performance of two main classes of advanced filtering techniques in comparison with the classical extended Kalman filter has been studied. Namely, the unscented and the adaptive filtering solutions were compared with each other from view of robustness and the achievable estimation accuracy. The filter algorithms have been applied to state estimation of a pilot plant of batch distillation process based on the column time series data. The experiments showed convincingly that the unscented algorithm is capable to provide at least the same but typically better estimation accuracy than that of the traditional EKF algorithm. Though computational burden is in par with the EKF, implementation of the unscented filter is simpler, not requiring the calculation of the derivatives. The increased robustness of the UF against changing and varying intensity noise and unmatched initial conditions, however, could not be substantiated in comparison to the standard EKF. Regarding the application of the adaptive filter a dual layer approach to the robust fault tolerant estimation of nonlinear processes was presented. The implementation and performance of the standard EKF solution in comparison with the adaptive and the fault tolerant adaptive EKF schemes, which make up the two operational levels of the filter, were comparatively studied. The results confirmed convincingly that the adaptive EKF is capable to increase robustness of the filter against nuisance variation of the measurement signals and, by the extension of the filtering scheme with a fault detection and reconfiguration mechanism, to enhance sensor fault tolerance and estimation accuracy, significantly.

References

T. Bailey, J. Nieto, J. Guivant, M. Stevens, and E. Nebot. Consistency of the EKF-SLAM algorithm. In: *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Syst.*, pages 3562-3568, 2006, Beijing, China.

M. Barolo and P. D. Cengio. Closed-loop optimal operation of batch distillation columns. *Computers and Chemical Engineering*, 25:561-569, 2001.

F. D. Busse, J. P. How, and J. Simpson. Demonstration of adaptive extended Kalman filter for low earth orbit formation estimation using CDGPS. In: *Proc. of Institute of Navigation GPS Meeting*, 2002, Portland, OR.

R. Campana and L. Marradi. GPS-based space navigation: Comparison of kalman filtering schemes. In: *Proc. of Institute of Navigation GPS Meeting*, 2000, Salt Lake City, UT.

N. P. Chopey (Ed.). *Handbook of Chemical Engineering Calculations* (3rd Edition). McGraw-Hill, 2004, ISBN 978-0-07-136262-7.

A. Edelmayer, M. Miranda, and V. Nebehaj. A cooperative federated filtering approach for enhanced position estimation and sensor fault tol- erance in ad-hoc vehicle networks. *IET Intell. Transp. Syst.*, **4**(1):82-92, 2010.

U. Frese. A discussion of simultaneous localization and mapping. *J. Autonomous Robots*, 20:25–42, 2006.

A. Gelb, J. F. Kasper, R. A. Nash, C. F. Price and A. A. Sutherland. *Applied optimal estimation*. MIT Press, MA, 1974.

S. Huang and G. Dissanayake. Convergence and consistency analysis for extended Kalman filter based SLAM. *IEEE Trans. on Robotics*, **23**(5):1552-3098, 2007.

S. Julier, J. Uhlmann, and H. F. Durrant-Whyte. A new method for the nonlinear transformation of means and covariances in filters and estima- tors. *IEEE Trans. Aut. Cont*., **45**(3):477-482, 2000.

S. J. Julier and J. K. Uhlmann. A counter example for the theory of simultaneous localization and map building. In: *Proc. of the IEEE Conf. on Robotics and Automation,* pages 4238–4243, 2001.

S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte. A new approach for filtering nonlinear systems. In: *Proc. IEEE American Control Confer- ence*, ACC'95, pages 1628–1632, 1995.

J. K. Kim and D. P. Ju. Shortcut procedure for multicomponent batch distillation with distillate reciever. *Industrial & Engineering Chemistry Research,* 38:1024-1031, 1999.

P. S. Maybeck. *Stochastic Models, Estimation, and Control*, Volume 2. Academic Press, 1982, New York.

R. Mehra. On the identification of variance and adaptive Kalman filtering. *IEEE Trans. Aut. Cont*., **15**(2):175–184, 1970.

R. Mehra. On-line identification of linear dynamic systems with applications to Kalman filtering. *IEEE Trans. Aut. Cont*., **16**(1):12–21, 1971.

T. Mejdell and S. Skogestad. Estimation of distillation compositions from multiple temperature measurements using partial-least-squares regression. *Industrial & Engineering Chemistry Research*, **30**(12):2543-2555, 1991.

C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review*, 20:801-836, 1978.

I. M. Mujtaba. Batch Distillation: Design and Operation. *Series on Chemical Engineering*, Vol. 3, Imperial College Press, 2004, London.

M. Nørgaard, N. K. Poulsen, and O. Ravn. New developments in state estimation for nonlinear systems. *Automatica*, 36:1627–1638, 2000.

R. M. Oisiovici and S. L. Cruz. State estimation of batch distillation columns using and extended Kalman filter. *Chemical Engineering Science*, 55:4667-4680, 2000.

R. M. Oisiovici and S. L. Cruz. Inferential control of high-purity multicomponent batch distillation columns using an extended Kalman filter. *Industrial & Engineering Chemistry Research*, 40:2628–2639, 2001.

E. Quintero-Marmol, W. L. Luyben, and C. Georgakis. Application of an extended luenberger observer to the control of multicomponent batch distillation. *Industrial & Engineering Chemistry Research*, 30:1870–1880, 1991

A. Romanenko and J. A. A. M. Castro. The unscented filter as an alternative to the ekf for nonlinear state estimation: a simulation case study. *Computers and Chemical Engineering*, 28:347-355, 2004.

M. Scalzo, G. Horvath, E. Jones, A. Bubalo, M. Alford, R. Niu, and P. K. Varshney. Adaptive filtering for single target tracking. In: *Proc. of the SPIE: Defense & Security Symposium*, Volume 4336, 2009. Orlando, FL.

H. W. Sorenson. *Kalman Filtering: Theory and Application*. IEEE Press, NY, 1985.

C. Venkateswarlu and S. Avantika. Optimal state estimation of multicomponent batch distillation. *Chemical Engineering Science*, 56:5771–5786, 2001.

E. A. Wan and R. van der Merwe. *The Unscented Kalman Filter*. In: Kalman Filtering and Neural Networks, (Ed.: S. Haykin), John Wiley & Sons Inc., ISBN: 0-471-36998-5, 2001, New York.

J. Wang. Stochastic modeling for real-time kinematic GPS/GLONASS positioning. *Navigation*, 46(4):297–305, 2000.

C. C. Yu and W. L. Luyben. Control of multicomponent distillation columns using rigorous composition estimators. *Int. Chem. Eng. Symp. Series*, 104:A29-A69, 1987.

APPLICATION OF ROTOR DYNAMICS TO ASSESSMENT OF LIQUID LUBRICANT'S PROPERTIES

Jozef RÉDL[1], Veronika VÁLIKOVÁ[1], Juraj MAGA[2]

[1]Department of Machine Design, Faculty of Engineering,
Slovak University of Agriculture in Nitra
[2]Department of Machines and Production Systems, Faculty of Engineering,
Slovak University of Agriculture in Nitra

Abstract

In this contribution we present the method of assessment of biolubricants' damping and stiffness properties used in agricultural machines and devices. Presented methodology is universally usable for all types of the liquid lubricants, not only the biolubricants. This method is based on the mathematical model of journal's motion in the sliding bearing. We assess here the stiffness and dumping properties of used biolubricant following its dynamic viscosity and Sommerfeld number. Stability of the journal's motion trajectory then defines properties of used lubricant, in this case MOL Farm UTTO Synt.

Key words

Biolubricants, stability of rotors, mathematical modeling, Sommerfeld number

1. Introduction

Mathematical description of rotor dynamics assembled on slide bearings was published by many authors. Between the most known authors belong the [3] who mathematically described damped and undamped Laval rotor on rigid bearings. There is also defined the Laval rotor with inside and outside damping. Analytical and numerical approach of the lubricated plane slider bearing was investigated by [5]. Important work of rotor dynamics was published by [2]. The effect of bearing cage run-out on the nonlinear dynamics of a rotating shaft was analyzed by [6].

2. Methods

Initial conditions
For simulation we have chosen a radial sliding bearing which parameters are described in the table 1. Slide bearing will be loaded by statically load $F = 1000N$. The nominal diameter of journal is $d = 30.10^{-3}m$ and its

material is constructional steel with modulus of elasticity $E = 2{,}1.10^5 MPa$. Material of the bush is CuSn alloy with modulus of elasticity $E_P = 1{,}1.10^5 MPa$ and material of the bearing body is constructional steel with modulus of elasticity $E_L = 2{,}1.10^5 MPa$. Rotational speed of the journal is $n = 1430\,\mathrm{min}^{-1}$, the fit of bush and bearing body is H7/r6. To calculate the bearing dimensions and parameters we used the methodology described by [1]. The average relative bearing clearance is $\overline{\Psi}_e = 3{,}003.10^{-3}$, bearing acting pressure $P_m = 1{,}111 MPa$. The Sommerfeld number was calculated from equation $S_O = \frac{P_m.\overline{\Psi}_e^2}{\eta_T.\omega}$, where $\eta_T = 8{,}871982.10^{-3}$ is a dynamic viscosity of lubricant at $100^\circ C$ and $S_O = 7{,}526$, $\varepsilon = 0{,}82718$, angular velocity $\omega = 149{,}75 s^{-1}$. Bearing clearance is $C = \frac{\overline{D} - \overline{d}}{2} = 69.9.10^{-6} m$. Each stiffness and damping coefficients of bearing that are dependent on the dynamic viscosity are solved with equations (1) and were published by [8].

$$k_{yy} = \frac{\varepsilon\omega\eta\overline{R}\overline{L}^3}{C^3(1-\varepsilon^2)^2},\ k_{yz} = -\frac{\pi\omega\eta\overline{R}\overline{L}^3(1+2\varepsilon^2)}{4C^3\sqrt{(1-\varepsilon^2)^5}},\ k_{zy} = \frac{\pi\omega\eta\overline{R}\overline{L}^3}{4C^3\sqrt{(1-\varepsilon^2)^3}},$$
$$k_{zz} = \frac{2\omega\eta\overline{R}\overline{L}^3\varepsilon(1+\varepsilon^2)}{C^3(1-\varepsilon^2)^3} \quad (1)$$

$$c_{yy} = \frac{\pi\eta\overline{R}\overline{L}^3}{2C^3\sqrt{(1-\varepsilon^2)^3}},\ c_{yz} = -\frac{2\varepsilon\eta\overline{R}\overline{L}^3}{C^3(1-\varepsilon^2)^2},\ c_{zy} = -\frac{2\varepsilon\eta\overline{R}\overline{L}^3}{C^3(1-\varepsilon^2)^2},$$
$$c_{zz} = \frac{\pi\eta\overline{R}\overline{L}^3(1+2\varepsilon^2)}{2C^3\sqrt{(1-\varepsilon^2)^5}}.$$

Table 1. Stiffness and damping values of bearing

Stiffness) $(N.m^{-1})$	Sommerfeld number	Damping $(N.s.m^{-1})$	Sommerfeld number
	7,526		7,526
k_{yy}	1,3017786.10^7	c_{yy}	92763,6447687
k_{yz}	-5,2095975.10^7	c_{yz}	-173860,2425619
k_{zy}	6945677,90205	c_{zy}	-173860,2425619
k_{zz}	1,3886485.10^8	c_{zz}	695772,6192458

Table 2. Dimensionless stiffness and damping coefficients of bearing

Stiffness	Sommerfeld number	Damping	Sommerfeld number
	7,526		7,526
γ_{yy}	0.9099432	β_{yy}	0.9710058
γ_{yz}	-3.6415086	β_{yz}	-1.8198864
γ_{zy}	0.4855029	β_{zy}	-1.8198864
γ_{zz}	9.7066533	β_{zz}	7.2830173

We solved the dimensionless parameters of stiffness and damping by equations (1) like [4],

$$\gamma_{ij} = \frac{C}{F} k_{ij}; (i, j = y, z),\ \beta_{ij} = \frac{C.\omega}{F} c_{ij}; (i, j = y, z) \quad . \tag{2}$$

Mathematical model

Mathematical model is based on the basic theory of rotor dynamics published by [3] and [2]. Solving differential equations (3) in matrix form gives us the trajectory of journal moving in the bearing.

$$\frac{S_O \,.\, C}{g} . \begin{bmatrix} \ddot{z}_J \\ \ddot{y}_J \end{bmatrix} + \frac{1}{\omega} . \begin{bmatrix} \beta_{zz} & \beta_{zy} \\ \beta_{yz} & \beta_{yy} \end{bmatrix} . \begin{bmatrix} \dot{z}_J \\ \dot{y}_J \end{bmatrix} + \begin{bmatrix} \gamma_{zz} & \gamma_{zy} \\ \gamma_{yz} & \gamma_{yy} \end{bmatrix} = \varepsilon.\omega^2 . \frac{S_O.C}{g} . \begin{bmatrix} \cos\omega t \\ \sin\omega t \end{bmatrix}. \tag{3}$$

Finally we get the limit angular velocity of the journal in the equation (4) below:

$$A_0 = \gamma_{zz}.\gamma_{yy} - \gamma_{yz}.\gamma_{zy},\ A_1 = \beta_{zz}.\gamma_{yy} + \beta_{yy}.\gamma_{zz} - \left(\beta_{zy}.\gamma_{yz} + \beta_{yz}.\gamma_{zy}\right),$$

$$A_2 = \beta_{zz}.\beta_{yy} - \beta_{yz}.\beta_{zy},$$

$$A_3 = \beta_{zz} + \beta_{yy},\ A_4 = \gamma_{zz} + \gamma_{yy},\ \omega_L = \sqrt{\frac{g}{S_0.C} . \frac{A_1.A_2.A_3}{A_1^2 - A_1A_3A_4 + A_0A_3^2}}. \tag{4}$$

3. Results and discussion

To solve differential equation (3) we chose the Runge-Kutta numerical method. We made the simulation with temperature 100°C for used biolubricant. The physical and mechanical properties of used biolubricants for simulating temperature were defined by the value of its Sommerfeld number. The differences in stiffness and damping are obvious when the temperature of the oil was 100°C and the values of Sommerfeld numbers corresponded with this temperature. The viscosity of the oil changes with the rising temperature. The corollary of this is the change of the stiffness and damping and the change of the behaviour of the journal in bearing. The trajectories of the bearing are depicted in figure 1. The maximum and minimum values of journal centre displacements are in table 3.

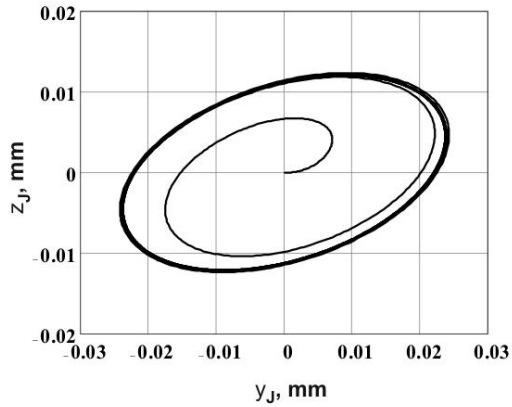


Figure 1. Stabilized trajectory of journal centre

Table 3. Maximum and minimum values of journal centre displacement

Displacement (mm)	Sommerfeld number
	7,526 ($n = 90\frac{\pi}{\omega}$)
$y_{J\max}$	$2{,}432.10^{-3}$
$y_{J\min}$	$-2{,}419.10^{-3}$
$z_{J\max}$	0,012
$z_{J\min}$	-0,012

The used rotational speed during the simulation was $n = 90\frac{\pi}{\omega}$. We can see that after the start, when the rotational speed is stabilized, the trajectory of the centre of the journal is stabilized, too.

4. Conclusion

We addressed the simulation of the simple rotor-bearing system in this contribution. We evaluated the parameters of the sliding terms following the actual methodologies. For lubrication of rotor – bearing system we used biolubricant, where we used Sommerfeld number to define the influence of the lubricant's properties and properties of the sliding couple. We substitute dimensionless parameters into the formed dynamic equations. These parameters describe the properties of lubricated sliding couple at temperature 100°C. Under authority of realized simulations we came to conclusions, that the direct stiffness terms, k_{yy} and k_{zz} become negligible beside the cross - coupled stiffness terms, k_{yz} and k_{zy}, the direct damping terms, c_{yy} and c_{zz} , tend to a non-zero limit, and the cross – coupled damping terms, c_{yz} and c_{zy} , incline to zero. The cross-coupled term k_{zy} must be negative in order to generate instability. The largest degree of instability occurs when $k_{yz} = -k_{zy}$. The system stability rapidly improves in case k_{yz} becomes negative. The greatest values of the deviation were observed in the direction with respect to *y* axes. The motion oriented to *z* axes was relatively consolidated. Presented methodology is suitable for assessment of the biolubricant influence on stability of the journal's motion in bearing. Critical angular velocity was solved form eq.4 and it is $\omega_L = 1070,9 s^{-1}$. We calculated critical angular velocity at which an unwanted resonant frequencies rising to the damage of the bearing and degradation of used lubricant would occur.

Acknowledgment

The study was supported by the VEGA (project number 1/1064/11 The analyse of the loading process of the grinding couple in the condition of tribologic experiment and their influence on the properties of chosen materials) Grant agency of Ministry of Education, Science and Sport of the Slovak Republic

References

FIALA, J. – SVOBODA, P. – ŠŤASTNÝ, K. 1989. Strojnické tabulky 3. Praha: SNTL. 704 s. ISBN 80-03-00151-X.

GENTA, G. 2005. Dynamics of Rotating Systems. New York: Springer Science + Business Media Inc.660 s. ISBN 0-387-20936-0.

GASCH, R. – PFÜTZNER, H. 1975. Rotordynamik – Eine Einführung. Berlin/Heideberg: Springer Verlag. 188 s. ISBN-13: 978-0387070469.

GOMEZ – MANCILLA, J. - NOSOV, V. – SILVA – NAVARRO, G. 2005. Rotor – Bearing System Stability Performance Comparing Hybrid versus Conventional Bearings. In International Journal of Rotating Machinery, roč.2005, 2005, č. 1 , s. 16-22. ISSN: 1023-621X.

MALVANO, R. – VATTA, F. – VIGLIANI, A. 1999. Lubricated Plane Slider Bearing : Analytic and Numerical Approach. In: Meccanica, roč. 34, 1999, č. 4, s. 237–250, ISSN: 1572-9648.

NATARAJ, C. – HARSHA, S.P. 2008. The effect of bearing cage run-out on the nonlinear dynamics of a rotating shaft. In.: Communications in Nonlinear Science and Numerical Simulation, roč. 13, 2008, č. 4, s. 822–838.

RÉDL, J. –KROČKO, V. 2009. Mathematical model of Jeffcott rotor supported on slide bearings assembly. In Annals of Warsaw University of Life Sciences - SGGW Agriculture. No. 53. 2009. p.85-91. ISSN 1898-6730.

WANG, J. K. – KHONSARI, M.M. 2006. A new derivation for journal bearing stiffness and damping coefficients in polar coordinates. In Journal of Sound and Vibration 290, roč.290, 2006, č. 1-2, s. 500–507. ISSN: 0022-460X.