

OPTIMISING A MEASUREMENT SETUP FOR DECISION MAKING

Kimmo Latva-Käyrä and Risto Ritala

Tampere University of Technology /
Measurement and Information Technology,
Tampere, Finland.
kimmo.latva-kayra@tut.fi

Abstract: Measurements are the key to rational decision making. Measurement information generates value, when it is applied in the decision making. An investment cost and maintenance costs are associated with each component of the measurement system. Clearly, there is – under a given set of scenarios – a measurement setup that is optimal in expected (discounted) utility. Contrary to process design, design of measurement and information systems has not been formulated as such an optimization problem, but rather been tackled intuitively. In this presentation we propose a framework for analyzing such an optimization problem. Our framework is based on that the basic mechanism of measurement is reduction of uncertainty about reality. Statistical decision theory serves as the basis for analyzing decision making. In this paper we apply the framework to a problem that is rather simple but of practical importance: how to arrange laboratory quality measurements optimally. In particular, we discuss a case in the paper making industry, in which the product quality is measured with automated quality analyzers and by laboratory measurements.

Keywords: decision support, design, optimization, uncertainty

1. INTRODUCTION

The process industries make use of hundreds of on-line and laboratory measurements to monitor and control the process [1]. The daily decision making about process and product quality by operators and engineers should be supported with information systems so that the best practice of operation can be achieved continuously. Measurements, soft sensors and process simulators form the basis for such decision support by reducing the uncertainty about the present state of the process and about its future evolution.

When making decisions or when combining information from various sources, the uncertainty of information is decisive and must be known. Ideally measurements, soft sensors and simulators should produce the probability distribution of the state of the process and of its predicted evolution. In present information systems such uncertainties

of data are not recorded, and rather often the uncertainty analysis has been altogether neglected.

Information from measurements, soft sensors and simulators generates value through improved decisions [2], because the uncertainty about the state of the process has been reduced. The amount of value generated depends on the goal set by the decision maker, including the decision maker's attitude towards risk. The concept of uncertainty is rather unfamiliar to process operators and engineers and hence risk is dealt with rather implicitly [3].

The optimal measurement system is such that it maximizes the value of information generated, under a given set of scenarios on external effects to the process. Optimal process design is well known [4], but the optimal design of information systems – measurements, actuators, control algorithms and data analysis methods – has emerged only recently [5-7].

This paper is organized as follows. In Section 2 we discuss statistical decision theory briefly as it is the framework within which we analyzed value generated. Furthermore, we discuss the generic problem of measurement setup and state that the relationship to optimal decision making is via specifying, how accurately the process state and/or product quality must be known. Section 3 discusses a practical way of designing the measurement setup and how that relates to optimal estimation in linear case. In Section 4 we formulate a case at papermaking process in which the task is to find an optimal laboratory measurement scheme to support quality management and discuss this case with real-life data. In Section 5 we discuss how our results can be generalized to other cases in process industries.

2. STATISTICAL DECISION THEORY

Decisions are based on available information about the system – measurements, silent information and a priori information. Decision making can be analyzed as an optimization task, a deterministic, stochastic, multigoal or game problem.

The formal statistical decision making problem consists of the following elements: a priori information about the state of the system, models of measurements, model for predicting the consequences of decision alternatives, and the utility of the consequences as the objective. To define these elements, descriptions for the system state (x), set of consequences (c) and set of allowable decisions (actions, a) must be set up. Note that x , c and a , all should be considered vectors, and that they may be past time series (x) or future time series (c, a). Figure 1 presents the decision making task: given the measurement $x^{(obs)}$, and the probabilistic elements, what is the action that yields maximal utility for DM [8,9].

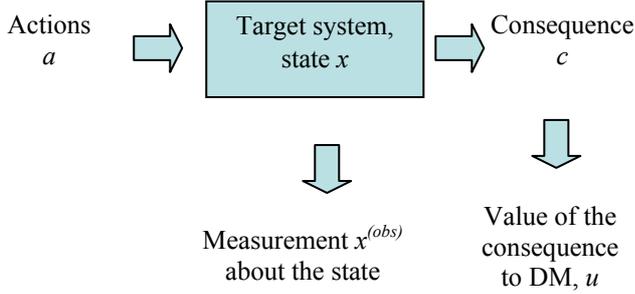


Figure 1. Action - consequence scheme of system.

Decision maker (DM) knows the state of the system, x , only probabilistically through uncertain measurements and possibly through some a priori information. The consequence c of the action a , given that system state is x , is known probabilistically. DM evaluates the system performance in terms of consequences. The value (utility) of the consequence to the decision maker, $u(c)$ uniquely determines the goodness of any action a – including the DMs attitude towards risk, assuming DM is rational.

Formally, the elements of a priori information, measurement models and prediction models are then, respectively, the probability distributions:

$$f_X^{(ap)}(x) \quad (1a)$$

$$f_{X^{(obs)}|x}^{(meas)}(x^{(obs)} | x) \quad (1b)$$

$$f_{C|a,x}^{(pred)}(c | a, x) \quad (1c).$$

Here $x^{(obs)}$ refers to the measured value of x . The probability distribution of consequence c , given that $x^{(obs)}$ has been measured and DM would decide a is then according to Bayes formula [10,11]

$$f_{C|a,x^{(obs)}}^{(pred)}(c | a, x^{(obs)}) = \quad (2)$$

$$N * \int_{domain(X)} f_{C|a,x}^{(pred)}(c | a, x) f_{X^{(obs)}|x}^{(meas)}(x^{(obs)} | x) f_X^{(ap)}(x) d^n x$$

where N is a normalization factor and n is the dimensionality of system state space description.

The optimal decision is the one that maximizes the expected utility, and the corresponding expected utility is the measure of performance [7-9, 12]:

$$a^*(x^{(obs)}) = \arg \max_{a \in A} \int_{domain(C)} u(c) f_{C|a,x^{(obs)}}^{(pred)}(c | a, x^{(obs)}) d^m c \quad (3).$$

$$U^*(x^{(obs)}) = \int_{domain(C)} u(c) f_{C|a^*(x^{(obs)}), x^{(obs)}}^{(pred)}(c | a^*(x^{(obs)}), x^{(obs)}) d^m c$$

Defining the objective of decision making, and in particular the attitude towards risk, is quite often the main challenge when formalizing operational decision making about production in papermaking and in other industrial processes. The most general approach to attitude towards risk is through a utility function. For a rational decision maker the utility function $u(c)$ is guaranteed to exist, but its most general identification method through finding certainty equivalents of “gambling cases” [6] is tedious and often not intuitive for the decision maker.

The setup of measurements affects the optimal expected utility $U^*(x^{(obs)})$ through how accurately the state is known when measurement $x^{(obs)}$ is made. Assuming that there is no measurement bias, this accuracy can be characterized by Σ_{xx} , the covariance matrix of measurement uncertainties. Hence we may write $U^*(x^{(obs)}, \Sigma_{xx})$. In order to achieve an accuracy Σ_{xx} a (life-time, discounted) cost $c(\Sigma_{xx})$ is caused. When designing a measurement setup, we assume that there will be a number of decision making situations, each with their specific action-consequence model and utility. A scenario occurs with frequency p_i , and the corresponding optimal utility, if measurement $x^{(obs)}$ is made, is $U_i^*(x^{(obs)})$. Furthermore, for each scenario, we can assess the a priori probability distribution of observing $x^{(obs)}$ to be $f_i(x^{(obs)})$. Then the design problem reduces into finding the optimal measurement accuracy maximizing the lifetime “profit”:

$$\Sigma_{xx}^* = \arg \max_{\Sigma_{xx}} \left[T * \left(\sum_{i=1}^I p_i \int_{domain(X)} U_i^*(x^{(obs)}, \Sigma_{xx}) f_i(x^{(obs)}) d^n x^{(obs)} \right) - c(\Sigma_{xx}) \right] \quad (4)$$

where T is the life time of the system and we have assumed that utility has been expressed in units comparable to those of costs.

Here we have considered only the direct effect of measurement setup (and accuracy) on value generated. Quite often the prediction model (1c) is identified and updated on the basis of the very same measurements. The better the accuracy of measurements the more accurate are the models and the better optimal utility in (3) can be achieved. Similarly, the a priori information about system state is based on long term statistics of the same measurements: the more accurate the measurements, the more accurate the a priori information and the better the decisions. The accumulating nature of this accuracy leads to complex discounting questions. Hence, we choose to neglect these indirect effects throughout the rest of the paper and concentrate on the direct effect only.

The measurement setup optimization described in (4) is extremely difficult to carry out in practice. We need to specify all decision making situations to arise during the life time of the system, their frequencies, and the utilities and prediction models associated with them. However, we should bear in mind that similar analysis is the basis of optimal process design and should thus be the goal of optimal measurements system design as well.

3. MINIMUM-COST MEASUREMENT SETUP FOR SPECIFIED ACCURACY OF STATE INFORMATION

The analysis above shows that optimizing the measurement setup is equivalent to finding the optimal accuracy, as described with Σ_{xx} . If we cannot solve the formal optimization problem (4), we may first seek to find “sufficient” accuracy for information about the system state through process expertise and then analyze by which measurement setups this accuracy is achieved and which of these setups is the one with lowest costs.

A measurement setup determines which measurements are to be made and when. In accordance with the analysis in Section 2 we shall consider that measurement setup is a measurement policy: the measurements are made at regular intervals independent on which are the measured values. However, we note that this could be further expanded for finding conditions for ad hoc measurements when the decisions are highly sensitive to the system state.

When estimating the system state the correlations between system state variables can be used reduce the number of measurements to be made and hence to reduce measurement costs. The elements in the measurement setup analysis are depicted in Figure 2.

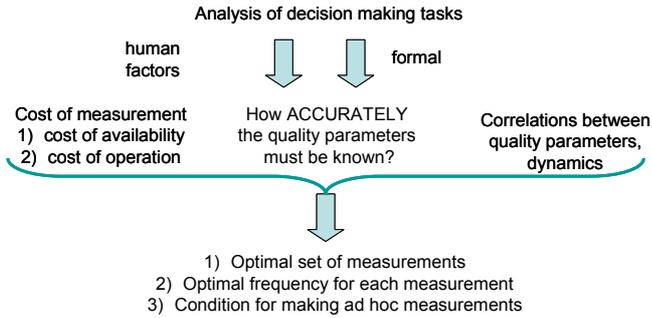


Figure 2. Measurement setup analysis scheme.

Next we shall consider the following case

- a priori information: system state is multivariate normally distributed, $X \sim N_n(\mu, \Sigma)$; no other a priori information
- measurement description: all state variables can be measured, the measurements are unbiased and distributed according to $X^{(obs)} | x \sim N_n(x, C)$; the description of any subset of measurements is obtained by marginalizing the full distribution with respect to the measurements not made.

Let us assume that we have analyzed the tasks that DM will be facing assisted with the measurement information system. With human experts we have concluded that the quality of decision making requires that at all instances the largest allowable uncertainty in state variable x_i is $\sigma_i^{(c)}$:

$$\left(\Sigma_{xx}^{(post)}\right)_{ii} < \left(\sigma_i^{(c)}\right)^2 \quad (5),$$

and we have chosen not limit the off-diagonal elements of $\Sigma_{xx}^{(post)}$, other than $\left|\left(\Sigma_{xx}^{(post)}\right)_{ij}\right| < \sigma_i^{(c)}\sigma_j^{(c)}$, guaranteed by (5).

In order to achieve the required accuracy, we may choose to measure once or several times some of the quality parameters and to estimate the other on the basis of the measurements made and the a priori correlations between the quality parameters. Assuming we know the cost associated to each of the measurements, we then may solve for optimal laboratory set up.

When the joint probability distribution of system state is multivariate Gaussian, we know from optimal estimation theory [13] that dividing the quality parameters into two groups, $x=[x_1 \ x_2]$ and measuring x_1 with measurement errors having a multivariate normal joint probability distribution $X_1^{(obs)} \sim N_d(x^{(1)}, C_{11})$, the estimates for x_2 and the covariance matrix describing their uncertainties are given by

$$\begin{aligned} \hat{x}^{(1)} &= x_1^{(obs)} \\ \Sigma_{11}^{(post)} &= C_{11} \\ \hat{x}^{(2)} &= \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1^{(obs)} - \mu_1) \\ \Sigma_{22}^{(post)} &= \Sigma_{22} - \Sigma_{21}(\Sigma_{11}^{-1} - \Sigma_{11}^{-1}C_{11}\Sigma_{11}^{-1})\Sigma_{12} \end{aligned} \quad (6).$$

Here Σ_{11} is the submatrix of Σ_{xx} for variable set x_1 , and respectively for other Σ_{ij} and μ_i .

The diagonal elements of the covariance matrix of estimates, $\Sigma_{ii}^{(post)}$, give the left hand side of equation (5). Then we may proceed to solve for the lowest cost laboratory setup satisfying the constraint of equation (5). If we repeat some of the measurements of $x^{(1)}$, this affects only the matrix C_{11} in the analysis above. Therefore measurement setups of repeats are also solved with the same approach, or formally:

$$\begin{aligned} [k]^* &= \arg \min_{[k]} c([k]) \\ s.t. & \end{aligned} \quad (7)$$

$$[k] = (k_1, \dots, k_n); \quad k_i = 0, 1, \dots$$

$$\left(\Sigma_{xx}^{(post)}([k])\right)_{ii} < \left(\sigma_i^{(c)}\right)^2; \forall i$$

When the a priori distribution is not Gaussian, the methods of nonlinear estimation need to be applied. However, following the principle outlined above.

The analysis above solved which measurements are to be made when only a priori information was the statistical dependence between the state variables. Therefore, it does not give us any information about how often the measurements are to be made. In process management, we on one hand the requirement that condition (5) must be satisfied at all times and on the other hand we have the additional a priori information from previous measurements and the state estimate based on those. Indeed, a need for new measurement arises, because this information no longer satisfies the condition (5).

The estimation theory provides the uncertainty of the estimate immediately after the measurement is made. It is intuitively obvious that as time progresses and now measurements are made, the estimate can still be considered as the estimate of the process state, but the uncertainty increases over time. The approach requiring least knowledge about the process is to assume that the process state undergoes a random walk so that the estimate uncertainty increases over time as [14-17]:

$$\begin{aligned} \left(\sum_{xx}^{(post)}\right)_{ii}(t_n + \tau) &= \left(\sum_{xx}^{(post)}\right)_{ii}(t_n) + D_{ii}\tau \\ 0 < \tau < t_{n+1} - t_n \end{aligned} \quad (8)$$

where t_n is the instant when the n th measurement/estimation was made and D_{ii} is the diffusion parameter of the random walk of process state.

With the assumption of equation (8), estimation method (e.g. equation (6)) and constraints, equation (5), we may formulate an optimization problem: which measurement we need to make and how often (interval τ^*) to keep the knowledge about the quality within the required accuracy:

$$\begin{aligned} ([k], \tau)^* &= \arg \min_{[k], \tau} c([k], \tau) \\ s.t. & \\ [k] &= (k_1, \dots, k_n); \quad k_i = 0, 1, \dots \\ \left(\sum_{xx}^{(post)}([k])\right)_{ii} + D_{ii}\tau &< \left(\sigma_i^{(c)}\right)^2; \quad \forall i \end{aligned} \quad (9).$$

Although defining the constraint for uncertainty (5), is extremely challenging for practical decision makers and the cost structure of making measurements may be much more complicated than each measurement having its cost independent from possible other measurements made at the same time, we claim the approach practical. We shall now proceed in applying the approach to analysis of quality management and related laboratory activities at a paper mill.

4. QUALITY MANAGEMENT AT A PAPER MILL

In process industries such as papermaking the quality management is commonly based on three level hierarchical measurement structure: accurate but costly and infrequent laboratory measurements, automated quality analyzers sampling more frequently and mimicking laboratory analyses, and indirect but frequent on-line measurements for

automatic control. The decisions supported with this information can be divided into three categories: continuous process and quality management, special actions, and configuration of the measurement information system itself.

At paper mills the three decision categories supported are active quality control through broke management and apportioning raw materials, detecting off-specifications products to be rejected, and calibration of on-line quality sensors. The accuracy constraints for quality information (5) can be derived from analysis of these decisions.

Hierarchy of measurements at paper mills consists of frequent on-line measurements and more accurate laboratory analysis. Between those the mills have laboratory analyzers that are like robots acting quite well according to the laboratory analysis standards. Measuring frequency of analyzers is usually once per machine reel, or 1-3 times an hour, whereas laboratory analyses are made at most 3 times a day. However, laboratory analyzers are less accurate. The hierarchy can be used as validation stairs or use laboratory analysis results to directly validate on-line measurements.

The laboratory activities should be derived by analyzing in which decisions and how the data will actually be used. There are only a few mills that have carried out such an analysis, and no mills that have applied decision analysis to specify how accurately the measured parameters should be known.

As a particular case we consider the management of optical properties of paper. Brightness, opacity and L-a-b color coordinates are the optical quality parameters of paper. The optical quality specifications of printing paper grades set by customers are tight as the visual appearance of printed products hinges on these quality parameters. The optical properties are well standardized and there exist laboratory devices of high accuracy to measure these parameters. However, such laboratory activities are labor intensive and also require investing in devices and systems. These properties can be measured with automated quality analyzers that have similar investment costs but the operational costs per sample are much lower. The practices of combining laboratory analyses and automated quality analyzers has been developed over time into their present form, and it may be questioned whether they are close to optimal. Should a green field production line be built, how should the laboratory activities be set up?

The optimal measurement setup of optical quality in paper is based on analysis of quality measurement data from a paper mill, over six months period from both a laboratory analyzer (27 quality measurements) and laboratory analysis (15 quality parameters). The analysis concentrated on one paper grade only. Cost of availability and operation for laboratory analyzer is same regardless of amount of quality parameters measured. Laboratory analysis is labor intensive and cost of availability and operation varies between quality parameters. The target is to find out if laboratory analyses are needed at all, and if they are needed, how often they

must be made to maintain the accuracy of optical quality information.

Following linear optimal estimation we divide quality parameters into two groups $x = [x_1 \ x_2]$ and measuring x_2 generate the estimates for x_1 .

We define the quality factor Q_i of measurement information as

$$Q_i = \frac{\left(\sum_{xx}^{(post)}\right)_{ii}}{\left(\sigma_i^{(c)}\right)^2} \quad (10)$$

so that the constraints (5) are expressed as $Q_i < 1$.

Objective was to find correlating quality parameters between these two – laboratory analysis and laboratory analyzer – methods and to find the group of quality parameters that can be estimated using measurement results from laboratory analyzer only, thus reducing the laboratory work. After that labor intensive laboratory analysis can be focused to those quality parameters that cannot be estimated using this model and are needed

Data was divided into two parts; first part (3 months) was used to estimate the optimal set of measurements and other part (the next three months) to validate the quality parameter estimation. All 27 of laboratory analyzer quality parameters were used in estimation. Figure 3 shows four histograms of actual optical quality parameters (laboratory analysis, black) and estimated optimal quality parameters (from laboratory analyzer, grey). The unexplained percentage of variance per estimated quality parameter is: [0.0875 0.1186 0.1772 0.1709]. As $(\sigma_i^{(c)})^2$ tends to be 0.25..0.01 of total variance, we conclude that the first two quality parameters are ones that may be estimated with single analyzer results, whereas the other two not.

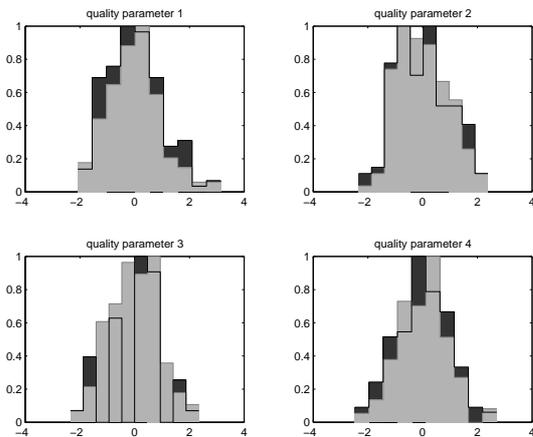


Figure 3. Histograms of four quality parameters, (laboratory analysis, black) and estimated quality parameters (from laboratory analyzer, grey).

The quality parameters that cannot be estimated only with quality analyzer results, must be measured in laboratory. There is a diffusion constant, see (8), associated with laboratory information. The diffused laboratory information and the information based on analyzer results can be fused to reduce the number of laboratory analyses. Our preliminary analysis shows, however, that the present practice on such measurements is close to optimal.

Figure 4 shows four histograms of quality parameters (laboratory analysis, black) and estimated quality parameters (from laboratory analyzer, grey) using validation data set.

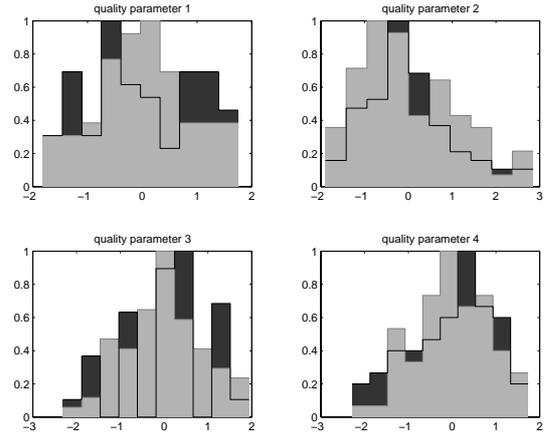


Figure 4. Histograms of four quality parameters, (laboratory analysis, black) and estimated quality parameters (from laboratory analyzer, grey) using validation data set.

The analysis of the validation shows in general that the predictive power of laboratory analyzer, using the same covariance matrix Σ_{12} , has degraded substantially. Therefore, occasional laboratory measurements are needed to dynamically validate the covariance matrix. The frequency of measurement is, according to our preliminary analysis, much lower than current practice. Hence, we have identified an opportunity to reduce laboratory work and focus it to where it generates most value.

The analysis with the presented framework – formal or expert derivation of accuracy of quality information required, optimal estimation analysis of opportunities to replace labor-intensive measurements by estimates, and dynamic degradation analysis to derive frequency of measurements - pinpoints critical measurements and concentrates more effort on them. In most cases the analysis process itself is of high importance: it provides a shared and documented view on performance requirements for the quality measurement activities; the accuracy of information, the availability and the costs related. Knowledge about the engineered accuracy and reliability of the measurements increases the operators trust at the quality parameters, thus supporting and improving decision making.

5. CONCLUSIONS

Measurements are uncertain and everything derived from them is uncertain. Decisions are always made under uncertainty. Value of measurement information is determined by the decisions made based on them.

In this paper we have outlined a framework for determining the value of measurement information and designing a measurement information system – what to measure and how often – based on the value generated under a defined set of scenarios. The framework relies on that we have explicitly defined utility for each decision making situation and that we have explicitly defined scenarios of external effects on the system, including their frequency of occurrence. Admittedly, these are strong – in most cases unrealistic – assumptions. Therefore we noted that the result of optimal design of measurement information system can be expressed as optimal accuracy of measurement information of state, and that such optimal accuracy can be obtained from human experts, in addition to formal approach.

We discussed the framework in a practical case of quality management at paper mills. We showed a potential for replacing some of the laborious laboratory measurements by estimates based on results from an automated laboratory analyzer. However, we also noted that for such an estimation to work over long period of time, occasional laboratory measurements must be made to keep the estimate of covariance matrix reliable. When outlining the framework we restricted ourselves to direct value generated by the measurements, i.e. how much the measurement improves decision making by providing more accurate information about the system state. The updating of covariance matrix is an example of indirect value generation: the measurement improves how we derive state information from other measurements. The framework will be expanded to tackle the indirect effects as well.

Decision making is difficult task, which can be made easier with more accurate and focused measurements. Effort for making measurements can be focused when measuring of every parameter is not needed. Thus more time and energy can be used for making those few measurements and thus also calculated parameters become more accurate.

Acknowledgements. The authors are grateful towards The Finnish Funding Agency of Technology (TEKES) through grant for project “Tuotanto 2010”. The authors thank Mr. Juuso Gren for processing the case study data.

REFERENCES

[1] K. Latva-Käyrä, “Dynamic Validation of On-line Consistency Measurement”. Ph.D. thesis, Tampere University of Technology, 105 p, Tampere 2003.
[2] see H. Raiffa, R. Schlaifer, “Applied Statistical Decision Theory”, Wiley Classics Edition (2000, originally published 1960).

[3] R. Ritala, J. Belle, K. Holmström, H. Ihalainen., J. Ruiz, M. Suojärvi, M. Tienari, ”Operations Decision Support based on Dynamic Simulation and Optimization”, Proceedings of Efficiency, PulPaper 2004, Conferences, - Energy, Coating, Efficiency, Helsinki, Finland. pp. 55-62.
[4] L. T. Biegler, I. E. Grossmann, A. W. Westerberg, “Systematic Methods of Chemical Process Design”. London: Prentice-Hall, (1997).
[5] C.A. Floudas, X. Lin, ”Continuous-time versus discrete-time approaches for scheduling of chemical processes: A review”, Computers and Chemical Engineering 28, 2109-2129 (2004) and references therein
[6] M. Morari, J. H. Lee, “Model predictive control: past present and future, Comp Chem. Eng, 23, 667-682, (2000), and references therein
[7] H. Jokinen, K. Latva-Käyrä, P. Pulkkinen, R. Ritala, ”Modelling of Uncertainty: Case Studies on Operation of Papermaking Process”, Proceedings MATHMOD 2006, Vienna, Austria, February 8-10, ARGESIM Report 30 (2006).
[8] S. French, D. Rios Insua, “Statistical Decision Theory”, Kendall’s Library of Statistics, Vol. 9, Arnold Publishing, 2000.
[9] for a collection of most import papers and relevant critique, see P. Gärdenfors, N.-E. Sahlin, “Decision, Probability and Utility – Selected Readings”, Cambridge University Press, 1988.
[10] J.O. Berger, “Statistical Decision Theory and Bayesian Analysis”. Springer, 617 p, USA 1980.
[11] J. M. Bernardo & A. F. M. Smith, “Bayesian Theory”. Wiley, 586 p, Great Britain 1994.
[12] H. Jokinen, R. Ritala, “Value assessment of measurements in large measurement information systems”. Proceedings of the 13th Int. Symp. on Measurements for Research and Industry Applications and the 9th European Workshop on ADC Modelling and Testing. Athens, Greece, pp. 367-372.
[13] see F. L. Lewis, “Optimal estimation: with an introduction to stochastic control theory”, Wiley-Interscience, 1986.
[14] see for example, G. R. Grimmett, D.R. Strizaker “Probability and Random Processes”, Ch. 13, Oxford University Press, 2001.
[15] H. Ihalainen, K. Latva-Käyrä, R. Ritala, ”Dynamic validation of on-line measurements: A probabilistic analysis”, Measurement, 39, 335-351 (2006).
[16] K. Latva-Käyrä, R. Ritala, “Sensor Diagnostics Based on Dynamic Characteristic Curve Estimation”. Proceedings of the 10th IMEKO TC10 International Conference, 9-10 June, 2005, Budapest, Hungary.
[17] K. Latva-Käyrä, R. Ritala, ”Dynamic validation of multivariate linear soft sensors with reference laboratory measurements”. In: Kappen, J., Manninen, J. & Ritala, R. (eds.) COST ACTION E36 Workshop: Modeling and simulation in pulp and paper industry. (URL:<http://www.vtt.fi/inf/pdf/>). p. 57-64, (Oct 6th, 2005).