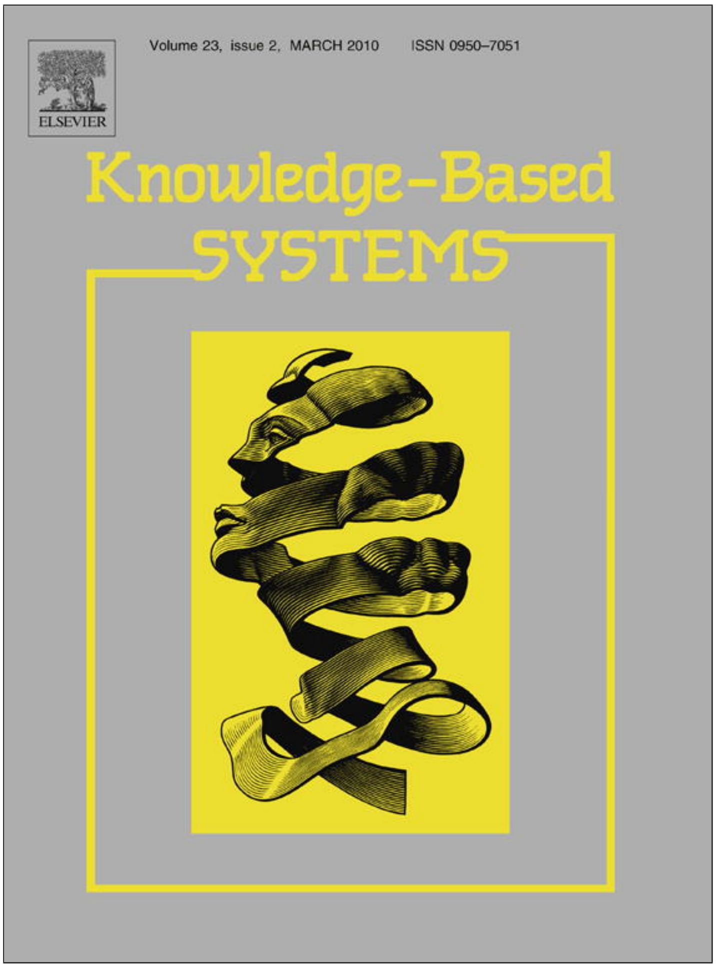
Volume 23, issue 2, MARCH 2010 ISSN 0950-7051
ELSEVIER
Knowledge-Based
SYSTEMS

Knowledge-Based Systems 23 (2010) 182–194

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Short communication

New results in modelling derived from Bayesian filtering

Claudiu Pozna [a], Radu-Emil Precup [b,*], József K. Tar [c], Igor Škrjanc [d], Stefan Preitl [b]

[a] Department of Product Design and Robotics, Transilvania University of Brasov, Bd. Eroilor 28, 500036 Brasov, Romania
[b] Department of Automation and Applied Informatics, "Politehnica" University of Timisoara, Bd. V. Parvan 2, 300223 Timisoara, Romania
[c] Institute of Intelligent Engineering Systems, Budapest Tech Polytechnical Institution, Bécsi út 96/B, H-1034 Budapest, Hungary
[d] Faculty of Electrical Engineering, Laboratory of Modelling, Simulation and Control, University of Ljubljana, Tržaška 25, 1000 Ljubljana, Slovenia

ARTICLE INFO

Article history:
Received 9 November 2008
Received in revised form 9 November 2009
Accepted 13 November 2009
Available online 4 December 2009

Keywords:
Knowledge
Model
Modelling algorithm
Plausible reasoning
Simulation

ABSTRACT

This paper suggests an original heuristic modelling algorithm expressed in terms of homogenous combinations of the classical system dynamics and the Bayesian degree of truth employed in modelling. The main benefits of the proposed approach compared to classical modelling are the increased transparency and alleviated computational time. Two case studies, dealing with a mobile robot and an unforced pendulum system, are included to exemplify and test the theoretical results. One of the case studies makes use of the definition and calculation of several discrete plausibilities.

1. Introduction

In essence, modelling is postulating assumptions how real-world behaves. One may refine indefinitely the model but the difference between the model and the real-world phenomenon, process or dynamical system which is subject to modelling cannot be avoided [10,22,34,45,49]. Increasing the quality of the models has been one of the favourite research directions during the last years. This involves the exponential growth of non-conventional modelling based on knowledge-based systems (KBS) tools including fuzzy logic [5,26,29,46,50,55], neural networks [1,33,51,52], genetic algorithms [14,24,38,48], data mining techniques [8,9,36,56] etc., and their merge resulting in hybrid models [2,6,7,25,30,39].

Another research direction regarding modelling deals with measuring the approximation capability of the models. More precisely not only the behaviour of the system modelled is calculated but also the degree of truth associated with the prediction ensured by the model. This approach is a merge between the traditional modelling and the Bayesian plausible reasoning rules [4] referred in [3]. If one is able to associate a certain degree of truth (plausibility) for the results corresponding to the model, then a decision can be taken emphasizing whether the results are either suitable or they become discordant [13]. In addition, it will be possible to increase in the appropriate moment the plausibility of the model in terms of observations. Nevertheless, the possibility to choose the appropriate time of observation and the need of measuring the plausibility of the observation considered are pointed out. In the framework of this second direction a set of rules, referred to as plausible reasoning rules, has been proposed in [28]. The proof has been done on the basis of the probability theory.

An epistemology has been suggested in [53] where the Bayesian filtering equations represent the core of the human knowledge model. One of the important aspects of this work is that it underlines the possibility of induction process modelling. Definitely the plausibility by observation is highlighted. The new idea of [53] is to use the concept of measuring the plausibility and offer an example regarding the increase of the degree of truth by observation. The results in causal reasoning [40] can be quoted in the same direction.

The approach based on just Bayesian rules has various applications including the analysis of generalization [47] or rational thinking [19] and econometrics [54]. The engineering applications are reported in data fusion [41], actuators control [32] and the mobile robotics field involving robust perception and risk assessment in highly dynamic environments [11,12].

Models which combine the plausible reasoning with traditional modelling have been proposed in [41]. It is important to highlight the practical aspects of this approach because the authors have constructed models based on Bayesian filtering. The models are implemented in real-time systems and emphasize a plausible

* Corresponding author. Tel.: +40 256 403229, +40 256 403226, +40 256 403230, +40 256 403240; fax: +40 256 403214.
E-mail addresses: cp@unitbv.ro (C. Pozna), radu.precup@aut.upt.ro (R.-E. Precup), tar.jozsef@nik.bmf.hu (J.K. Tar), igor.skrjanc@fe.uni-lj.si (I. Škrjanc), stefan.preitl@aut.upt.ro (S. Preitl).

0950-7051/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2009.11.015

reasoning problem construction. The structure of such a construction is important and it consists of two levels, the problem description and the question. The first level contains also two parts, the specification of the model and the identification of its parameters. The main drawback of this construction is its computational cost. However the structure of the problem construction is suitable for a plausible reasoning expert but it is difficult for an engineer who is usually familiar only with traditional modelling.

The first aim of this paper is to give a new and attractive derivation of Bayesian filtering equations for the case of discretized probability measures. The derived forms of the equations hold even for partially ordered sets rather than probability measures. However use is made of generally acknowledged valid probability measures in the illustrative examples offered in the paper.

The second aim is to suggest an algorithm that will enable gaining knowledge on system dynamics expressed as the set of vectors with two elements, the behaviour of the system (the output of the system i.e. its dynamic response) and the plausibility (the degree of truth) of its result. That set of vectors is defined as follows:

$$\left\{ \begin{bmatrix} y_i \\ P(y_i) \end{bmatrix} \Bigg|_{i=\overline{1,n}} \right\}, \tag{1}$$

where y_i represents the output of the dynamical system which is subject to modelling, $P(y_i)$ is the plausibility of the sentence "the output of the system has the value y_i" and n stands for the total number of model iterations with the iteration index i. This algorithm is referred to as modelling algorithm and the third aim of this paper is to simulate its behaviour.

The paper is organized as follows in order to achieve the three aims. The following section is focused on the definition of the model which combines the classical state-space representation of the dynamical model with the plausibility of its result i.e. the system output. Next, Section 3 is dedicated to the new heuristic modelling algorithm. Section 4 validates the theoretical approach by the presentation of two case studies, and Section 5 concludes the paper.

2. Model elements

The model elements to be constructed as follows are defined in terms of (1). This section will treat aspects concerning that definition by presenting the homogenous combination between the classical state-space model characterizing the system dynamics and the plausibility of its results obtained by Bayesian filtering. The Section consists of two parts, the first one is focused on the computation of output and output plausibility and the second one presents the computation of the state vector and its plausibility.

The theoretical aspects regarding the plausibility in modelling are extracted from [28]. They are formulated in terms of the points 1 and 2:

1. The representation of plausibility is given by the plausibility function

   $p : \Phi \to [0, 1], p(A|X) = y,$

   where Φ is an ordered set of sentences defined over a linear normed space and $p(A|X)$ is a continuous and monotonic function which assigns a certain degree of truth to the sentence A under the condition that the sentence X is true.
2. The consistency of the common sense requires that the function p should fulfil the following properties:

$$p(AB|X) = p(A|X)p(B|X),$$
$$p(A|X) + p(\neg A|X) = 1,$$
$$p(A + B|X) = p(A|X) + p(B|X) - p(AB|X),$$
$$p((A_i|X)) = 1/n, i = \overline{1, n},$$

where $\{A_i\}_{i=\overline{1,n}}$ is a complete set of mutually exclusive sentences.

Four comments are highlighted in relation with 1 and 2, to be presented as follows. First, by consistency we consider:

- Each possible way of reasoning concerning a certain sentence must lead to the same result.
- All equivalent sentences have an equal degree of plausibility.
- In order to obtain the plausibility of a sentence one must account for all available evidence.

Second, the plausibility can involve:

- A priori sentences. For example, the plausibility of the sentence "the output of the system will be y" can take a value which is nonzero although a posteriori the output of the system is z, because it is known that the model is an approximation of the reality in which unknown disturbances may occur.
- A posteriori sentences. For example, the plausibility of the sentence "the output of the system is y" can take a value which is nonzero even it has been measured that the output value is z, because it is known that the sensors are subject to disturbances.

Third, according to [15–17] the plausibility theory can be viewed as a generalization of the probability theory in terms of replacing the probability measure by a partially ordered set. Hence, if the chosen set meets the Kolmogorov axioms (related to the points 1 and 2 presented before), the resulting theory is equivalent to the probability calculus. With this regard, the probability theory is well established in model construction and identification. A great advantage of Bayesian statistics applied here is that it offers a way how to test the validity of the modelling assumptions mentioned in Section 1 in the light of data. Thus, the model construction is reduced to the choice of model structure in the considered approach.

Fourth, the plausibility and the probability have the same axiomatic basis and the main difference of these concepts is epistemological. In connection with the epistemological aspects of the concept of probability it may be noted that probabilities have been accepted in [27] as measurable physical quantities e.g. masses or velocities. The real basis of this philosophy was the unlimited repeatability of certain experiments. In other words, a priori one can say "the probability or the plausibility of the sentence "the dice will be 1" takes the value 1/6" but a posteriori it is not suitable to say "the probability of the sentence "the dice is 1" is...".

The plausibility will be used in the sequel as connected to the disturbances. This point of view will be employed in both modelling and measurement.

2.1. Computation of system output, output estimation and plausibility of output measurement

The accepted dynamical systems can be modelled in the general multi input-multi output (MIMO) case by the following state-space mathematical model:

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} = \mathbf{g}(\mathbf{x}, \mathbf{u}) \end{cases}, \tag{2}$$

where $\mathbf{x} \in R^n$ is the system state vector, $\mathbf{u}$ is the system input vector, $\mathbf{y}$ stands for the system output vector and $\mathbf{f}$ and $\mathbf{g}$ represent generally nonlinear vector functions. Without reducing the generality of the algorithm mentioned before, the case of single input-single output (SISO) linear time-invariant (LTI) systems will be considered in this paper, so (2) can be expressed in its particular form

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u \\ y = \mathbf{c}^T\mathbf{x} + du \end{cases}, \tag{3}$$

with $\mathbf{A} \in R^{n\times n}$, $\mathbf{b} \in R^{n\times 1}$, $\mathbf{c} \in R^{n\times 1}$, $d \in R$ and the superscript T denotes matrix transposition.

The computation of system output relies on (3) making use of analytical or numerical methods. Since the analytical ones are rather complex even in case of higher-order SISO LTI systems, the numerical methods are preferred. That is the reason why the first-order differential equations in (3) will be transformed into difference ones:

$$\begin{cases} \frac{\mathbf{x}_i - \mathbf{x}_{i-1}}{\Delta t} = \mathbf{A}\mathbf{x}_{i-1} + \mathbf{b}u_{i-1} \\ y_i = \mathbf{c}^T\mathbf{x}_i + du_i \end{cases}, \tag{4}$$

equivalent to the recurrent equations:

$$\begin{cases} \mathbf{x}_i = \mathbf{A}^*\mathbf{x}_{i-1} + \mathbf{b}^*u_{i-1} \\ y_i = \mathbf{c}^T\mathbf{x}_i + du_i \end{cases}, \quad \mathbf{A}^* = \mathbf{A}\Delta t - \mathbf{I}, \mathbf{b}^* = \mathbf{b}\Delta t, \tag{5}$$

where $\mathbf{I} \in R^{n\times n}$ is the unity matrix and Δt represents the sampling time. Using (5) one may compute iteratively the values of the state vectors $\mathbf{x}_i, i = \overline{1,n}$. The initial state vector of the dynamical system $\mathbf{x}_0$ and the input u_i are needed in order to do those iterations.

At this level it is intended to improve the quality of the model (5) by including the unknown or eluded aspects as mentioned earlier. All these aspects are expressed by the output disturbances p_i^{est}. Inserting the disturbances will transform the output of the system into the estimated output. The connection between the system output and the output estimation is

$$y_i^{\text{est}} = y_i + p_i^{\text{est}}, \tag{6}$$

where y_i^{est} is the output estimation i.e. the output computed with modelling disturbance accounted for, y_i is the output of model (5) with no disturbances considered in this model and p_i^{est} stands for uncertainty or output disturbance.

The value of the disturbance is not known a priori but one may obtain by experiments the statistical distribution of p_i^{est}, referred to as $P(p_i^{\text{est}})$. A discrete statistical distribution function is defined with this respect:

$$P(p_i^{\text{est}}) = \{P(p_{i,j}^{\text{est}}) | j = \overline{1,m}\}, \tag{7}$$

with m – the number of elements in the domain of P. Specifically, $P(p_{i,j}^{\text{est}})$ is the truth value of the sentence "at iteration i the system output is y_j", where y_j is an element belonging to the output domain. Since (6) results in

$$\begin{aligned} y_i^{\text{est}} - y_i &= y_i^{\text{est}} - \left(\mathbf{c}^T\mathbf{x}_i + du_{i-1}\right) \\ &= y_i^{\text{est}} - \mathbf{c}^T(\mathbf{A}^*\mathbf{x}_{i-1} + \mathbf{b}^*u_{i-1}) - du_{i-1}, \end{aligned} \tag{8}$$

it is justified to accept that:

$$P(p_i^{\text{est}}) = P(y_i^{\text{est}} - y_i) \equiv P(y_i^{\text{est}} | + \mathbf{x}_{i-1}), \tag{9}$$

where (9) highlights that the degree of the plausibility of y_i^{est} in the condition of accepting that the state vector $\mathbf{x}_{i-1}$ equals the statistical distribution of the model uncertainty. In addition, making use of the Bayesian plausible reasoning rules [4], $P(y_i^{\text{est}})$ can be computed in terms of

$$P(y_i^{\text{est}}) \propto \sum_{\mathbf{x}_{i-1}} \left(P(\mathbf{x}_{i-1})P(y_i^{\text{est}} | \mathbf{x}_{i-1})\right), \tag{10}$$

where: $P(y_i^{\text{est}})$ – the plausibility of the output estimation, $P(\mathbf{x}_{i-1})$ – the plausibility of the state vector $\mathbf{x}_{i-1}$, $P(y_i^{\text{est}} | \mathbf{x}_{i-1})$ – the plausibility of the estimated output when the state vector $\mathbf{x}_{i-1}$ is known, $\propto$ – symbol standing for proportionality.

Doing the observation means that the system output is measured and the process of measurement yields the measured output y_i^{meas}. To measure means to use a sensor that can be characterized also by a mathematical model according to

$$y_i^{\text{meas}} = y_i^{\text{est}} + e(y_i) + p_i^{\text{meas}}, \tag{11}$$

where $e(y_i)$ is the sensor error function and p_i^{meas} is the measurement disturbance. This value is unknown but its statistical distribution, $P(p_i^{\text{meas}})$, is in fact known by experiments:

$$P(p_i^{\text{meas}}) = P(y_i^{\text{meas}} - y_i^{\text{est}} - e(y_i)) \equiv P(y_i^{\text{meas}} | y_i^{\text{est}}). \tag{12}$$

The interpretation of (12) is that the statistical distribution of measurement disturbance equals the plausibility of the measured output y_i^{meas} when the estimated value of the output y_i^{est} is known. Use is made again of the Bayesian plausible reasoning rules leading to

$$P(y_i^{\text{meas}}) \propto P(y_i^{\text{est}})P(y_i^{\text{meas}} | y_i^{\text{est}}), \tag{13}$$

where: $P(y_i^{\text{meas}})$ – the plausibility of the output measurement, $P(y_i^{\text{est}})$ – the plausibility of the output estimation, $P(y_i^{\text{meas}} | y_i^{\text{est}})$ – the plausibility of the measured output when the output estimation is assumed to be known.

Some remarks are outlined in relation with the results presented above. First, a distribution of output estimation and measurements has been obtained using (10) or (13). Second, the degree of truth of each sentence "at iteration i the estimated or the measured output is y_j" has been calculated. This result represents the second element in the vectors defined in (1), and the first element in those vectors is the more plausible result y_j. Definitely, from all possibilities $j = \overline{1,m}$ the one that gives the maximum degree of truth of the mentioned sentence is selected. Since (10) involves the plausibility of the dynamical system state vectors, this aspect needs also further discussion.

2.2. Computation of state vectors

It is important to outline that the state vector of the dynamical system is used in (10). This requires that the state vectors $\mathbf{x}_{i-1}$ should be known before computing y_i^{est} and y_i^{meas}. Usually an observer is used to compute the current state vector $\mathbf{x}_i$ from the values of y_i^{meas} and u_i. The observer design is based on (5) and makes use of

$$\begin{cases} \mathbf{s}_i = \mathbf{F}^*\mathbf{s}_{i-1} + \mathbf{G}^*[y_i^{\text{meas}} \quad u_i]^T \\ \mathbf{x}_i^{\text{com}} = \mathbf{H}\mathbf{s}_i + \mathbf{J}[y_i^{\text{meas}} \quad u_i]^T \end{cases}, \tag{14}$$

with the matrices and vectors $\mathbf{s}_i \in R^{nx1}$, $\mathbf{F}^* \in R^{n\times n}$, $\mathbf{G}^* \in R^{n\times 2}$, $\mathbf{H} \in R^{n\times n}$, $\mathbf{J} \in R^{nx2}$. Similarly to (5), the quality of the observer model can be increased by adding the model disturbance:

$$\mathbf{x}_i = \mathbf{x}_i^{\text{com}} + \mathbf{p}_i^{\text{est}}, \tag{15}$$

where $\mathbf{x}_i$ is the state vector of the dynamical model, $\mathbf{x}_i^{\text{com}}$ represents the computed state vector associated with the dynamical model and $\mathbf{p}_i^{\text{est}}$ stands for a disturbance vector. Each element in $\mathbf{p}_i^{\text{est}}$ is computed as function of the output disturbance p_i^{est}.

The statistical distribution of the model disturbance p_i^{est} is known. Under these conditions, straightforward calculations regarding (14) and (15) will result in

$$\mathbf{x}_i - \mathbf{x}_i^{\text{com}} = \mathbf{x}_i - \mathbf{H}\left(\mathbf{F}^*\mathbf{s}_{i-1} + \mathbf{G}[y_i^{\text{meas}} \quad u_i]^T\right) - \mathbf{J}[y_i^{\text{meas}} \quad u_i]^T. \tag{16}$$

Therefore the plausibility of the state vector $\mathbf{x}_i$ is known accepting that the state vector $\mathbf{x}_{i-1}$ and the output y_i^{meas} are known:

$$P(\mathbf{p}_i^{\text{est}}) = P(\mathbf{x}_i - \mathbf{x}_i^{\text{com}}) \equiv P(\mathbf{x}_i|\mathbf{x}_{i-1}, y_i^{\text{mas}}). \quad (17)$$

Concluding, the application of the Bayesian plausible reasoning rules will lead to

$$P(\mathbf{x}_i) \propto \sum_{y_i^{\text{meas}}} P(y_i^{\text{meas}}) \sum_{\mathbf{x}_{i-1}} (P(\mathbf{x}_{i-1})P(\mathbf{x}_i|\mathbf{x}_{i-1}, y_i^{\text{meas}})), \quad (18)$$

where $P(\mathbf{x}_i)$ is the plausibility of the state vector $\mathbf{x}_i$, $P(\mathbf{x}_{i-1})$ is the plausibility of the state vector $\mathbf{x}_{i-1}$, $P(y_i^{\text{meas}})$ is the plausibility of the measured output y_i^{meas} and $P(\mathbf{x}_i|\mathbf{x}_{i-1}, y_i^{\text{meas}})$ is the plausibility of the state vector $\mathbf{x}_i$ when $\mathbf{x}_{i-1}$ and y_i^{meas} are known.

The result expressed in terms of (18) completes the previous outcomes presented in (10) and (13) because it offers the possibility to compute the plausibility of the state vector. In order to make these results more intelligible and more suitable the next Section is devoted to the development of a modelling algorithm in the Bayesian filtering framework.

3. Modelling algorithm

As shown in the previous section the plausibility is based on statistical distributions. That result can be interpreted as after a statistical analysis of the experimental data one may choose one of the well-known distributions [18,35]. One short presentation with this regard is summarized in Table 1.

Gaussian distributions are generally the most frequently used ones [20,37]. Without reducing the generality of the modelling algorithm to be presented in the sequel, only these distribution types will be accepted and expressed as

$$P(p_{i,j}) = N \exp\left(-\frac{p_{i,j}^2}{2\sigma^2}\right), \quad (19)$$

with: $p_{i,j}$ – the (independent) variable of the considered distribution, $P(p_{i,j})$ – the plausibility of the variable $p_{i,j}$, σ – the distribution variance, $i = \overline{1,n}$ – the iteration index employed in the computation of the output, $j = \overline{1,m}$ is the index of the current element in the set on which distribution is defined with a total number of m elements of this set, and the parameter N is obtained in terms of

$$\sum_{j=1}^{m} P(p_{i,j}) = 1. \quad (20)$$

Since the Gaussian distribution is defined for any real $p_{i,j}$, one challenge is to choose an appropriate domain which will not lose information and will also admit a convenient computational time. The following condition has been imposed to solve this problem:

$$\frac{\min_j (P(p_{i,j}))}{\max_j (P(p_{i,j}))} = 5\%, \quad j = \overline{1,m}. \quad (21)$$

Using (21) the width of the definition domain D, referred to as $l(D)$, results as follows as related to the distribution variance σ (Fig. 1):

$$l(D) = 2\sqrt{6}\sigma. \quad (22)$$

Since the simulation domain D^*, with $y_i \in D^*$, is larger than D i.e. $l(D^*) > l(D)$ and the distribution must be defined over this domain, the distribution $P(p_{i,j})$ has been extended by inserting zero elements according to Fig. 2 and

$$P(p_{i,j}) = 0, \forall p_{i,j} \in D^* \setminus D. \quad (23)$$

In order to accomplish one of three aims of this paper consisting of an algorithm offering knowledge on system dynamics, the previous results must be analyzed and appropriate mathematical forms should be expressed. Also, because the main drawback of the plausible model is its computational time, mathematical forms which will decrease the computational time will be derived in the sequel. With this regard the first step is to define a less dense distribution (23). In the same context, the calculations start with (10) that can be re-expressed as

$$\begin{aligned} P\left(y_{i,j}^{\text{est}}\right) &\propto \sum_{\mathbf{x}_{i-1,j}} \left(P(\mathbf{x}_{i-1,j})P\left(y_{i,j}^{\text{est}}|\mathbf{x}_{i-1,j}\right)\right) \\ &= \left[P\left(y_{i,1}^{\text{est}}|\mathbf{x}_{i-1,1}\right)\ P\left(y_{i,1}^{\text{est}}|\mathbf{x}_{i-1,2}\right)\ \ldots\ P\left(y_{i,1}^{\text{est}}|\mathbf{x}_{i-1,m}\right)\right] \\ &\quad \cdot \left[P(\mathbf{x}_{i-1,1})\ P(\mathbf{x}_{i-1,2})\ \ldots\ P(\mathbf{x}_{i-1,m})\right]^T. \end{aligned} \quad (24)$$

Next, (24) can be transformed into the following matrix form:

Table 1
Frequently used distributions.

Distribution	Distribution laws, average and dispersion	Comments
Binomial	$P_n(k) = C_n^k p^k (1-p)^{n-k}$, $\mu = np$, $\sigma^2 = np(1-p)$	The distribution is used when the number of experiments and the variety of results is small. *Example*: The 6 points dice. P_n is the distribution law; μ is average; σ is the dispersion
Poisson	$\psi_n(k) = \frac{a^k e^{-a}}{k!}$, $\mu = a$, $\sigma^2 = a$	The distribution is used when we have made many experiments and the variety of results is also numerous. *Example*: accidents. Ψ_n is the distribution law; μ is average; σ is the dispersion
Gaussian	$p(x) = \frac{1}{a\sqrt{2\pi}} \exp\left(-\frac{(x-b)^2}{2a^2}\right)$, $\mu = b$, $\sigma^2 = a^2$	The distribution is used when we have made many experiments and the variety of results is not so numerous. *Example*: measuring errors. p is the distribution law; μ is average; σ is the dispersion
Uniform	$u_n(k) = a$	This distribution is used when we can not obtain any knowledge on the output value. The only thing we know in this case is that the output belongs to a particular domain but each value of the mentioned domain can be reached with the same plausibility. *Example*: the probability that the dice will be on 1, 2, 3, 4, 5, 6. u is the distribution law

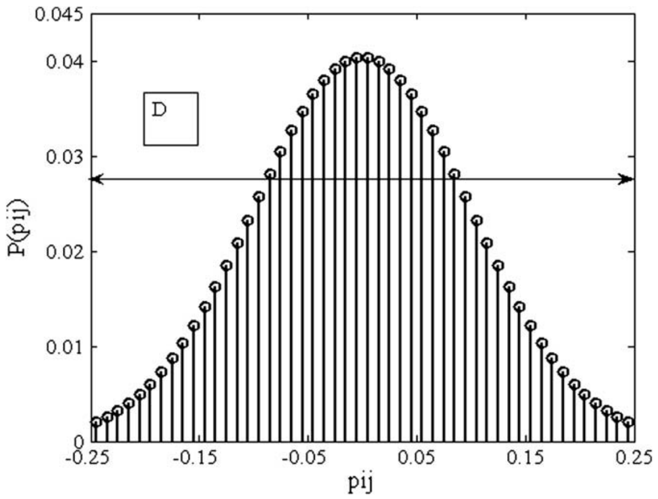

Fig. 1. Distribution $P(p_{i,j})$ versus $p_{i,j}$ for $p_{i,j} \in D$.

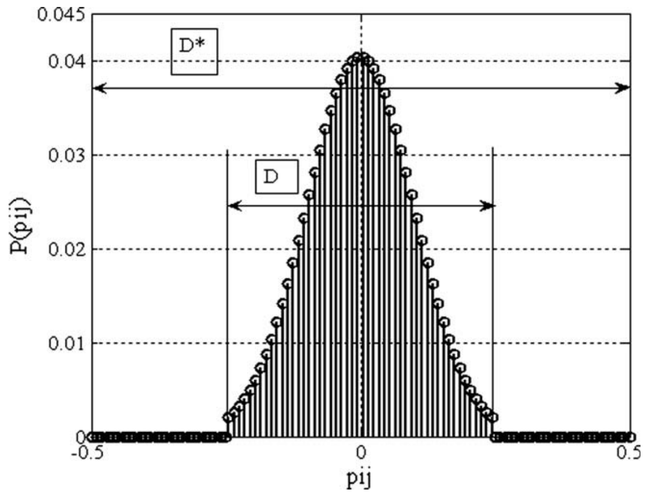

Fig. 2. Distribution $P(p_{i,j})$ versus $p_{i,j}$ for $p_{i,j} \in D^*$.

$$\mathbf{p}(y_i^{est}) \propto \mathbf{P}^{est}\mathbf{p}(\mathbf{x}_{i-1}),$$
$$\mathbf{P}^{est} = \begin{bmatrix} P\left(y_{i,1}^{est}|\mathbf{x}_{i-1,1}\right) & P\left(y_{i,1}^{est}|\mathbf{x}_{i-1,2}\right) & \dots & P\left(y_{i,1}^{est}|\mathbf{x}_{i-1,m}\right) \\ P\left(y_{i,2}^{est}|\mathbf{x}_{i-1,1}\right) & P\left(y_{i,2}^{est}|\mathbf{x}_{i-1,2}\right) & \dots & P\left(y_{i,2}^{est}|\mathbf{x}_{i-1,m}\right) \\ \dots & \dots & \dots & \dots \\ P\left(y_{i,m}^{est}|\mathbf{x}_{i-1,1}\right) & P\left(y_{i,m}^{est}|\mathbf{x}_{i-1,2}\right) & \dots & P\left(y_{i,m}^{est}|\mathbf{x}_{i-1,m}\right) \end{bmatrix}, \quad (25)$$
$$\mathbf{p}(y_i^{est}) = \left[P\left(y_{i,1}^{est}\right) \quad P\left(y_{i,2}^{est}\right) \quad \dots \quad P\left(y_{i,m}^{est}\right)\right]^T,$$
$$\mathbf{p}(\mathbf{x}_{i-1}) = \left[P(\mathbf{x}_{i-1,1}) \quad P(\mathbf{x}_{i-1,2}) \quad \dots \quad P(\mathbf{x}_{i-1,m})\right]^T.$$

where $\mathbf{p}(y_i^{est})$ is the plausibility vector of the estimated output, $\mathbf{p}(\mathbf{x}_{i-1})$ is the plausibility vector of the state vector of the dynamical system, and $\mathbf{P}^{est} = \mathbf{P}(y_i^{est}|x_{i-1}), i = \overline{1,n}$, is called the output estimation matrix. It is important to highlight that $\mathbf{P}^{est} = \text{const}$ i.e. if simulations are done, the matrix can be computed before the simulations.

A lot of multiplications can be avoided to simplify the computations in (25) making use of the fact that $\mathbf{P}^{est}$ is a band matrix. Unfortunately this property can result in a relative alleviation of the number of zero elements in $\mathbf{p}(y_i^{est})$ according to Fig. 3a which illustrates a first method to decrease the computational time. Fig. 3a is based on the hypothesis that the output estimation matrix is a band matrix. This hypothesis is achieved for non-oscillatory systems, and shows the mechanism of multiplication viz. the decrease of the number of zero elements after each iteration. This means that after each iteration the number of nonzero elements in $\mathbf{p}(y_i^{est})$ is doubled.

Another possibility is to approximate the result by using only the maximum element in the vector $\mathbf{p}(x_{i-1})$ and the correspondent column in $\mathbf{P}^{est}$. This can be expressed under the form of the second method to decrease the computational time presented in Fig. 3b. The number of nonzero elements in $\mathbf{p}(y_i^{est})$ will be equal to that of the nonzero elements of the k-th column in the estimation matrix. Fig. 3b illustrates a diagram of the multiplication process when the mentioned maximum is the k-th element. This means that the number of after nonzero elements in $\mathbf{p}(y_i^{est})$ can be kept constant at each iteration. Fig. 3 illustrates the variation of the number of nonzero elements in $\mathbf{p}(y_i^{est})$ according to the matrix multiplication rules (in Fig. 3a) and proposed approximation (in Fig. 3b).

Making use of the second method (25) can be transformed into

$$\mathbf{p}(y_i^{est}) \propto \mathbf{p}_k^{est} \max_{j=\overline{1,m}} (\mathbf{p}(\mathbf{x}_{i-1})), \quad (26)$$

where $\mathbf{p}_k^{est} = \mathbf{P}_{\overline{1,m},k}^{est}$ is a vector representing the k-th column in the output estimation matrix and $k, k = \overline{1,m}$, is the index of the maximum element in the vector $\mathbf{p}(\mathbf{x}_{i-1})$, $P(\mathbf{x}_{i-1,k}) = \max_{j=\overline{1,m}}(P(\mathbf{x}_{i-1,j}))$.

The justification for using the maximum values of $\mathbf{p}(x_{i-1})$ in (26) is that in the simulations the interest is focused on the value of the output which has the maximum plausibility as it will be exemplified and illustrated in the sequel. Using this approximation the plausibility value is not modified. However this aspect does not affect the ability of decision making on the acceptance of our results because the minimum value of the confidence, $P(y_i)_{\min}$, is defined in terms of this approximation and included in the suggested modelling algorithm.

The proposed approach applied to the matrices in (3) results in the substitution of $(d+1)^2$ multiplications and d^2 additions by d multiplications at each iteration. Therefore the complexity of the algorithm will be reduced.

The next step deals with the transformation of (13) into its equivalent expression:

$$\left[P\left(y_{i,1}^{meas}\right) \quad P\left(y_{i,2}^{meas}\right) \quad \dots \quad P\left(y_{i,m}^{meas}\right)\right]$$
$$\propto \left[P\left(y_i^{meas}|y_{i,1}^{est}\right)P\left(y_{i,1}^{est}\right) \quad P\left(y_i^{meas}|y_{i,2}^{est}\right)P\left(y_{i,2}^{est}\right) \quad \dots \quad P\left(y_i^{meas}|y_{i,m}^{est}\right)P\left(y_{i,m}^{est}\right)\right]. \quad (27)$$

Therefore, introducing the output measurement matrix $\mathbf{P}^{meas}$:

$$\mathbf{P}^{meas} = \begin{bmatrix} P\left(y_{i,1}^{meas}|y_{i,1}^{est}\right) & P\left(y_{i,1}^{meas}|y_{i,2}^{est}\right) & \dots & P\left(y_{i,1}^{meas}|y_{i,m}^{est}\right) \\ P\left(y_{i,2}^{meas}|y_{i,1}^{est}\right) & P\left(y_{i,2}^{meas}|y_{i,2}^{est}\right) & \dots & P\left(y_{i,2}^{meas}|y_{i,m}^{est}\right) \\ \dots & \dots & \dots & \dots \\ P\left(y_{i,m}^{meas}|y_{i,1}^{est}\right) & P\left(y_{i,m}^{meas}|y_{i,2}^{est}\right) & \dots & P\left(y_{i,m}^{meas}|y_{i,m}^{est}\right) \end{bmatrix}, \quad (28)$$

the generalization of (13) becomes:

$$\mathbf{p}(y_i^{meas}) \propto \text{diag}(\mathbf{p}_r^{meas})\mathbf{p}(y_i^{est}), \quad (29)$$

where $\mathbf{p}_r^{meas} = \mathbf{P}_{r,\overline{1,m}}^{meas}$ is a vector equal to the r-th line in the measurement matrix $\mathbf{P}^{meas}$ and $r, r = \overline{1,m}$, is the index of the output measurement in the domain of y. It should be highlighted that $\mathbf{P}^{meas} = \text{const}$ and $\mathbf{P}^{meas}$ can be computed before the simulation, too. Summarizing, (25) and (29) lead to

$$\mathbf{p}(y_i^{meas}) \propto diag(\mathbf{p}_r^{meas})\mathbf{P}^{est}\mathbf{p}(\mathbf{x}_{i-1}). \quad (30)$$

The following approximation of (30) can be obtained if (26) is employed to approximate (25):

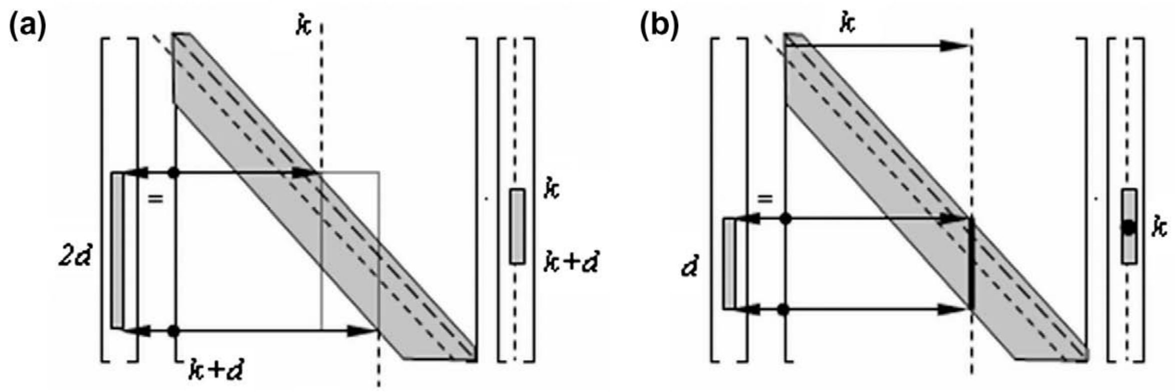

Fig. 3. Two methods to decrease the computational time: nonzero elements in $\mathbf{p}(y_i^{est})$ for matrix multiplication rules (a) and proposed approximation (b).

$$\mathbf{p}(y_i^{meas}) \propto \mathrm{diag}(\mathbf{p}_r^{meas})\mathbf{p}_k^{est} \max_{j=\overline{1,m}}((\mathbf{x}_{i-1})). \tag{31}$$

Once again this approach applied to the matrices in (3) results in the substitution of $(d+1)^2$ multiplications and d^2 additions by d multiplications at each iteration (Fig. 3).

The focus will be now on (18), where the second sum can be expressed as

$$\begin{aligned}&\sum_{\mathbf{x}_{i-1}}\left(P(\mathbf{x}_{i-1})P\left(\mathbf{x}_{i,j}|\mathbf{x}_{i-1},y_{i,j}^{meas}\right)\right)\\ &=\left[P\left(\mathbf{x}_{i,j}|\mathbf{x}_{i-1,1},y_{i,j}^{meas}\right)\quad P\left(\mathbf{x}_{i,j}|\mathbf{x}_{i-1,2},y_{i,j}^{meas}\right)\quad\ldots\quad P\left(\mathbf{x}_{i,j}|\mathbf{x}_{i-1,m},y_{i,j}^{meas}\right)\right]\\ &\cdot\left[P(\mathbf{x}_{i-1,1})\quad P(\mathbf{x}_{i-1,2})\quad\ldots\quad P(\mathbf{x}_{i-1,m})\right]^T,\end{aligned} \tag{32}$$

and it is transformed further into

$$\begin{bmatrix}\sum_{\mathbf{x}_{i-1}}\left(P\left(\mathbf{x}_{i,1}|\mathbf{x}_{i-1},y_{i,1}^{meas}\right)P(\mathbf{x}_{i-1})\right)\\ \sum_{\mathbf{x}_{i-1}}\left(P\left(\mathbf{x}_{i,2}|\mathbf{x}_{i-1},y_{i,2}^{meas}\right)P(\mathbf{x}_{i-1})\right)\\ \cdots\\ \sum_{\mathbf{x}_{i-1}}\left(P\left(\mathbf{x}_{i,m}|\mathbf{x}_{i-1},y_{i,1}^{meas}\right)P(\mathbf{x}_{i-1})\right)\end{bmatrix}=\mathbf{P}_j^{sta}\mathbf{p}(\mathbf{x}_{i-1}),$$

$$\mathbf{P}_j^{sta}=\begin{bmatrix}P\left(\mathbf{x}_{i,1}|\mathbf{x}_{i-1,1},y_{i,j}^{meas}\right) & P\left(\mathbf{x}_{i,1}|\mathbf{x}_{i-1,2},y_{i,j}^{meas}\right) & \ldots & P\left(\mathbf{x}_{i,1}|\mathbf{x}_{i-1,m},y_{i,j}^{meas}\right)\\ P\left(\mathbf{x}_{i,2}|\mathbf{x}_{i-1,1},y_{i,j}^{meas}\right) & P\left(\mathbf{x}_{i,2}|\mathbf{x}_{i-1,2},y_{i,j}^{meas}\right) & \ldots & P\left(\mathbf{x}_{i,2}|\mathbf{x}_{i-1,m},y_{i,j}^{mas}\right)\\ \cdots & \cdots & \cdots & \cdots\\ P\left(\mathbf{x}_{i,m}|\mathbf{x}_{i-1,1},y_{i,j}^{meas}\right) & P\left(\mathbf{x}_{i,m}|\mathbf{x}_{i-1,2},y_{i,j}^{meas}\right) & \ldots & P\left(\mathbf{x}_{i,m}|\mathbf{x}_{i-1,m},y_{i,j}^{meas}\right)\end{bmatrix},$$
$$j=\overline{1,m}. \tag{33}$$

Finally the use of (18) and (33) results in

$$\begin{aligned}&\mathbf{p}(\mathbf{x}_i)\propto\Pi^{sta}\mathbf{p}(y_i^{meas}),\\ &\Pi^{sta}=\left[\mathbf{P}_1^{sta}\mathbf{p}(\mathbf{x}_{i-1})\quad\mathbf{P}_2^{sta}\mathbf{p}(\mathbf{x}_{i-1})\quad\ldots\quad\mathbf{P}_m^{sta}\mathbf{p}(\mathbf{x}_{i-1})\right]^T\\ &\qquad=\mathbf{P}^{sta}\mathbf{p}(\mathbf{x}_{i-1}),\mathbf{P}^{sta}=\left[\mathbf{P}_1^{sta}\quad\mathbf{P}_2^{sta}\quad\ldots\quad\mathbf{P}_m^{sta}\right]^T,\\ &\mathbf{p}(y_i^{meas})=\left[P\left(y_{i,1}^{meas}\right)\quad P\left(y_{i,2}^{meas}\right)\quad\ldots\quad P\left(y_{i,m}^{meas}\right)\right]^T,\end{aligned} \tag{34}$$

with $\mathbf{P}^{sta}=\mathrm{const}$ – the state observation matrix, which can be computed before the simulation.

Two comments are outlined in relation with (34). First, because the measurements are not done and it is intended to continue the estimation of the output, (34) may be used in terms of replacing the measured output by the estimated one:

$$\mathbf{p}(\mathbf{x}_i)\propto\Pi^{sta}\mathbf{p}(y_i^{est}). \tag{35}$$

Second, according to the approximation method presented before and illustrated in Fig. 3b, the computational time can be reduced in (34) by transforming it into its approximated version:

$$\mathbf{p}(\mathbf{x}_i)\propto\pi_h^{sta}\max_{j=\overline{1,m}}\left(\mathbf{p}(y_i^{meas})\right), \tag{36}$$

where the vector $\pi_h^{sta}=\Pi_{\overline{1,m},h}^{sta}$ stands for the h-th column in the matrix Π^{sta} and h is index of the maximum element in the vector $\mathbf{p}(y_i^{meas})$. In other words, $P(y_{i,h}^{meas})=\max_{j=\overline{1,m}}(P(y_{i,j}^{meas}))$.

The combination between the plausible reasoning formalism and the traditional dynamical modelling allows obtaining the dynamical system output and the degree of truth (the plausibility) of the result. The combination can be defined, as mentioned in Section 1, like knowledge on system dynamics expressed in terms of the set of vectors

$$\left\{\begin{bmatrix}y_i\\ P(y_i)\end{bmatrix}\Bigg|_{i=\overline{1,n}}\right\}=\left\{\begin{bmatrix}y_{i,v}\\ P\left(y_{i,v}^{meas}\right)\end{bmatrix}\Bigg|i=\overline{1,n},P\left(y_{i,v}^{meas}\right)=\max_{j=\overline{1,m}}\left(P\left(y_{i,j}^{meas}\right)\right)\right\}, \tag{37}$$

where $i, i=\overline{1,n}$, is the iteration index used in the proposed modelling algorithm. In addition, the two components of the knowledge on system dynamics can be illustrated in Fig. 4. One may remark that the knowledge on system dynamics is a homogenous result because both elements in the vectors in (37) are obtained from the plausibility distribution of system behaviour viz. system output. It must be emphasized that this is a more comprehensive result than just the system output itself because it associates the designer's degree of confidence in this value, and this idea has been incorporated in the proposed algorithm. The strategy of the modelling algorithm starts with setting a certain minimum value of confidence in the system output, $P(y_i)_{\min}$. Then the model is run until the confidence reaches its minimum value. Next, in order to continue, measurements are necessary because the plausibility of the results will increase.

Since the crisp values of plausibility are needed in the simulation the proportionality relations derived in this section must be transformed into equations. In order to do this the vectors will be normalized after each iteration according to (38)–(40):

$$\mathbf{p}(y_i^{est})=\frac{\mathbf{p}(y_i^{est})}{\sum_{j=1}^m p_j(y_i^{est})},\quad i=\overline{1,n}, \tag{38}$$

$$\mathbf{p}(y_i^{meas})=\frac{\mathbf{p}(y_i^{meas})}{\sum_{j=1}^m p_j(y_i^{meas})},\quad i=\overline{1,n}, \tag{39}$$

$$\mathbf{p}(\mathbf{x}_i)=\frac{\mathbf{p}(\mathbf{x}_i)}{\sum_{j=1}^m p_j(\mathbf{x}_i)},\quad i=\overline{1,n}. \tag{40}$$

The flowchart of this iterative algorithm is presented in Fig. 5. The following comments concerning the steps of the algorithm are pointed out:

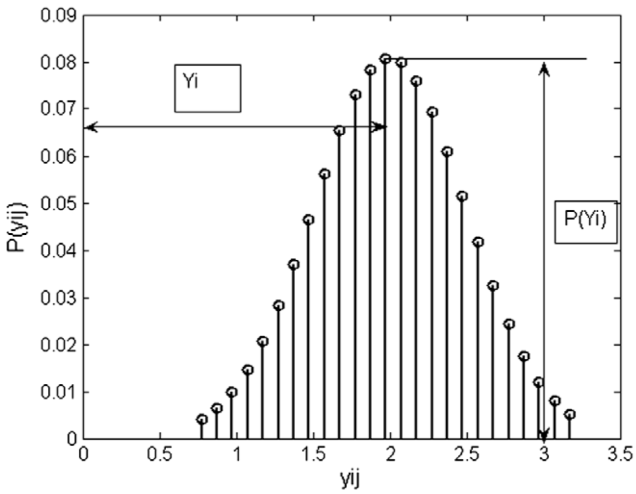

Fig. 4. The two components of knowledge on system dynamics.

- Set the initial data:
  - identify the model parameters and write the state-space model (5),
  - do experiments leading to the estimation matrix $\mathbf{P}^{est}$,
  - do computations yielding the state matrix $\mathbf{P}^{sta}$,
  - choose a sensor, do experiments and set the output measurement matrix $\mathbf{P}^{meas}$,
  - set an acceptable plausibility for the output $P(y_i)_{\min}$,
  - set an input signal u_i.

- Compute the output estimation:
  - make use of (25) or (26) and compute the estimation vector $\mathbf{p}(y_i^{est})$,
  - employ (38) to normalize the vector $\mathbf{p}(y_i^{est})$,
  - if the plausibility of the results is acceptable i.e. $P(y_i) \geqslant P(y_i)_{\min}$, then continue with the state estimation process, else do measurements.

- Do the output measurements:
  - make use of (30) or (31) and compute the measurement vector $\mathbf{p}(y_i^{meas})$,
  - employ (39) to normalize the vector $\mathbf{p}(y_i^{meas})$,
  - if the plausibility of the results is acceptable i.e. $P(y_i) \geqslant P(y_i)_{\min}$, then continue with the state estimation process, else do measurements.

- Do the state estimation:
  - if measurements have been done, then make use of (35) or (36) and compute the plausibility vector of the state vector of the model $\mathbf{p}(\mathbf{x}_i)$,
  - if no measurements have been done, then make use of (35) and compute the state of the model $\mathbf{p}(\mathbf{x}_i)$,
  - employ (40) to normalize the vector $\mathbf{p}(\mathbf{x}_i)$.

The key idea is to iterate the algorithm until the confidence in the predicted output reached the desired value. The degrees of freedom in the algorithm are represented by the initial data. Therefore it is guaranteed that the predicted output will always increase $P(y_i)$ over the desired level and the algorithm will not enter any infinite loop.

4. Case studies

This Section is dedicated to the validation of the heuristic modelling algorithm proposed in Section 3 by two case studies.

We have chosen the first case study because it is a non-oscillatory system and the state, estimation and measurement matrices are band matrices. This means that we have been able to exemplify our proposed algorithm with or without the proposed approximation. The second case study is an oscillatory system. The reason of that additional case presentation is to give a clear picture of the proposed algorithm. That is also the reason why in the first case study we have chosen a Gaussian distribution and in the second case study a uniform one.

In the first case study it will be shown that if a model is used iteratively, then the plausibility of results will decrease. It is obvious that using iteratively a model the confidence in results will be smaller and smaller. The knowledge on system dynamics suggested in this paper offers one tool to measure the confidence in results.

Another important issue concerns the measuring effect. Absolutely, the confidence in results increases by measuring. With this regard, the knowledge on system dynamics provides one mathematical tool to measure this increase concerning the confidence in results.

Contrarily, since the second case study concerns a system which converges to a particular value, it will be shown that during the simulation the plausibility of the result increases when the system is closer to the equilibrium state.

4.1. Case study 1

The system studied in this case study is represented by the simplified model of a mobile robot with the block diagram presented in Fig. 6. The model parameters are the mobile mass $m = 1$ kg and the viscous friction $b = 0.1$ N s^2/m. The model input is the constant force $u = 1$ N and the model output is the robot speed x_2. The speed can be measured using a sensor but the position x_1 of the vehicle robot must be computed.

The dynamical model of the robot can be characterized in terms of either the differential equations (continuous time state-space model)

$$\begin{aligned} \dot{\mathbf{x}} &= \begin{bmatrix} 0 & 1 \\ 0 & -0.1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \\ y_i &= [0 \quad 1]\mathbf{x} \end{aligned}, \tag{41}$$

where $\mathbf{x} = [x_1 \quad x_2]^T$ is the robot state vector, or the recurrent equations (discrete time state-space model)

$$\begin{aligned} \mathbf{x}_i &= \begin{bmatrix} 1 & \Delta t \\ 0 & 1 - 0.1\Delta t \end{bmatrix} \mathbf{x}_{i-1} + \begin{bmatrix} 0 \\ \Delta t \end{bmatrix} u, \\ y_i &= [0 \quad 1]\mathbf{x}_i \end{aligned} \tag{42}$$

where $\mathbf{x}_i = [x_{i,1} \quad x_{i,2}]^T$ is the robot state vector at the i-th iteration, $x_{i,1}$, and $x_{i,2}$ stand for the position and the velocity of the robot, respectively, at the i-th iteration. The initial state vector of the system is accepted to be $\mathbf{x}_0 = [0 \quad 0]^T$ and the value of the sampling time is set to $\Delta t = 0.5$s. In order to define the discrete plausibility distribution the following values have been chosen: $x_{i,1} = 0.1l, l = \overline{0,n}$, $x_{i,2} = 0.1l, l = \overline{0,n}$, $n = 200$.

The statistical distribution of output disturbance$P(p^{est})$, the statistical distribution of measurement disturbance $P(p^{meas})$, the sensor error function $e(y)$ and the statistical distribution of state vector disturbance $P(p^{sta})$ have been set as follows:

$$\begin{aligned} &P(p^{est}) \propto \exp\left(-\frac{(p^{est})^2}{2(0.2)^2}\right), P(p^{meas}) \propto \exp\left(-\frac{(p^{meas})^2}{2(0.2)^2}\right), \\ &e(y) = 0.5, P(p^{sta}) \propto \exp\left(-\frac{(p^{sta})^2}{2(0.2)^2}\right). \end{aligned} \tag{43}$$

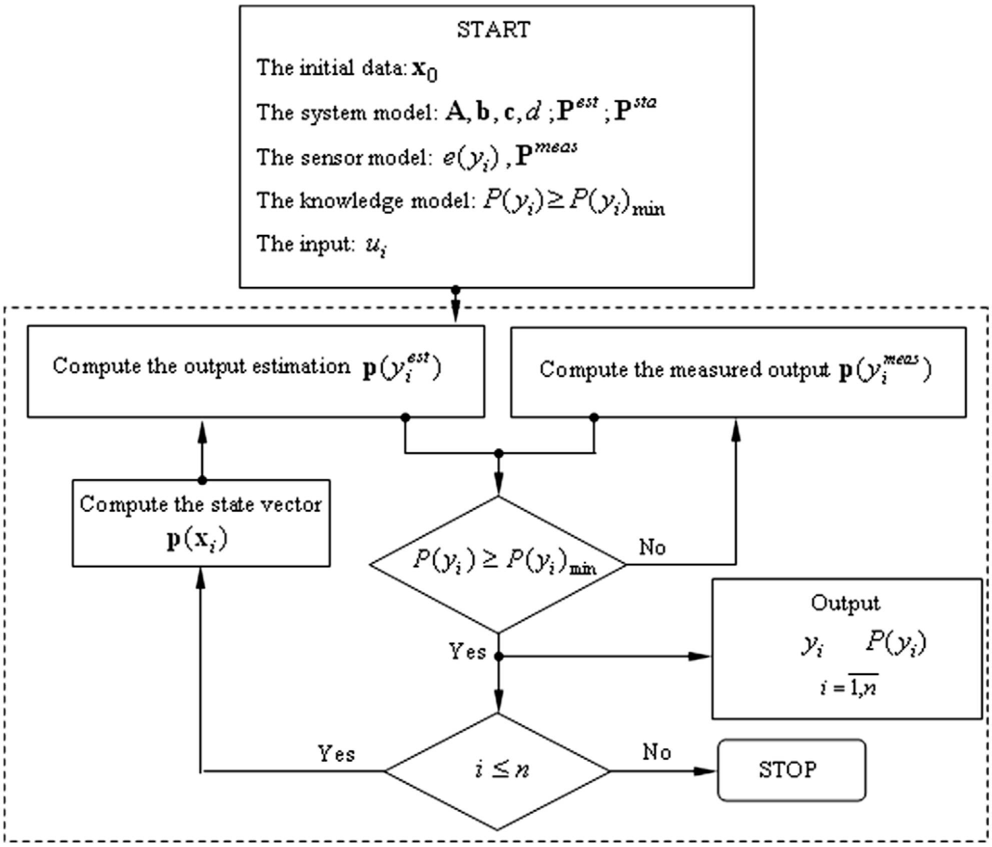

Fig. 5. Flowchart of modelling algorithm.

The knowledge on system dynamics is illustrated in Fig. 7 accepting the conditions of no measurements during the simulation process. The results have been obtained after eight iterations and a Gaussian distribution has been plotted per iteration.

One may observe that the plausibility decreases. Although this information is expected and obvious, the importance of the suggested algorithm is that the degree of confidence can be measured and decisions can be taken according to the observations. Fig. 8 illustrates the knowledge on system dynamics if several measurements are done. The results presented in Fig. 8 have been obtained for four iterations. After a certain velocity has been computed three successive measurements were done, and the computed value of the output is confirmed each time a measurement was done. Fig. 8 suggests that the measurement will increase the confidence in results. That information is obvious, but the advantage of the new algorithm is the possibility to measure the increase in plausibility and make decisions based on it.

Next, Fig. 9 shows a comparison between the knowledge on system dynamics in two cases, simulation without (in Fig. 9a) and with (in Fig. 9b) the proposed approximation described in (30) and (36). The curves in Fig. 9 have been obtained for 4 successive iterations. One iteration consists of estimations followed by measurements; two plausibility distributions are plotted after each iteration.

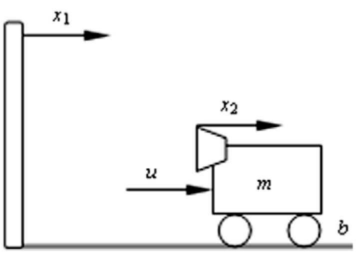

Fig. 6. Block diagram of mobile robot.

The simulations in both cases deal with the following sequence of operations:

- the transitions from the state vector $\mathbf{x}_{i-1}$ to the state vector $\mathbf{x}_i$,
- measuring i.e. digital simulation of this process, and
- computing the state vectors.

A comparison of these results reveals that no differences between the system output y_i can be observed, but a careful analysis can highlight the differences between the plausibility degrees $P(y_i)$. One of the main benefits of the proposed approximation technique is the decreasing of the computational time. That time decrease can be assessed and illustrated by the decrease of the number of elementary arithmetic operations. In this case study using (25) for d=10 (Fig. 3) leads to the results presented in Table 2.

Fig. 10 points out a comparison between the system outputs y_i for three types of simulations: using the differential Eq. (41), the recurrent Eq. (42) and extracting the output from the knowledge on system dynamics.

The errors between the models based on differential and recurrent equations can be alleviated by decreasing Δt. The results obtained from the knowledge on system dynamics track the results obtained from the model based on differential equations. Finally Fig. 11 illustrates the simulation results according to the proposed algorithm obtained for setting the parameter $P(y_i)_{\min} = 0.1$. The simulation starts with computing the estimated velocity and doing

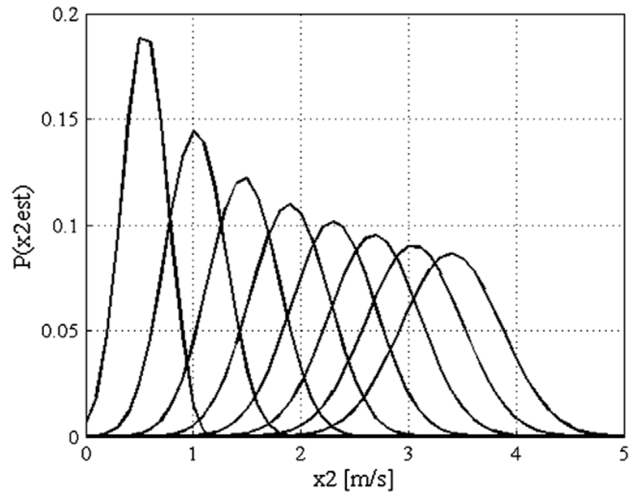

Fig. 7. Plausibility of output versus output for eight successive iterations.

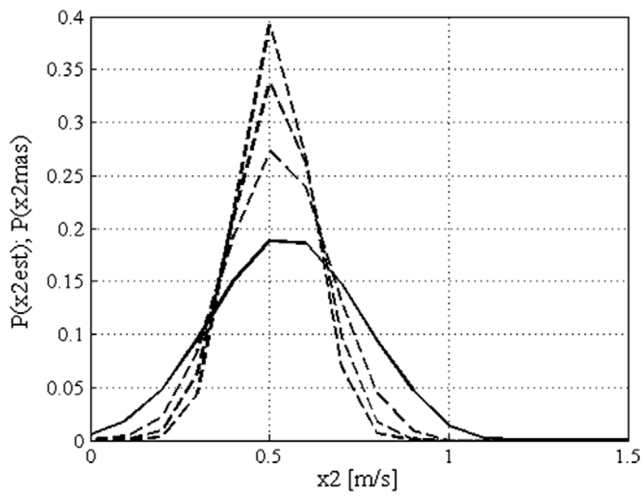

Fig. 8. Plausibility of output versus output for three successive measurements.

several iterations (five in the accepted case) until the plausibility of the output decreased bellow the confidence value (0.1 in this case). Next, a measurement was done in order to increase the plausibility of our results. That succession has been repeated two times. It is important to note that the algorithm sets automatically the moment when the measurement should be done in order to obtain trusted results.

4.2. Case study 2

The second case study considers the unforced pendulum system. The evolution of the output (the angular position of the pendulum) is caused by the initial conditions of the pendulum including its initial position. This example has been chosen due to its special behaviour that makes the output converge to an equilibrium state (position).

The dynamical model can be derived in terms of (41) using the definitions of parameters and variable according to Fig. 12:

$$ml^2\ddot{q} = m\ g\ l\sin(q) - b\dot{q}, \tag{44}$$

where m is the pendulum mass, l is the pendulum length, q is the angular position, b is the joint viscous friction and g is the gravitational acceleration, and the initial conditions are not inserted for the sake of simplicity. The following parameters have been chosen for this application: $m = 1\text{kg}$, $l = 1\text{m}$, $b = 0.1\text{Nms}$, $g = 9.8\text{m/s}^2$.

The linearization of (44) in the vicinity of the equilibrium position leads to the model:

$$ml^2\ddot{q} - m\ g\ l\ q + b\dot{q} = 0. \tag{45}$$

Digital simulations have been done in order to compare the results of the models (44) and (45). The digital simulation results obtained for the initial conditions $\{q(0) = 1, \dot{q}(0) = 0\}$ are presented in Fig. 13.

It is accepted from (5) that:

$$q_i^{est} = q_i + p_i^{est}. \tag{46}$$

The strategy will be modified in this case by the fact that the decision on the (in)accuracy of model (45) is made and trusted only on the basis of the envelope of the dynamical system response. Therefore (46) will be transformed into

$$q_i^{est} = p_i^{est}. \tag{47}$$

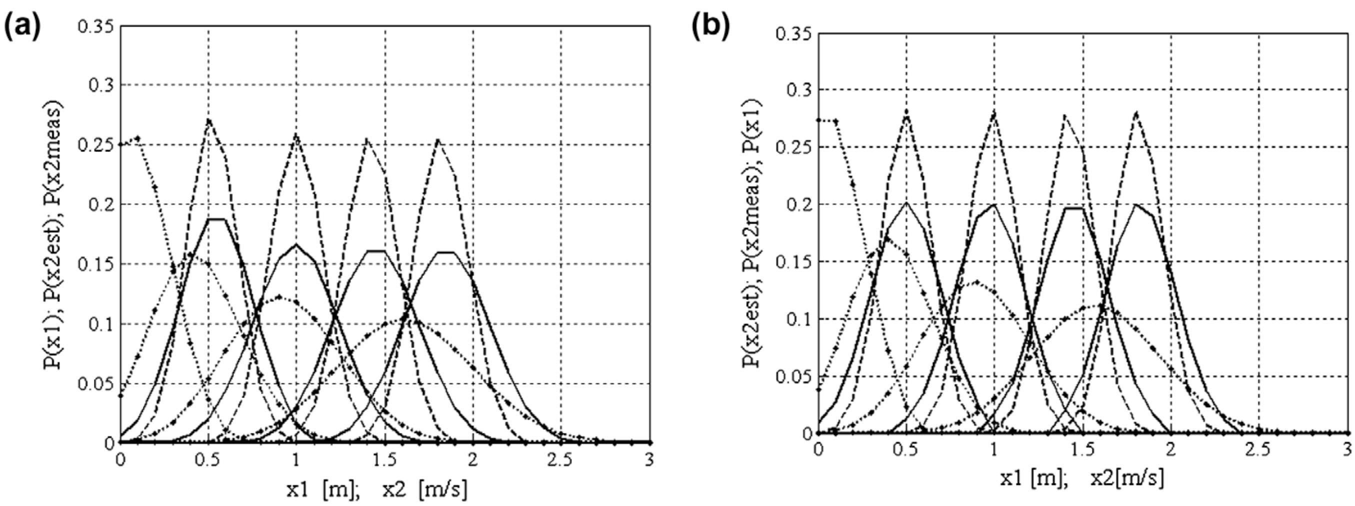

Fig. 9. Plausibility of position and velocity versus position and velocity value. Comparison between the results of two simulations without (a) and with approximation: (b): x_1 (...), x_2 after measuring (- - -), x_2 after estimation (___).

Table 2
Number of operations without and with the proposed approximation technique.

		Iteration 1	Iteration 2	Iteration 3	Iteration 4	Total
Without approximation	Multiplications	121	441	1681	6561	8804
	Additions	100	400	1600	6400	8500
With approximation	Multiplications	10	10	10	10	40
	Additions	0	0	0	0	0

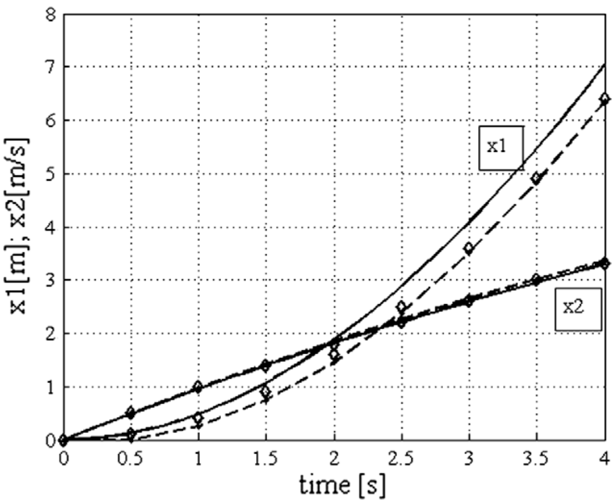

Fig. 10. Comparison between three solutions: differential equation (___), recurrent equation (- - -), model based on knowledge on system dynamics (◊).

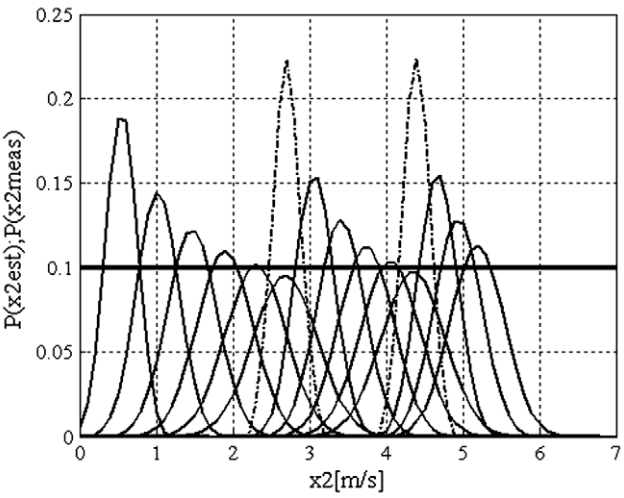

Fig. 11. Simulation results for the proposed algorithm.

The envelope of the dynamic response of the pendulum has been decided to be employed to obtain the disturbance estimation. The time scale has been divided up into several intervals. Each time interval has a maximum and minimum value of the angular position. These values represent the boundary region of the values of the angular positions. There is no knowledge on the angular position inside a certain interval i.e. it is just known that this value is inside the mentioned boundary region. Hence, the plausibility distribution is uniform. It is important to highlight that since the interval length is decreasing with respect to time (the pendulum position approaches its equilibrium position) and because the distribution is normalized, the plausibility will increase during the

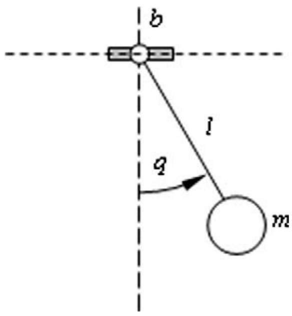

Fig. 12. Definition of variable and parameters related to the unforced pendulum system.

time. Thus, the pendulum will converge to the equilibrium position with respect to time and the designer will be more and more sure on this aspect.

In order to define the discrete plausibility the time moments $t = 0.1w, w = \overline{0,100}$, and the angle $q = -1 + 0.1l, l = \overline{0,20}$, have been chosen and the a priori information about the envelope of the pendulum angle has been set to $q(t) \in \{I_i | T_i < t \leqslant T_{i+1}; \dot{q}(T_i) = \dot{q}(T_{i+1}); i = \overline{1,11}\}$, where the values of I_i and T_i can be obtained in terms of either (45) or measurements. Furthermore, due to the convergence the relationship $I_i > I_{i+1}, i = \overline{1,11}$, is valid. All these aspects are illustrated in Fig. 14 by means of the definition of intervals in Fig. 14a and the plausibility for each interval in Fig. 14b.

Next, (9) and the strategy as mentioned earlier will result in

$$P(p_i^{\mathrm{est}}) \equiv P(q_i^{\mathrm{est}}|t) = \frac{1}{\mathrm{card}(I_k)}, t \in (T_k, T_{k+1}), \tag{48}$$

where $\mathrm{card}(I_k)$ stands for the number of elements $q \in I_k$. Accepting the conditions presented before, (10) becomes:

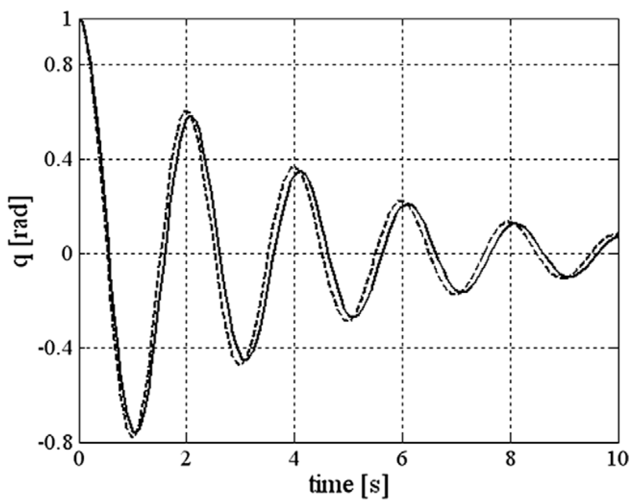

Fig. 13. Comparison between nonlinear (___) and linear model (- - -).

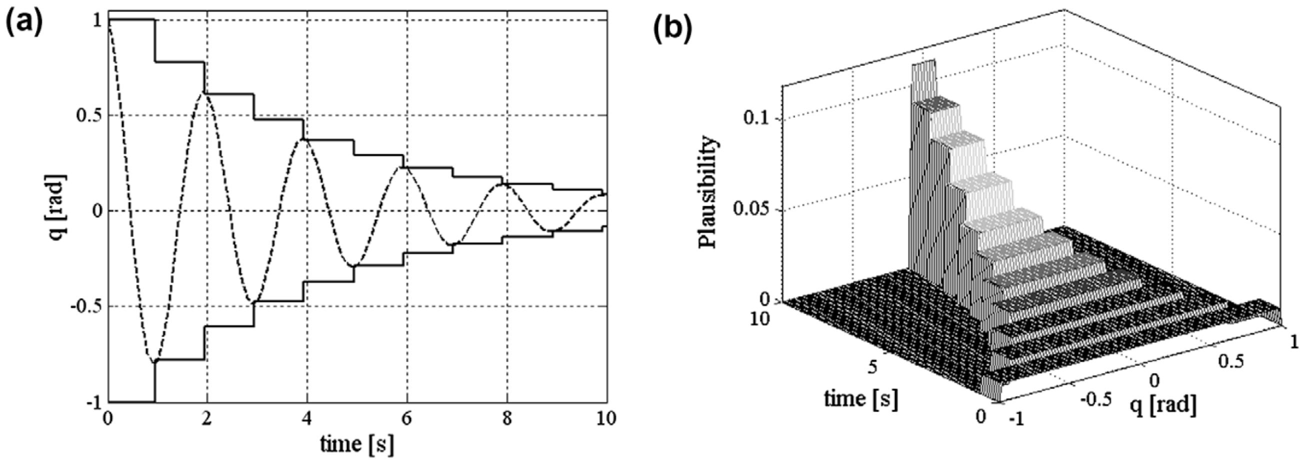

Fig. 14. Definition of time intervals as part of the time scale (a) and plausibility for each interval (b).

$$P(q_i^{est}) \propto \sum_t P(t)P(q_{i-1}^{est}|t), \tag{49}$$

where $P(q_i^{est})$ is the plausibility of the output estimation at the i-th iteration, $P(t)$ is the plausibility of the time i.e. the degree of truth that from the initial moment a certain time interval has past and $P(q_{i-1}^{est}|t)$ is the plausibility of the estimated value of time when the time at the iteration i–1 is known. The interpretation of (49) states that in order to obtain an estimation of the angular position estimations of time and next position are needed.

The final step is represented by the measurement of the position and obtaining the plausibility of the measured position. Use is made of (13) with this respect, where the Gaussian distribution:

$$P(p^{meas}) \propto \exp\left(-\frac{(p^{meas})^2}{2(0.75)^2}\right), \tag{50}$$

has been employed. The results consist of the output value and the plausibility of this value as indicated in (1). Fig. 15a illustrates a comparison between the output of the dynamical model and the output component of the model based on knowledge on system dynamics; the output variation versus time has been obtained by the digital simulation of the behaviour of (44). Fig. 15b presents the evolution of the plausibility versus time during the simulation of the model based on knowledge on system dynamics. It must be highlighted that the plausibility is the second component of the knowledge on system dynamics.

Two comments are emphasized in relation with the results presented in Fig. 15. First, the comparison in Fig. 15a illustrates the accuracy of the proposed algorithm. Second, it is important to underline that in this case the plausibility will increase with respect to time during the system evolution to the equilibrium state.

5. Conclusions

This paper merges two approaches to modelling. The first one is the traditional modelling where assumptions are postulated on the behaviour of a certain dynamical system. The second one is based on the well-known Bayesian plausible reasoning rules.

Compared to other previously used combinations [11,41] our contribution concerns the alleviation of the complexity of the model based on an original matrix-based formulation. The new modelling algorithm consists of three steps:

- prediction of the output (estimation),
- correction by measurement,
- computation of system's state variables.

Each step can make use of the information computed before simulation i.e. the output estimation matrix, the measurement ma-

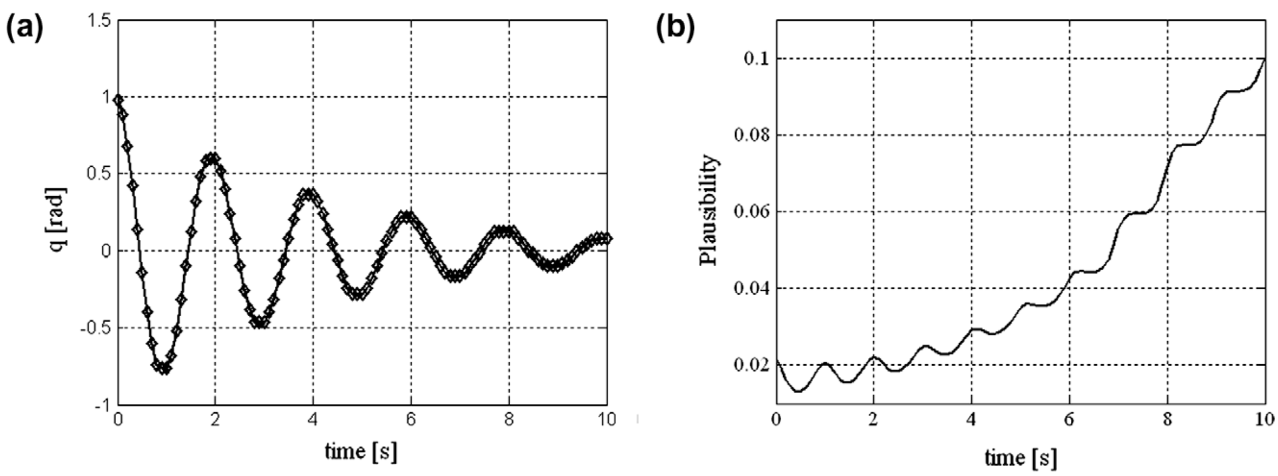

Fig. 15. (a) Comparison between the outputs versus time computed by the differential model (_) and the model based on knowledge on system dynamics (◊); (b) plausibility of output versus time.

trix and the observation matrix aiming a low computational time. Analyzing the computations necessarily to be done online, the authors have proposed approximations which decrease even more the amount of necessary computations.

The plausibility of results gives confidence in the suggested modelling approach. It is important to highlight that the plausibility (the degree of truth) is application-dependent. According to the proposed distributions the plausibility value depends on the disturbance variance and the domain sampling. The variance depends on the quality of the model, but the domain sampling represents user's choice. If the user's option is a dense sampling the precision of the model will increase and at the same time the (relative) plausibility will decrease. This has been illustrated by means of the two case studies.

Concluding, a particular domain of plausibility should be accepted for each model. In addition, the plausibility of two models should not be compared generally because each domain characterizes a certain model under a specific approach.

The heuristic modelling algorithm suggested here has been presented an application-oriented manner. Therefore the properties of the algorithm, consistency considerations and convergence proofs were not given. The particular choices made in the case study 1 i.e. linear state-space model with Gaussian disturbances are consistent with the well accepted Kalman filter. Those choices are backed up first by the fact that the Kalman filtering techniques are algorithms that perform filtering on certain linear dynamical system models to estimate the state vector. The linear models are viewed as partially observed stochastic processes with linear dynamics and linear observations, both subject to Gaussian noise. Second, the Kalman filter implements a predictor–corrector type estimation (making use of the filter time and measurement update equations) that is optimal, similarly to our modelling algorithm, in the sense that it minimizes the error covariance when certain conditions are met. Making use of these two aspects, the method suggested in this paper benefits of the well acknowledged advantages of Kalman filtering techniques including their computational efficiency and accuracy of state vector estimation. However the extensions of our method to nonlinear and MIMO systems can take over the results from Kalman filtering such as extended, unscented and particle filtering, and the scalability should be addressed.

The limitations of the proposed method and algorithm deal with the relatively large amount of heuristics and small number of variables in applications. Hence future research will deal with solving the modelling problem for complex systems with nonlinearities and large numbers of state variables. This must be accompanied by the proper definition of all variables involved and the derivation of other approaches accounting for the vicinity of the current position. The sensitivity analysis of results with respect to the parameter settings is needed and may result in optimal systems [42].

Other limitation of the paper consists of the fact that it eludes the problem of input data selection. The input data must be sufficiently excitatory such that to avoid losing the information on the dynamic behaviour of the system, and this study is worthwhile. The results of this paper will be combined with fuzzy models to be integrated in real-world applications [21,23,31,43,44].

Acknowledgements

This work was supported by the Romanian Ministry of Education, Research and Innovation through the PNII Idei Project 842/2008, and the cooperation between Budapest Tech Polytechnical Institution, "Politehnica" University of Timisoara and University of Ljubljana in the framework of the Hungarian–Romanian and Slovenian–Romanian Intergovernmental Science and Technology Cooperation Programs. The support from the CNCSIS and CNMP of Romania is dully acknowledged.

References

[1] A. Alessandri, C. Cervellera, M. Sanguineti, Design of asymptotic estimators: an approach based on neural networks and nonlinear programming, IEEE Transactions on Neural Networks 18 (1) (2007) 86–96.
[2] P. Angelov, X. Zhou, Evolving fuzzy-rule-based classifiers from data streams, IEEE Transactions on Fuzzy Systems 16 (6) (2008) 1462–1475.
[3] G.A. Barnard, Studies in the history of probability and statistics: IX. Thomas Bayes's essay towards solving a problem in the doctrine of chances, Biometrika 45 (1958) 296–315.
[4] T. Bayes, An essay towards solving a problem in the doctrine of chances, Philosophical Transactions of the Royal Society of London 53 (1763) 370–418.
[5] S. Behbahani, C.W. de Silva, Mechatronic design quotient as the basis of a new multicriteria mechatronic design methodology, IEEE/ASME Transactions on Mechatronics 12 (2) (2007) 227–232.
[6] S. Blažič, D. Matko, G. Geiger, Simple model of a multi-batch driven pipeline, Mathematics and Computers in Simulation 64 (6) (2004) 617–630.
[7] P.-C. Chang, C.-H. Liu, C.-Y. Fan, Data clustering and fuzzy neural network for sales forecasting: a case study in printed circuit board industry, Knowledge-Based Systems 22 (5) (2009) 344–355.
[8] Y.-L. Chen, C.-H. Weng, Mining fuzzy association rules from questionnaire data, Knowledge-Based Systems 22 (1) (2009) 46–56.
[9] B.-H. Chu, M.-S. Tsai, C.-S. Ho, Toward a hybrid data mining model for customer retention, Knowledge-Based Systems 20 (8) (2007) 703–718.
[10] P.W.H. Chung, L.Y.C. Cheung, C.H.C. Machin, Compliance flow – managing the compliance of dynamic and complex processes, Knowledge-Based Systems 21 (4) (2008) 332–354.
[11] C. Coué, C. Pradalier, C. Laugier, T. Fraichard, P. Bessière, Bayesian occupancy filtering for multitarget tracking: an automotive application, International Journal of Robotics Research 25 (1) (2006) 19–30.
[12] P. Dangauthier, P. Bessière, A. Spalanzani, Auto-supervised learning in the Bayesian programming framework, in: Proceedings of the 2005 IEEE International Conference on Robotics and Automation, Barcelona, Spain, 2005, pp. 3066–3071.
[13] R.O. Duda, J. Gaschnig, P.E. Hart, Model design in the prospector system for mineral exploration, in: D. Michie (Ed.), Expert Systems in the Micro Electronic Age, Edinburgh University Press, Edinburgh, 1979, pp. 153–167.
[14] M.E. ElAlami, A filter model for feature subset selection based on genetic algorithm, Knowledge-Based Systems 22 (5) (2009) 356–362.
[15] N. Friedman, J.Y. Halpern, Modeling beliefs in dynamic systems. Part I: foundations, Artificial Intelligence 95 (2) (1997) 257–316.
[16] N. Friedman, J.Y. Halpern, Modeling beliefs in dynamic systems. Part II: revision and update, Journal of AI Research 10 (1999) 117–167.
[17] N. Friedman, J.Y. Halpern, Plausibility measures and default reasoning, Journal of the ACM 48 (4) (2001) 648–685.
[18] C.M. Grinstead, J.L. Snell, Introduction to Probability, American Mathematical Society, Providence, RI, 1997.
[19] E. Heit, Properties of inductive reasoning, Psychonomic Bulletin & Review 7 (4) (2000) 569–592.
[20] R. Herzallah, D. Lowe, A mixture density network approach to modelling and exploiting uncertainty in nonlinear control problems, Engineering Applications of Artificial Intelligence 17 (2) (2004) 145–158.
[21] D. Hladek, J. Vaščák, P. Sinčák, Multi-robot control system for pursuit-evasion problem, Journal of Electrical Engineering 60 (3) (2009) 143–148.
[22] A.H. Hong, P.W.H. Chung, J. McDonald, J. Madden, Automated cause and effect analysis for process plants, in: Proceedings of 9th International Conference on Chemical and Process Engineering, Rome, Italy, 2009, pp.1281–1287.
[23] L. Horváth, I.J. Rudas, Modeling and Problem Solving Methods for Engineers, Academic Press, Burlington, MA, 2004.
[24] Y.-C. Hu, Incorporating a non-additive decision making method into multi-layer neural networks and its application to financial distress analysis, Knowledge-Based Systems 21 (5) (2008) 383–390.
[25] M.-J. Huang, Y.-L. Tsou, S.-C. Lee, Integrating fuzzy data mining and fuzzy artificial neural networks for discovering implicit knowledge, Knowledge-Based Systems 19 (6) (2006) 396–403.
[26] K.-A. Hwang, C.-H. Yang, Assessment of affective state in distance learning based on image detection by using fuzzy fusion, Knowledge-Based Systems 22 (4) (2009) 256–260.
[27] L. Jánossy, Theory and Practice of the Evaluation of Measurements, Clarendon Press, Oxford, 1965.
[28] E.T. Jaynes, G.L. Bretthorst, Probability Theory: The Logic of Science, Cambridge University Press, Cambridge, 2003.
[29] Z.C. Johanyák, A. Ádám-Majo, Fuzzy modeling of the relation between components of thermoplastic composites and their mechanical properties, in: Proceedings of 5th International Symposium on Applied Computational Intelligence and Informatics, Timisoara, Romania, 2009, pp. 481–486.
[30] A. Keles, M. Kolcak, A. Keles, The adaptive neuro-fuzzy model for forecasting the domestic debt, Knowledge-Based Systems 21 (8) (2008) 951–957.
[31] G. Klančar, D. Matko, S. Blažič, Wheeled mobile robots control in a linear platoon, Journal of Intelligent and Robotic Systems 54 (5) (2009) 709–731.
[32] K. Kording, D. Wolpert, Bayesian integration in sensorimotor learning, Nature 427 (6971) (2004) 244–247.

[33] Y. Kultur, B. Turhan, A. Bener, Ensemble of neural networks with associative memory (ENNA) for estimating software development costs, Knowledge-Based Systems 22 (6) (2009) 395–402.
[34] R.R. Leitch, Q. Shen, G.M. Coghill, M.J. Chantler, Choosing the right model, IEE Proceedings – Control Theory and Applications 146 (5) (1999) 435–449.
[35] C.E. Manski, Partial Identification of Probability Distributions, Springer-Verlag, New York, 2003.
[36] K. Marsolo, M. Twa, M.A. Bullimore, S. Parthasarathy, Spatial modeling and classification of corneal shape, IEEE Transactions on Information Technology in Biomedicine 11 (2) (2007) 203–212.
[37] F. Matía, A. Jiménez, B.M. Al-Hadithi, D. Rodríguez-Losada, R. Galán, The fuzzy Kalman filter: state estimation using possibilistic techniques, Fuzzy Sets and Systems 157 (16) (2006) 2145–2170.
[38] G. Nicosia, S. Rinaudo, E. Sciacca, An evolutionary algorithm-based approach to robust analog circuit design using constrained multi-objective optimization, Knowledge-Based Systems 21 (3) (2008) 175–183.
[39] S.-K. Oh, W. Pedrycz, H.-S. Park, Rule-based multi-FNN identification with the aid of evolutionary fuzzy granulation, Knowledge-Based Systems 17 (1) (2004) 1–13.
[40] J. Pearl, Causality: Models, Reasoning, and Inference, Cambridge University Press, Cambridge, 2000.
[41] C. Pradalier, F. Colas, P. Bessière, Expressing Bayesian fusion as a product of distributions: applications in robotics, in: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, Las Vegas, NV, 2003, pp. 1851–1856.
[42] R.-E. Precup, S. Preitl, Optimisation criteria in development of fuzzy controllers with dynamics, Engineering Applications of Artificial Intelligence 17 (6) (2004) 661–674.
[43] R.-E. Precup, S. Preitl, PI-fuzzy controllers for integral plants to ensure robust stability, Information Sciences 177 (20) (2007) 4410–4429.
[44] R.-E. Precup, S. Preitl, J.K. Tar, M.L. Tomescu, M. Takács, P. Korondi, P. Baranyi, Fuzzy control system performance enhancement by iterative learning control, IEEE Transactions on Industrial Electronics 55 (9) (2008) 3461–3475.
[45] A. Puerta, J. Eisenstein, Towards a general computational framework for model-based interface development systems, Knowledge-Based Systems 12 (8) (1999) 433–442.
[46] A. Sala, T.M. Guerra, R. Babuška, Perspectives of fuzzy systems and control, Fuzzy Sets and Systems 156 (3) (2005) 432–444.
[47] R.N. Shepard, Towards a universal law of generalization for psychological science, Science 237 (1987) 1317–1323.
[48] L. Tan, D. Taniar, Adaptive estimated maximum-entropy distribution model, Information Sciences 177 (15) (2007) 3110–3128.
[49] E. Tsui, B.J. Garner, S. Staab, The role of artificial intelligence in knowledge management, Knowledge-Based Systems 13 (5) (2000) 235–239.
[50] J. Vaníček, I. Vrana, S. Aly, Fuzzy aggregation and averaging for group decision making: a generalization and survey, Knowledge-Based Systems 22 (1) (2009) 79–84.
[51] W. Wen, Y.H. Chen, H.H. Pao, A mobile knowledge management decision support system for automatically conducting an electronic business, Knowledge-Based Systems 21 (7) (2008) 540–550.
[52] B. Yu, Z. Xu, C. Li, Latent semantic analysis for text categorization using neural network, Knowledge-Based Systems 21 (8) (2008) 900–904.
[53] E. Yudkowsky, Creating Friendly AI, Singularity Institute for Artificial Intelligence, Santa Clara, CA, 2001.
[54] A. Zellner, Generalizing the standard product rule of probability theory and Bayes's theorem, Journal of Econometrics 138 (1) (2007) 14–23.
[55] G. Zhang, J. Lu, T. Dillon, Decentralized multi-objective bilevel decision making with fuzzy demands, Knowledge-Based Systems 20 (5) (2007) 495–507.
[56] X.-D. Zhu, Z.-Q. Huang, Conceptual modeling rules extracting for data streams, Knowledge-Based Systems 21 (8) (2008) 934–940.