Hierarchical Modelling of Industrial System Reliability with Probabilistic Logic

Kamil Dedecius[1] and Pavel Ettler[2]
[1]*Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic*
Pod Vodárenskou věží 4, 182 08 Prague, Czech Republic
[2]*COMPUREG Plzeň, s.r.o., Nádražní 18, 306 34 Plzeň, Czech Republic*
dedecius@utia.cas.cz, ettler@compureg.cz

Keywords: Bayesian Analysis, Subjective Logic, Reliability.

Abstract: The use of Bayesian methods in dynamic assessment of system reliability is inevitably limited by computational difficulties arising from non-conjugate prior distributions. This contribution proposes an alternative framework, based on the combination of Bayesian methods and the subjective logic. The advantage of the former – consistent and exhaustive representation of available statistical knowledge, is extended by the latter, allowing computationally feasible combination of this knowledge at any level of the observed system using logic operations. The resulting methodology is currently under development in order to enlarge the capability of an intended novel industrial hierarchical condition monitoring system.

1 INTRODUCTION

We study the topic of hierarchical evaluation of system conditions (in terms of realiability and fault detection), based on the Bayesian paradigm. Its overwhelming influence in many areas of scientific data analyses naturally led to its more or less extensive use in the reliability theory. For example, (Hamada et al., 2008) give a comprehensive survey of Bayesian methods for assessment of components and systems reliability, including the techniques exploiting degradation data, assurance testing, regression models in reliability, Bayesian fault trees and network models.

The main difficulty with Bayesian modelling of dynamic systems is associated with the computation of posterior distributions (Gelman et al., 2003). If the prior distributions expressing the knowledge of the inferred variable of interest are not conjugate to the data model, the posteriors may take intractable forms. Computationally demanding approximations of these posteriors become inevitable, but many methods (e.g. Markov chain Monte Carlo) cannot reach a feasible result in the available timespan. This problem is yet accented when hierarchical models are involved.

The purpose of this contribution is twofold: first, it proposes a method for computationally tractable Bayesian inference of beta-distributed system reliability. Second, the subjective logic (Jøsang, 2001; Jøsang, 2008) provides, besides the gained tractability resulting from logical operations, the means for intelligible representation of reliability at any level of the monitored system, from individual components up to the system as a whole. Since the methods are somewhat different, we first present them concisely, giving examples along the way.

Elaboration of the methodology is not autotelic – combination of both approaches is being exploited within the international project aiming to develop a novel type of condition monitoring system and to test its achievements in the industrial environment. The paper presents work-in-progress results, whose assessment with respect to the existing methodologies are a part of the future research.

2 BAYES AND SUBJECTIVE LOGIC

The proposed framework consists of two methodologies – the Bayesian modelling, already well-established paradigm, and the subjective logic, a very recent probabilistic logic paradigm. While the Bayesian modelling allows theoretically consistent and versatile approach to reliability modelling[1] at any level of the system of interest, the subjective logic

[1]We understand the reliability to coincide with the probability that the studied system works well.

provides means for fast composition of conclusions among these levels. The result gives rise a to novel framework for system health monitoring, exploiting the intriguing aspects of both involved theories.

2.1 Principles of Bayesian Modelling

The principles of Bayesian modelling consist in specification of a probabilistic model for observable data y and a prior distribution for this model's unobservable parameters θ. In other words, we assume that the data y obey some distribution with a probability density function $f(y|\theta)$, or, under existence of observable exploratory variables, $f(y|x,\theta)$.

The prior distribution $\pi(\theta)$ statistically summarizes all *a priori* available information about the inferred parameter θ. It can be obtained from past measurements, from an expert or alternatively has a non-informative form. The prior pdf is updated by new information provided by x and y according to the Bayes' rule,

$$\pi(\theta|x,y) = \frac{f(y|x,\theta)\pi(\theta)}{\int f(y|x,\theta)\pi(\theta)d\theta}, \tag{1}$$

where the integral

$$q(y|x) = \int f(y|x,\theta)\pi(\theta)d\theta = \int f(y,\theta|x)d\theta \tag{2}$$

is taken over the space of θ. It serves as a normalizing constant, assuring that the resulting posterior distribution $\pi(\theta|x,y)$ is proper. By careful inspecting of (2) it is possible to notice, that $q(y|x)$ can play even more fundamental role than only the normalizing one. It is also a predictive density of y given x, which is obtained as an expected value $\mathbb{E}_\theta[f(y|x,\theta)]$, that is, over all admissible values of θ.

The resulting posterior pdf in (1), namely $\pi(\theta|x,y)$, involves both the prior information and the contribution from the observed data.

An important thought not always applicable property related to the Bayesian modelling is conjugacy. A model $f(y|x,\theta)$, being chosen from a suitable class of distributions (called the exponential family), guarantees the existence of the conjugate prior pdf $\pi(\theta)$. This ensures that the posterior pdf $\pi(\theta|x,y)$ lies in the same class of distributions as $\pi(\theta)$, the Bayesian update is analytically tractable, and the posterior can serve as the prior for a subsequent update when new data is obtained. This salient feature is clearly practical for real time modelling of dynamical systems. Then, denoting $t = 1,2,\ldots$ the time index,

$$\pi(\theta|x_{1:t},y_{1:t}) \propto f(y_t|x_t,\theta)\pi(\theta|x_{1:t-1},y_{1:t-1}), \tag{3}$$

where $y_{1:t} = \{y_1,\ldots,y_t\}$ (analogously for $x_{1:t}$) and $\propto$ stands for proportionality.

More on Bayesian modelling can be found, e.g., in (Gelman et al., 2003); its application to dynamic modelling is thoroughly treated in (Peterka, 1981).

Example 1. *Assume the regression model $y = x'\theta + \varepsilon$ where $y \in \mathbb{R}^n$ is a regressand (dependent or response variable), $x \in \mathbb{R}^{n\times m}$ is a regressor (independent explanatory variable), $\theta \in \mathbb{R}^m$ denotes regression coefficients and ε is a vector of independent identically distributed noise terms from a 0-centered normal distribution with a known variance, $\mathcal{N}(0,\sigma^2)$. Then, the probabilistic model $f(y|x,\theta)$ for y has the form*

$$y \sim \mathcal{N}(\theta^\intercal x, \sigma^2).$$

If the prior pdf $\pi(\theta)$ is normal, then the posterior pdf $\pi(\theta|x,y)$ is also normal. That is, the form of the prior pdf is preserved, prior pdf is conjugate to the model and dynamic setting (3) *is possible. The predictive distribution with a pdf $q(y|x)$ takes the form of a generalized Student's t distribution.*

2.1.1 Aspects of the Bayesian Approach

The Bayesian approach has many advantages, arising from its theoretical consistency and to a significant degree balancing its frequent computational burden. Some of the advantages, important for reliability modelling, are:

Uncertainty – an inherent aspect of information – is consistently involved in modelling. For instance, the posterior distribution of θ naturally expresses our uncertainty in terms of variance.

Asymptotics – while frequentist paradigm heavily relies on asymptotic results, the Bayesian approach does not. It yields results with any sample size. This is possible due to the concept of uncertainty.

Dynamic Modelling – an important aspect connected with conjugate priors. For instance, the Kalman filter or autoregression have their Bayesian interpretations. But there are many Bayesian models without equivalent non-Bayesian counterparts. Dynamic modelling inevitably calls for parameter tracking. The Bayesian paradigm allows its consistent treatment, e.g. (Dedecius et al., 2012).

Model Selection and Combination – it is easily possible to discriminate among several candidate models or even combine their results in order to further improve (e.g. stabilize) the whole modelling, see, e.g. (Raftery et al., 2010).

2.2 Principles of Subjective Logic

Subjective logic (Jøsang, 2001; Jøsang, 2008) is a novel probabilistic logic theory for treatment of uncertain propositions. Like other logic theories, it provides operations with these propositions, for instance logical negation (NOT), conjunction (AND), disjunction (OR), implications, modi ponens and tolens and many others, some of which do not have their counterparts in other logics, e.g. discounting.

Subjective opinions, binomial or Dirichlet, express beliefs about proposition under uncertainty. Since we are working with dichotomous variable – failure absent/present – we focus on binomial opinions. These opinions are represented by a quadruple

$$\omega = (b, d, u, a),$$

where b denotes the mass of belief in support of the variable of interest being true, d disbelief with the opposite meaning, u uncertainty, the mass complementing belief and disbelief, and finally a stands for the base rate, which is, to some degree, similar to prior information in Bayesian inference. The terms satisfy the following conditions:

$$b, d, u, a \in [0,1] \quad \text{and} \quad b + d + u = 1.$$

Opinions in subjective logic have a straight connection with binary logic and probability. For instance

$b = 1$ – equivalent to binary logic TRUE;

$d = 1$ – equivalent to binary logic FALSE;

$u = 0$ – equivalent to traditional probability.

The mean value of a binary proposition about X is simply the mass of belief plus a proportion of uncertainty assigned by the base rate,

$$\mathbb{E}[X] = b + au. \tag{4}$$

Obviously, binomial opinions live in a convex hull, concretely 2-simplex, an equilateral triangle with edges of unit norm.

Although the representation of the binomial opinions in terms of $\omega = (b, d, u, a)$ is intuitive, we will prefer a more convenient form of a beta distribution $B(\alpha, \beta)$ with parameters $\alpha, \beta > 0$. The beta distribution has a pdf of the form

$$f(p|\alpha,\beta) \propto p^{\alpha-1}(1-p)^{\beta-1}, \quad p \in [0,1], \tag{5}$$

where p denotes probability and $\propto$ is proportionality. The beta distribution of a binomial opinion maps the parameters as $\alpha = r + Wa$ and $\beta = s + Wa$, where W is a prior weight, usually set equal to 2, guaranteeing the uniform pdf with $a = 0.5$ and $r, s = 0$. The equivalent of (4) is the mean value of the beta-distributed variable,

$$\mathbb{E}[p] = \frac{\alpha}{\alpha+\beta} = \frac{r+Wa}{r+s+W}.$$

Bijective mapping between the binomial representation and the opinion has the following form:

$$\left(b = \frac{r}{\Xi}, \quad d = \frac{s}{\Xi}, \quad u = \frac{W}{\Xi}\right)$$
$$\updownarrow$$
$$\left(r = \frac{Wb}{u}, \quad s = \frac{Wd}{u}\right)$$

where $\Xi = W + r + s$. We omit the trivial case $u = 0$ and consider standard $W = 2$, preserving uniform distribution with $b, d = 0$.

Subjective Logic Conjunction and Disjunction (AND/OR). Given two independent opinions $\omega_x = (b_x, d_x, u_x, a_x)$ and $\omega_y = (b_y, d_y, u_y, a_y)$, the conjunction (AND) binomial opinion $\omega_{x\wedge y}$ is given by

$$b_{x\wedge y} = b_x b_y + \frac{(1-a_x)a_y b_x u_y + (1-a_y)a_x b_y u_x}{1-a_x a_y}$$
$$d_{x\wedge y} = d_x + d_y - d_x d_y$$
$$u_{x\wedge y} = u_x u_y + \frac{(1-a_y)b_x u_y + (1-a_x)b_y u_x}{1-a_x a_y}$$
$$a_{x\wedge y} = a_x a_y.$$

while the logical disjunction (OR) is given by

$$b_{x\vee y} = b_x + b_y - b_x b_y$$
$$d_{x\vee y} = d_x d_y + \frac{(1-a_y)a_x d_x u_y + (1-a_x)a_y d_y u_x}{a_x + a_y - a_x a_y}$$
$$u_{x\vee y} = u_x u_y + \frac{a_y u_y d_x + a_x u_x d_y}{a_x + a_y - a_x a_y}$$
$$a_{x\vee y} = a_x + a_y - a_x a_y.$$

The proofs can be found in (Jøsang and McAnally, 2005). For other logical operations refer, e.g., (Jøsang, 2008).

Example 2. *Suppose that we monitor a real system using the active probing technique, consisting in periodic sending short request messages and counting the number of received responses. A typical example is the Internet Control Message Protocol (ICMP) Echo Request, commonly known as* ping*, which should be answered by ICMP Echo Reply. Simple counting of the proportion of received and all requests is an example of application of the beta distribution. An illustration is depicted in Fig. 1.*

Example 3. *Let us have a system consisting of three blocks – A, B and C. Let A and B be interchangeable in the sense that it is satisfactory if at least one of them works well and let C be a critical system (Fig. 2). An example of such setting is a redundant disk array where A and B are two mirrored hard drives*

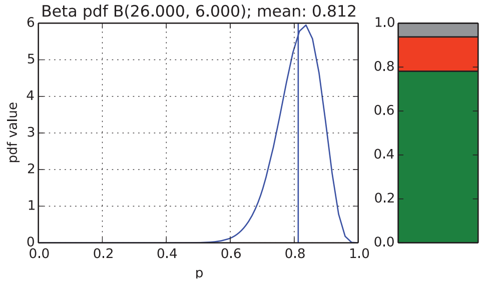

Figure 1: ICMP active probing example: $r+s=30$ requests, $r=25$ received replies, base rate $a=0.5$ and standard $W=2$. Left: beta pdf with depicted mean value. Right: proportions of belief (green), disbelief (red) and uncertainty (grey).

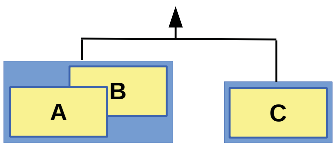

Figure 2: Example of a three-block network with redundancy.

and C is another hard drive. We are interested in the condition monitoring of the whole setting, that is, in

$$\omega^* = (\omega_A \vee \omega_B) \wedge \omega_C \tag{6}$$

For example, if

$$\begin{aligned}\omega_A &= (0.95, 0.02, 0.03, 0.5)\\ \omega_B &= (0.3, 0.6, 0.1, 0.5)\\ \omega_C &= (0.9, 0.05, 0.05, 0.5),\end{aligned}$$

we obtain $\omega^* = (0.89, 0.07, 0.05, 0.38)$. *This is also graphically depicted in Fig. 3 for the block A–B and in Fig. 4 for the whole system A–B–C.*

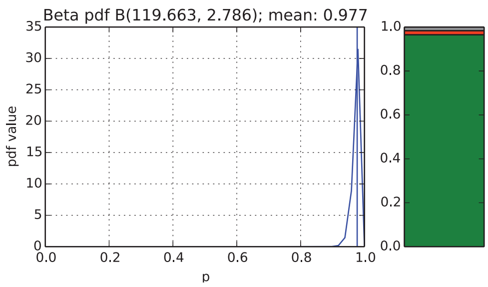

Figure 3: Subsystem A–B. The resulting $\omega_A \vee \omega_B = (0.97, 0.02, 0.02, 0.75)$ with (b, d, u) depicted in green, red and grey, respectively.

3 FUSION FRAMEWORK

Fusion of the theories of Bayesian modelling and subjective logic is relatively straightforward, due to the

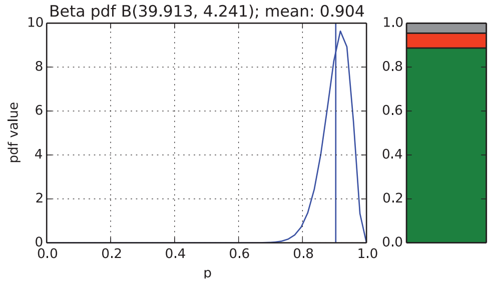

Figure 4: System A–B–C. The resulting $(\omega_A \vee \omega_B) \wedge \omega_C = (0.89, 0.07, 0.04, 0.38)$ with (b, d, u) depicted in green, red and grey, respectively.

beta representation of binomial opinions. Following the aforementioned principles of the Bayesian update, equations (1) or (3), we only need to know the information generating model. A natural choice is the generalized binomial distribution, since the beta distribution is conjugate to it and the resulting posterior will be beta as well. Remind that this is advantageous in dynamic settings.

The generalized binomial distribution has a pdf of the form

$$f(y|p) \propto p^r (1-p)^s, \tag{7}$$

where $p \in [0,1]$ is a parameter (probability) and $r, s > 0$ are statistics, e.g. r successes and s failures in $r+s$ trials. If r and s are positive integers, then one talks about the usual binomial distribution $\text{Binom}(r+s, r)$. Indeed, the normalizing term then does not take the form of the commonly known binomial coefficient in the generalized form. Clearly, (7) is conjugate to (5), the quantities r, s coincide and Wa in (5) is absorbed to the normalizing term.

The Bayesian update with newly obtained r_t and s_t takes the form

$$\pi(p|r_{1:t}, s_{1:t}) \propto f(r_t, s_t|p)\pi(p|r_{1:t-1}, s_{1:t-1}). \tag{8}$$

That is, the posterior pdf is fully characterized by statistics

$$r_{1:t} = r_{1:t-1} + r_t \tag{9}$$

$$s_{1:t} = s_{1:t-1} + s_t. \tag{10}$$

3.1 Accounting for Variability

The ordinary Bayes' rule (1) assumes invariability of the inferred parameters. A direct exploitation of the prior and model for varying parameters inevitably spoils the inference. A simple way around this issue is the use of forgetting. A thorough overview of the forgetting theory state of the art can be found in (Dedecius et al., 2012). Here, we propose to use the most simple approach – the exponential forgetting, introduced in the Bayesian setting in (Peterka, 1981). This means flattening of the prior pdf before new data is incorporated into it,

$$\pi(p|r_{1:t-1}, s_{1:t-1}) \leftarrow \pi(p|r_{1:t-1}, s_{1:t-1})^{\lambda}, \qquad (11)$$

where $\lambda \in (0,1)$ is the forgetting factor, usually not lower than 0.95. Obviously,

$$r_{1:t-1} \leftarrow \lambda r_{1:t-1}$$
$$s_{1:t-1} \leftarrow \lambda s_{1:t-1}$$

Some selected pairs of values of λ and the number d of effective samples are depicted in Table 1.

Table 1: Number d of effective samples given forgetting factor λ.

λ	0.999	0.998	0.995	0.99	0.98	0.95
d	1000	500	200	100	50	20

3.2 Sources of Opinion Information

The model (7) can be viewed in the scope of the classification theory: the information, generated by the underlaying process, belongs either to the class “success” (good operating conditions) or to the class “failure”. Furthermore, this classification can be hard (the information pertains to only one class at a time), or it can be soft (the information may belong to both classes to some extent). The former case is typical for strictly dichotomous information (the system is either working or not). The latter applies when there is some uncertainty present (it works but not perfectly).

There are several possible sources of information (data) for update of opinions. Below, we list our initial results. However, this topic still deserves considerable amount of research effort.

3.3 Hard Classification

Strictly binomial data naturally arise in many situations, where only success or failure are observed, e.g. monitored subsystem working/not working. A typical example is the basic active probing technique illustrated above. Then, the Bayesian update (3) simply incorporates direct counts of successes $r_t \in \mathbb{Z}_+$ and failures $s_t \in \mathbb{Z}_+$ into the prior distribution via equations (10).

Often, it is possible to hard-classify with respect to some preset threshold. In Section 4.2, we give an illustration of this. The measurements of a noise-corrupted signal are classified based on the signal-to-noise ratio (SNR) resulting either from good or bad (system) conditions. This dichotomous classification exploits a criterion set by an expert.

3.4 Soft Classification

Here the situation is much more complicated, as we deal with the need of extraction of information in favor of both classes. In the scope of the Bayesian analysis, we propose to exploit the predictive pdf $q(y|x)$, equation (2). It is a natural source of information, measuring the fit of the data y with respect to the a posteriori available information about the model and its parameters. For assigning the data to the classes, we can compare their localisation with respect to a suitable statistics, e.g. the mean value, median or other. A popular approach independent of our framework is to divide the support of $q(y|x)$ into interval by multiplies of the standard deviation σ starting from the mean μ. For instance, the $1-3\sigma$ intervals are given by $\mu \pm \sigma, \mu \pm 2\sigma, \mu \pm 3\sigma$, minus the mutual intersections. These intervals are assigned probabilities, inherited as r_t and $s_t = 1 - r_t$ for the Bayesian update (3). The probabilities can be obtained as the proportion of data within each interval. Roughly, given by the Chebyshev’s inequality

$$\Pr(|Y-\mu| \geq c\sigma) \leq c^{-2}, \qquad c \in \mathbb{R}^+.$$

Precisely, knowledge of the functional form of the pdf allows determination of exact proportions from the distribution function $Q(x)$,

$$\Pr(|Y-\mu| \leq c\sigma) = \int_{\mu-c\sigma}^{\mu+c\sigma} q(y|x)dy$$
$$= Q(\mu+c\sigma) - Q(\mu-c\sigma).$$

Example 4. *Let us consider specification of* (r_t, s_t) *for the Bayesian update of the beta distribution according to* (8). *If the assignment is based on the assumption of normality of* $y|x \sim \mathcal{N}(\mu, \sigma^2)$, *then* $r_t = 1 - s_t$ *equivalent to the relative amount of data within intervals is*

$r_t = 0.68$ *for* $y \in (\mu-\sigma, \mu+\sigma)$;
$r_t = 0.27$ *for* $y \in (\mu-2\sigma, \mu+2\sigma) \setminus (\mu-\sigma, \mu+\sigma)$;
$r_t = 0.04$ *for* $y \in (\mu-3\sigma, \mu+3\sigma) \setminus (\mu-2\sigma, \mu+2\sigma)$;
$r_t = 0.01$ *elsewhere.*

The above-given values equivalently come from the $3\sigma^2$ *rule.*

4 EXAMPLES

Two examples are given below. The first one depicts the evolution of the binomial opinion updated by several data. The latter considers modelling with real data.

4.1 Bayesian Update

This example considers the Bayesian update of the binomial opinion represented by the beta pdf. It is initialized with $(b,d,u,a) = (0.4,0.3,0.3,0.5)$. Each of the six subsequent updates is based on 10 measurements with 5, 2, 5, 7, 9 and 9 successes. The scheme corresponds to hard classification (Section 3.3) with binomial data, typical, e.g., for the active probing monitoring.

The evolution of the beta pdf and the opinions is depicted in Figure 5 (green is belief, red disbelief and grey uncertainty). Consistently with our expectation, the gradually increasing kurtosis of the beta pdf is connected with the diminishing opinion uncertainty.

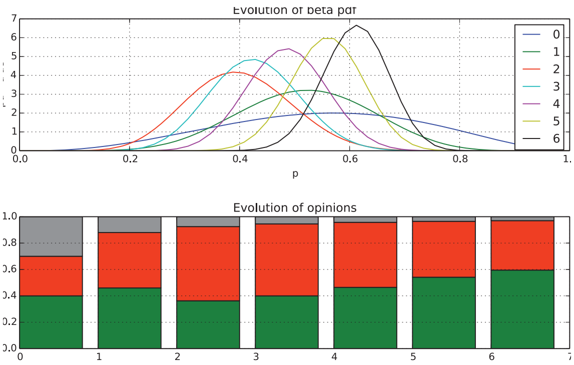

Figure 5: Bayesian update: Evolution of the beta pdf and opinion during 6 updates.

4.2 Example from Metal Processing Industry

This example relates to the operating conditions of a cold rolling mill. The output thickness of a processed metal strip is one of the key measures of the product quality. Its deviation h_2 from the nominal value is essential for the thickness control. If h_2 is not measured correctly, be it due to the sensor failure or problems with the rolled strip, the control may potentially result in deterioration of the final product quality.

Figure 6 shows an example of a spurious thickness measurement, affected by a dirt on the strip surface. The dirt implies the superposition of an additive noise due to the jitter of sensor's measuring tips (top, in blue). Assuming normality of the additive noise, the signal (h_2) can be filtered using a normal first order autoregressive model AR(1) of the form

$$h_{2,t} \sim \mathcal{N}\left(\begin{bmatrix} h_{2,t-1} \\ 1 \end{bmatrix}^\intercal \begin{bmatrix} \beta_{0,t} \\ \beta_{1,t} \end{bmatrix}, \sigma_t^2 \right), \qquad t = 1,2,\ldots$$

where the scalars $\beta_{0,t}$ and $\beta_{1,t}$ are the regression coefficients (the slope and intercept) and σ_t^2 denotes the zero-mean noise variance. The parameters $\beta_{0,t}, \beta_{1,t}$ and $\sigma^{2,t}$ are estimated in the Bayesian framework with the normal inverse-gamma prior pdf in the Peterka's form $\mathcal{N}i\mathcal{G}(V_{t-1}, \nu_{t-1})$ with a symmetric positive definite information matrix $V_{t-1} \in \mathbb{R}^{3\times 3}$ and the scalar degrees of freedom $\nu_{t-1} \in \mathbb{R}_+$, see (Peterka, 1981). Their update (1) with exponential forgetting reads

$$V_t = \lambda V_{t-1} + \begin{bmatrix} h_{2,t} \\ h_{2,t-1} \\ 1 \end{bmatrix} \begin{bmatrix} h_{2,t} \\ h_{2,t-1} \\ 1 \end{bmatrix}^\intercal$$

$$\nu_t = \lambda \nu_{t-1} + 1.$$

The estimators of β_0, β_1 and σ^2 are given by

$$\begin{bmatrix} \widehat{\beta}_{0,t} \\ \widehat{\beta}_{1,t} \end{bmatrix} = V_{t,[2:3,2:3]}^{-1} V_{t,[2:3,1]} \tag{12}$$

and

$$\widehat{\sigma}_t^2 = \frac{1}{\nu} \left(V_{t,[1,1]} - \begin{bmatrix} \widehat{\beta}_{0,t} \\ \widehat{\beta}_{1,t} \end{bmatrix}^\intercal V_{t,[2:3,1]} \right), \tag{13}$$

where the indices of the type $[i:j,k]$ denote blocks on rows $i:j$ and column k. Thorough inspection of (12) and (13) reveals ordinary least squares estimators with the Tichonoff-type regularization, that is, the least-squares estimators in the Bayesian sense.

We exploit the hard classification approach based on the signal-to-noise ratio (SNR). The threshold value between normal and abnormal state is 10dB (determined by an expert). The value of SNR is directly used to update the beta distribution.

The prior pdf was initialized with V_0 with diagonal $[0.1, 0.01, 0.01]$ and zeros elsewhere and $\nu_0 = 3$ corresponding to a flat distribution. The variability of parameters and opinions is driven by the forgetting factors, set to 0.97 for the estimation of the autoregressive model and 0.95 for the beta updates, (11). Since these factors influence the reaction time, a method for their adaptive tuning would further improve the properties of the modelling and condition evaluation.

The results from the modelling and determination of h_2 conditions are depicted in Figure 6 as well as the original data. The first subfigure shows the evolution of h_2 (blue) and the filtered signal (red). Obviously, there exist two segments with significant superposed noise. The evolution of the estimates follows. The SNR values is in accordance with the expected behavior with the exception of the first circa 400 steps. We stress, that most of this period is a stabilization, i.e., transition from a very flat prior (i.e., high variance, low information) pdf to an informative one. The last subfigure depicts the evolution of the tripple (b,d,u) of the binomial opinion. Again, the beginning is very skeptical due to the flat prior. We see, that (up to the stabilization), the framework quite promptly reacted to the worsening of the signal quality. The loss of information from the noisy signal becomes particularly evident after $t \approx 1500$, where even the autoregressive model looses its filtering ability. Correspondingly, $d \to 1$.

Sometimes, this approach may be too stringent. Then, the update may be based on other (e.g. transformed) information. Also, the responsiveness may be tuned by the forgetting factor.

We remind, that the subjective logic opinion is intended to enter the logical operations for further analysis of the whole system.

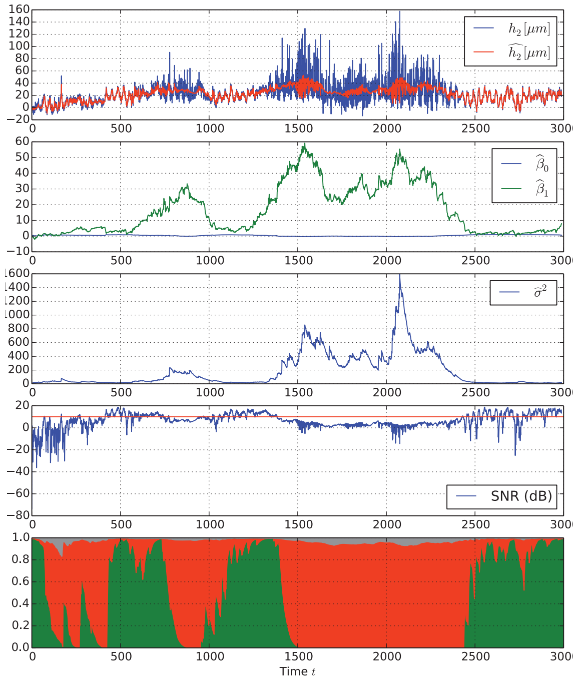

Figure 6: Evolution of condition of the output strip thickness deviation h_2. From top: strip thickness deviation h_2 (blue) and its filtered value $\widehat{h}_2$ (red); evolution of the regression coefficients estimates $\hat{\beta}_0$ and $\hat{\beta}_1$; evolution of the noise variance estimate $\widehat{\sigma}^2$; evolution of SNR with respect to the 10dB criterion; evolution of opinion (in the same colors as above).

5 CONCLUSION

The proposed novel framework, combining the Bayesian paradigm for information processing and the subjective logic for its combination and representation, provides intriguing methods for hierarchical modelling of a system reliability. While not dictating the particular form of combination, it allows to exploit the best of both theories where and when necessary. The approach is under its initial development and a lot needs to be done. First, the Bayesian update of the opinion beta pdf is fine up to the need of achieving binomial information from not necessarily binomial data. This point deserves a lot of focus yet.

ACKNOWLEDGEMENTS

The research project is supported by the grant MŠMT 7D12004 (E!7262 ProDisMon).

REFERENCES

Dedecius, K., Nagy, I., and Kárný, M. (2012). Parameter tracking with partial forgetting method. *International Journal of Adaptive Control and Signal Processing*, 26(1):1–12.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). *Bayesian Data Analysis, Second Edition*. Chapman & Hall/CRC.

Hamada, M. S., Wilson, A. G., Reese, C. S., and Martz, H. F. (2008). *Bayesian Reliability*.

Jøsang, A. (2001). A Logic for Uncertain Probabilities. *Int. J. Unc. Fuzz. Knowl. Based Syst.*, 09(03):279–311.

Jøsang, A. (2008). Conditional Reasoning with Subjective Logic. *Journal of Multiple-Valued Logic and Soft Computing*, 15(1):5–38.

Jøsang, A. and McAnally, D. (2005). Multiplication and comultiplication of beliefs. *International Journal of Approximate Reasoning*, 38(1):19–51.

Peterka, V. (1981). Bayesian approach to system identification In P. Eykhoff (Ed.) Trends and Progress in System Identification, Oxford, U.K.: Pergamon Press, 239–304.

Raftery, A., Kárný, M., and Ettler, P. (2010). Online Prediction Under Model Uncertainty via Dynamic Model Averaging: Application to a Cold Rolling Mill. *Technometrics*, 52(1):52–66.