On Evaluating Monitoring Design Effectiveness

Daniele Zonta, Carlo Cappello, Matteo Pozzi, Branko Glisic

ON EVALUATING MONITORING DESIGN EFFECTIVENESS

Daniele Zonta[1], Carlo Cappello[1], Matteo Pozzi[2], Branko Glisic[3]

[1] *Department of Civil, Environmental and Mechanical Engineering, University of Trento, via Mesiano 77, 38123 Trento, Italy*

[2] *Department Civil & Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213-3890, USA*

[3] *The Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08544, USA*

daniele.zonta@unitn.it

ABSTRACT

When designing a structure such as a bridge or a building, a civil engineer follows a well-established, rational procedure, whereby the performance of the design concept is predicted through structural analysis and quantitatively assessed with respect to the target performance. On the contrary, when an engineer designs a monitoring system, the approach is often heuristic with performance evaluation based on common sense or experience, rather than on quantitative analysis. In this paper, we describe a rational procedure for the design of monitoring systems, keeping in mind an analogy between structural and monitoring design. Whereas the structural design objective is to achieve stability with an appropriate level of safety, the object of monitoring is to acquire knowledge with an appropriate level of confidence. Herein, we illustrate the motoring design procedure with a simple practical example and discuss a possible metric for evaluating the performance of a design concept.

KEYWORDS : *Bayesian inference, Monitoring, Metrology, Error propagation, Cable stayed bridge*

INTRODUCTION

When designing a structure such as a bridge or a building, a civil engineer follows a well-established, rational procedure, whereby the performance of the design concept is predicted through structural analysis and quantitatively assessed with respect to the target performance. On the contrary when an engineer designs a monitoring system, the approach is often heuristic with performance evaluation based on common sense or experience, rather than on quantitative analysis. In this paper we describe a rational procedure for the design of monitoring systems, keeping in mind an analogy between structural and monitoring design (Table 1). The objective of structural design is to ensure stability under a given load with an appropriate safety factor. The structural design process includes: definition of design loads; calculation of stresses (structural demand) using a structural model; choice of a technological solution that offers the required strength (or capacity). The design is satisfactory if the strength is greater than given load (or when capacity is greater than demand). The performance of the structural system is evaluated by calculation of the probability of collapse. In structural health monitoring, we acquire data using sensors to understand the condition state of a structure. We can say that to monitor is to deduce the state of a system (the structure) using observations (the instrumental data) and assumptions as to the state of the structure (prior knowledge) and as to the relationship with the observations (model). In other words, whereas the structural design objective is to achieve stability with an appropriate level of safety, the object of monitoring is to acquire knowledge with an appropriate level of confidence. The monitoring design process includes: definition of the target performance of monitoring (for example the accuracy of knowledge); calculation of the required accuracy of instrumental data, using a model; choice of sensor technology. The design is satisfactory if knowledge accuracy is equal or better than the demand. In logical terms, structural health monitoring is formally identical to the metrology problem of indirect measurement, where the measurand is indirectly estimated based on observation of other physical quantities linked to the measurand. In analogy with the metrology problem, in this contribution we use error propagation technique, based on Bayesian logic [1,2], to judge a priori the effectiveness of a monitoring method. We demonstrate the approach in estimating the tension of a cable stay based on vibrational measurements

Table 1: Equivalence between structural design and monitoring system design.

	Structural Design	**Monitoring System Design**
Objective	Structural stability with appropriate safety	Knowledge of structural state with appropriate confidence
Design target	Load	State knowledge accuracy
Demand	Stress	Measurement accuracy demand
Model	Relationship between stress and load	Relationship between state parameters and measurements
Capacity	Resistance	Sensor accuracy
Limit state	Stress vs. Resistance	Measurement accuracy demand vs. sensor accuracy
Performance metric	Probability of failure	Probability of state misidentification

1 PROBLEM STATEMENT AND FORMULATION

From a logical standpoint, structural health monitoring is an inference problem where we attempt to gain information on the state of a structure based on sensor observations and the expected relationship between observations and state, normally encoded in a mechanical or heuristic model. Written in mathematical terms, the estimation of the state variables is formally identical to the metrology problem of indirect measurement, where the measurand is indirectly estimated based on observation of other physical quantities linked to the measurand. First of all, let us define the basic concepts of sensor, state observation and model that we will use in the rest of this paper.

Sensor: there are many uses of the term sensor; in this paper we will refer to a sensor in logical terms as suggested by Hall and McMullen [3]: a logical sensor is defined as ‘any device which functions as a source of information’.

Observation: in general we use objective information in our inference problem; in structural health monitoring, the raw data acquired by the sensors are observations.

State: an object which represents the condition of the system. The state of the system can be specified by a discrete variable or class (for example, a concrete beam can be cracked or not cracked), by a set of continuous parameters $\boldsymbol{\theta}$ (for example, Young modulus of concrete, the opening of a crack) or by a combination of the two.

Model: the way we assume the observations are correlated to the state. Typically, in structural engineering problems the model has a physical background (e.g.: a Finite Element or analytical model). Sometimes the model can be heuristic. Very often the model depends on some state parameters; the model can be probabilistic in the sense that the parameters involved are random, and represented by a distribution.

Going back to the definition introduced above, the basic monitoring problem is to infer the state S of the structure based on a set of instrumental observations, or measurements, $\mathbf{y}$. In general the state S is identified by set of variables which characterize the state. To start with, we assume that the state is defined by a set of continuous parameters $\boldsymbol{\theta}$, The joint distribution $p_{\boldsymbol{\theta}\mathbf{y}}(\boldsymbol{\theta},\mathbf{y})$ of $\boldsymbol{\theta}$ and $\mathbf{y}$ reads either in any of the two following expressions:

$$p_{\boldsymbol{\theta}\mathbf{y}}(\boldsymbol{\theta},\mathbf{y})=p(\boldsymbol{\theta}|\mathbf{y})\cdot p(\mathbf{y})=p(\mathbf{y}|\boldsymbol{\theta})\cdot p(\boldsymbol{\theta}). \tag{1}$$

We can we can rewrite Equation (1) in the classical form:

$$p(\boldsymbol{\theta}|\mathbf{y})=\frac{p_{\boldsymbol{\theta}\mathbf{y}}(\boldsymbol{\theta},\mathbf{y})}{p(\mathbf{y})}=\frac{p(\mathbf{y}|\boldsymbol{\theta})\cdot p(\boldsymbol{\theta})}{p(\mathbf{y})}, \tag{2}$$

which shows that the posterior distribution $p(\boldsymbol{\theta}|\mathbf{y})$ of the parameters is proportional to the likelihood $p(\mathbf{y}|\boldsymbol{\theta})$ and to the prior distribution $p(\boldsymbol{\theta})$ of the parameters. The evidence can be seen as a normalization constant, which must be calculated by integrating over the parameter space:

$$p(\mathbf{y})=\int_{D\boldsymbol{\theta}} p_{\boldsymbol{\theta}\mathbf{y}}(\boldsymbol{\theta},\mathbf{y})\cdot d\boldsymbol{\theta}=\int_{D\boldsymbol{\theta}} p(\mathbf{y}\,|\,\boldsymbol{\theta})\cdot p(\boldsymbol{\theta})\cdot d\boldsymbol{\theta}, \tag{3}$$

where $D\boldsymbol{\theta}$ means the domain of the parameters $\boldsymbol{\theta}$.

Equation (2) can be generally seen as an equation in the joint space of the parameters $\boldsymbol{\theta}$ and of the observations $\mathbf{y}$. So for example $p(\boldsymbol{\theta}|\mathbf{y})$ is a function of $\mathbf{y}$ and $\boldsymbol{\theta}$, which returns for any individual value of parameters and any realization the posterior mass density of the parameters.

Normally, we use Bayes' theorem to update the parameter distribution after a specific realization $\mathbf{y}_0$ of $\mathbf{y}$, which is to say, in simpler words, after acquiring a specific dataset $\mathbf{y}_0$ from the monitoring system. In this case the posterior distribution $p(\boldsymbol{\theta}|\mathbf{y}_0)$ simply writes:

$$p(\boldsymbol{\theta}|\mathbf{y}_0)=\frac{p(\mathbf{y}_0|\boldsymbol{\theta})\cdot p(\boldsymbol{\theta})}{p(\mathbf{y}_0)}. \tag{4}$$

Observe that here, unlike in Equation (1), $\mathbf{y}_0$ is a determinate vector, thus posterior $p(\boldsymbol{\theta}|\mathbf{y}_0)$ and likelihood $p(\mathbf{y}_0|\boldsymbol{\theta})$ are now mere functions of the parameters, while $p(\mathbf{y}_0)$ is a constant. In other words, the posterior as in Equation (4) is simply a section on the joint distribution $p_{\boldsymbol{\theta}\mathbf{y}}(\boldsymbol{\theta},\mathbf{y})$ for $\mathbf{y}=\mathbf{y}_0$, being the evidence $p(\mathbf{y}_0)$ a normalization factor which makes the integral of the posterior distribution equal to 1.

Equation (4) is the classical formulation to update the state parameter after acquiring monitoring data. However this is not our problem. Here we want to infer a priori features of the posterior distribution in the assumption that we are going to acquire some data, but we do not have these data yet. We refer to this analysis as *prior estimation of the posterior* as to *pre-posterior analysis*.

2 SINGLE PARAMETER CASE

In general the distribution of the parameter θ, posterior to an observation $\mathbf{y}$, has an expected value $\mu_{\theta|\mathrm{y}}$ and variance $\sigma^2_{\theta|\mathrm{y}}$, both depending on the observation $\mathbf{y}$.

$$\mu_{\theta|\mathrm{y}}(\mathbf{y})=\frac{\int_{D\theta} p_{\boldsymbol{\theta}\mathbf{y}}(\theta,\mathbf{y})\cdot\theta\cdot d\theta}{\int_{D\theta} p_{\boldsymbol{\theta}\mathbf{y}}(\theta,\mathbf{y})\cdot d\theta}=\frac{\int_{D\theta} p_{\boldsymbol{\theta}\mathbf{y}}(\theta,\mathbf{y})\cdot\theta\cdot d\theta}{p(\mathbf{y})}; \tag{5}$$

$$\sigma^2_{\theta|\mathrm{y}}(\mathbf{y})=\frac{\int_{D\theta} p_{\boldsymbol{\theta}\mathbf{y}}(\theta,\mathbf{y})\cdot\left(\theta-\mu_{\theta|\mathrm{y}}(\mathbf{y})\right)^2\cdot d\theta}{p(\mathbf{y})}. \tag{6}$$

Since the observation is unknown a priori, these quantities can be seen as function of the observation $\mathbf{y}$. Thus the posterior variance can be seen as a random variable with expected value

$$\sigma^2_{\theta,pp} = E\left[\sigma^2_{\theta|\mathbf{y}}(\mathbf{y})\right] = \int_{\mathrm{D}\mathbf{y}} \sigma^2_{\theta|\mathbf{y}}(\mathbf{y}) \cdot \mathrm{p}(\mathbf{y}) \cdot \mathrm{d}\mathbf{y} = \int_{\mathrm{D}(\theta,\mathbf{y})} \mathrm{p}_{\theta\mathbf{y}}(\theta,\mathbf{y}) \cdot \left(\theta - \mu_{\theta|\mathbf{y}}(\mathbf{y})\right)^2 \cdot \mathrm{d}\theta \cdot \mathrm{d}\mathbf{y}\,. \tag{7}$$

In practice quantity $\sigma^2_{\theta,pp}$ represents what we expect variance to be after monitoring is performed. We will refer to this quantity as *prior estimation of the posterior variance*, or *pre-posterior variance* in short. Observe that the pre-posterior variance is independent on any observation, so being at all affect a quantity a priori. Compared with the 'proper' pior variance σ^2_θ, the pre-posterior gives an idea on how useful is the monitoring to gain information on the unknown parameter: if the pre-posterior variance is much smaller than the prior, this means that we expect our monitoring method to improve our knowledge of the parameter. On the contrary, when the pre-posterior and prior are equal or similar, this means that we do not expect monitoring to brings any significant knowledge improvement, thus in this case our monitoring method is useless. A candidate indicator of the expected effectiveness is:

$$\eta = \sqrt{\frac{\sigma^2_\theta}{\sigma^2_{\theta,pp}}} = \frac{\sigma_\theta}{\sigma_{\theta,pp}}\,. \tag{8}$$

In simple words, when the effectiveness is, say, 10 this means that we expect monitoring to estimate the parameter ten time better than t. An alternate way to express the same quantity is to transform parameter θ to a normalized parameter φ with standard Gaussian prior distribution (this is commonly known as Hasofer-Lind transformation [4] in the reliability community):

$$\varphi = \frac{\theta - \mu_\theta}{\sigma_\theta}\,. \tag{9}$$

Since the prior variance σ^2_φ of the standard parameter is one, Equation (8) simply reduces to:

$$\eta = \sqrt{\frac{1}{\sigma^2_{\varphi,pp}}} = \frac{1}{\sigma_{\varphi,pp}}, \tag{10}$$

meaning that the effectiveness of the monitoring system is represented by the inverse pre-posterior variance of the parameter, normalized to variable with standard Gaussian prior distribution.

3 MULTI PARAMETER CASE

At this point we want to generalize the concepts introduced in the previous section for single parameter to the a case where monitoring address the knowledge of a set of parameters $\boldsymbol{\theta}$. To start, assume that the prior distribution is a multivariate Gaussian with mean vector $\boldsymbol{\mu}_{\boldsymbol{\theta}}$ and covariance $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$. Using Bayes theorem, we can calculate the posterior distribution its expected value vector $\boldsymbol{\mu}_{\boldsymbol{\theta}|\mathbf{y}}$ and a covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}|\mathbf{y}}$, both function of the observation $\mathbf{y}$. In general, there may be two types of parameter: state parameters that actually represent the state of the system and auxiliary parameters that serve to define the correlation between observations into the probabilistic model used to calculate the likelihood. Here we assume that the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}|\mathbf{y}}$ has been marginalized with respect to the state parameters. In this case Equation (7) generalize to

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta},pp} = E\left[\boldsymbol{\Sigma}_{\boldsymbol{\theta}|\mathbf{y}}(\mathbf{y})\right] = \int_{\mathrm{D}\mathbf{y}} \boldsymbol{\Sigma}_{\boldsymbol{\theta}|\mathbf{y}}(\mathbf{y}) \cdot \mathrm{p}(\mathbf{y}) \cdot \mathrm{d}\mathbf{y}\,. \tag{11}$$

where matrix $\mathbf{\Sigma}_{\mathbf{\theta},pp}$ is here labelled *pre-posterior covariance matrix*. Here again we can judge the effectivness of our monitoring method by comparing the preposterior with the proper prior covariance matrix $\mathbf{\Sigma}_{\mathbf{\theta}}$.

Similar to the single parameter case, we might want to transform the original variables to a standard normal set $\boldsymbol{\varphi}$, which prior distribution has mean value $\boldsymbol{\mu}_{\boldsymbol{\varphi}} = \mathbf{0}$ and covariance $\mathbf{\Sigma}_{\boldsymbol{\varphi}} = \mathbf{I}$, being $\mathbf{I}$ the identity matrix. If the parameters are mutually uncorrelated a priori (i.e., the prior covariance is diagonal):

$$\varphi_i = \frac{\theta_i - \mu_{\theta i}}{\sigma_{\theta i}}. \tag{12}$$

while in the most general case we must transform a multivariate Gaussian

$$p(\mathbf{\theta}) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma_{\mathbf{\theta}})}} \exp\left[-\frac{1}{2}(\mathbf{\theta} - \mu_{\mathbf{\theta}})^T \Sigma_{\mathbf{\theta}}^{-1} (\mathbf{\theta} - \mu_{\mathbf{\theta}})\right], \tag{13}$$

to a standard multivariate

$$p(\boldsymbol{\varphi}) = \frac{1}{\sqrt{(2\pi)^n}} \exp\left[-\frac{1}{2}\boldsymbol{\varphi}^T \boldsymbol{\varphi}\right]. \tag{14}$$

By comparison of the two:

$$\boldsymbol{\varphi} = \Sigma_{\mathbf{\theta}}^{-1/2} (\mathbf{\theta} - \mu_{\mathbf{\theta}}), \tag{15}$$

where $\Sigma_{\mathbf{\theta}}^{-1/2}$ is the Cholesky decomposition of the prior covariance. To judge the effectiveness of the monitoring strategy, it is usefull to introduce a metric encoding the level of uncertainty of the posterior parameters. This metric must be invariant with respect to a linear transformation in the parameters space, thus trace or determinant in the covariance matrix are both possible candidate. Between the two, trace looks more appropriate, because it reduces to zero only if all the parameters are deterministically known, while determinant zeroes when any individual parameter is deterministically known. Thus, the extension of Equation (10) to the multi parameter case reads:

$$\eta = \left(\mathrm{tr}\left(\mathbf{\Sigma}_{\boldsymbol{\varphi},pp}\right)\right)^{-1/2}. \tag{16}$$

4 EXAMPLE

Here we show with a simple example how judging monitoring effectiveness in a single parameter case. Suppose we want to measure the tension force T of a cable-stay based on its first natural frequency f. To keep the example clear and simple to the reader, let's make some simplifying assumptions. Particularly, assume that sag and bending stiffness are negligible, so that the relationship between f and T is that provided by classical string theory (see for example [5,6]):

$$f^2 = \frac{T}{4mL^2}, \tag{17}$$

where L is the cable length, m is its linear mass. Further assume that cable length and linear mass are deterministically known and equal (let make them in round figures) to m=100 kg and L=100 m, respectively. These assumptions basically imply that all the uncertainty in T descends from the measurement error of f (which is to say that once the frequency is measured with no error the tension is also exactly known). This is obviously not necessarily the case in the real world, where in fact uncertainties also stem from model, physical parameters, uniformity and boundary conditions of the cable and environmental effects. Yet, to keep the example clearer, we will accept these simplifications.

Although we don't know it exactly, even before carrying out the measurement we always have some idea of what the tension might be. This is to say, in the Bayesian jargon, that we have a prior knowledge of T. To start, assume that the we know the tension to be around 7000 kN, but it could be equally well a ten percent more or less. We can formalize this initial knowledge stating that the prior distribution of T is Gaussian with mean value μ_T=7000 kN and standard deviation σ_T =700 kN. For convenience, this prior distribution p(T) is represented in Figure 1c in the form of a 2-dimensional function of T and f, although it is understood that it is independent on the measurement f.

We have to decide whether carrying out a vibrational measurement make sense or not. To do so, we can estimate how much we expect to improve our prior knowledge of the tension, once performed a vibrational test. We know that monitoring allows to identify the tension, with an uncertainty that depends on the measurement method. Assume that the standard deviation of the frequency measurement is σ_f =0.01 Hz, and that this uncertainty is independent on the frequency f itself. Given a determinate value of T, the expected value of f is, overturning Equation (17):

$$E\left(f \mid T\right)=\sqrt{\frac{T}{4mL^2}} \tag{18}$$

Further assuming, for simplicity, that the likelihood is Gaussian, its expression is:

$$\mathrm{p}\left(f \mid T\right)=\mathrm{Norm}\left(\sqrt{\frac{T}{4mL^2}},\sigma_f^2;f\right). \tag{19}$$

We can interpret Equation (19) as a 2-dimensional function of T and f, which graph is shown in figure 1b. Given a specific measurement f_0, the cross-section of this 2-dimensional function is the likelihood of T. Observe that because the relationship between T and f is nonlinear, the standard deviation $\sigma_{f|T}$ of T does depend on the specific measurement f_0.

Now, the product of prior p(T) and likelihood p(f|T) for any possible pair of f and T produces the joint distribution $p_{Tf}(T,f)$ shown in Figure 1a. A section of this joint distribution for an individual observation f represent the unnormalized posterior with integral equal to the evidence p(f). Figure 2a shows how evidence changes with the observation. Known the measurement f, we can also calculate the posterior variance $\sigma_{T|f}^2$ of T. This is shown in figure 2b as a function of f. However, the problem is that before carrying out the test, we don't know the measurement and therefore we don't have the exact value of the posterior variance. However, we can always plot the variance of T for any possible individual value of f. This is shown in figure 2b as a function of f.

Now, we can calculate the pre-posterior variance through Equation (7):

$$\sigma_{T,pp}^2=\int_0^{+\infty}\sigma_{T|f}^2\left(f\right)\cdot\mathrm{p}\left(f\right)\cdot\mathrm{d}f\ . \tag{20}$$

A possible way to solve this integral is: (i) drawn a set of observations **f** using pseudo-random algorithms; (ii) calculate the mean $\mu_{T|\mathbf{f}}$ and the standard deviation $\sigma_{T|\mathbf{f}}$ of the posterior distribution corresponding to the set **f** (e.g., using Metropolis-Hastings algorithm [7]); (iii) calculate the posterior variance; then (iv) repeat step i to iii; and finally (v) calculate the pre-posterior variance. Solved numerically, the integral provides the prepisterior variance $\sigma_{T,pp}^2$ =10975 kN^2, corresponding to a pre-posterior standard deviation of $\sigma_{T,pp}$ =104.76 kN. (Observe that the square root of the expected value of variance does not correspond to the expected value of standard deviation). The pre-posterior standard deviation can be compared with the prior σ_T =700 kN. This provides an effectiveness index of η=6.68, meaning that in this case monitoring will significantly improve the quality of our knowledge.

It is important to observe how the effectiveness of monitoring depends on the prior knowledge. As a second example, let's repeat the same calculation assuming the tension is know a priori with a better approximation, say σ_T =140 kN and a measurement uncertainty is higher, say 0.05 Hz. In this case, likelihood, joint distribution and evidence are those shown in Figure 3. Following the same path, we obtain a pre-posterior standard deviation of $\sigma_{T,pp}$ =135.32 kN and an effectiveness index of η=1.03, meaning that in this second case monitoring does not improuve our knowledge of *T*, thus it's virtually useless.

Classical error propagation theory [8] can be seen as an approximated way to estimate the pre-posterior variance. The idea is to calculate the posterior variance for the expected observation value:

$$\mu_f \cong \sqrt{\frac{\mu_T}{4mL^2}} \tag{21}$$

Assuming Gaussian distributions, the posterior variance can be estimated with the inverse variance rule:

$$\sigma_{T|f}^{-2} = \sigma_{f|T}^{-2} + \sigma_T^{-2} \tag{22}$$

where the likelihood variance I approximated with a first order error propagation ther:

$$\sigma_{f|T} \cong \left[\frac{\mathrm{d}T}{\mathrm{d}f}\right]_{\mu_T} \sigma_f = 8mL^2\mu_f\sigma_f \tag{23}$$

Referring to the first example, we obtain in our case μ_f =1.32 Hz, $\sigma_{f|T}$ =105.8 Hz and eventually $\sigma_{T|f}$ =104.61 Hz, which is very close to the 'exact' pre-posterior estimated.

CONCLUSION

We proposed a rational framework to estimate monitoring system accuracy in prior condition. The presented approach enables the estimator to obtain the 'pre-posterior', which is an estimate of the posterior distribution, calculated with the prior distribution of parameters $\boldsymbol{\theta}$. Covariance matrix of posterior distribution contains the information regarding the accuracy of estimation, however, its components have values that depends on dimension. Therefore, we proposed to apply the Hasofer-Lind transformation in order to obtain a set of dimensionless parameters $\boldsymbol{\varphi}$. The Hasofer-Lind transformation has to be carried out using the mean and covariance matrix of prior distribution. We defined the trace of covariance matrix of $\boldsymbol{\varphi}$ as an indicator of the posterior uncertainty. The value of the trace calculated in this way can be used to compare different monitoring systems as it is a dimensionless parameter scaled on the prior distribution.

REFERENCES

[1] D. Zonta, B. Glisic and S. Adriaenssens. Value of information: impact of monitoring on decision-making, *Structural Control and Health Monitoring*, 21:1043–1056, 2014.

[2] D. Sivia and J. Skilling, *Data Analysis: A Bayesian Tutorial*, Cambdrige University Press, 2006.

[3] D. L. Hall and S. A. H. McMullen. *Mathematical Techniques in Multisensor Data Fusion*, Artech House, 2004.

[4] H.O. Madsen, S. Krenk, N.C. Lind. *Methods of Structural Safety*, Dover Publications, 2006.

[5] B. H. Kim and T. Park. Estimation of cable tension force using the frequency-based system identification method, *Journal of Sound and Vibration* 304, 2007.

[6] D. Zonta, F. Bruschetta, R. Zandonini, M. Pozzi, M. L. Wang, B. Glisic, D. Inaudi, D. Posenato, Y. Zhao. Sensor fusion on structural monitoring data analysis: application to a cable-stayed bridge, *Key Engineering Materials*, 569-570:812-819, 2013.

[7] W. M. Bolstad, *Understanding Computational Bayesian Statistics*, Wiley, 2010.

[8] L. Kirkup and R. Frenkel, *An Introduction to Uncertainty in Measurement*, Cambridge University Press, 2010.

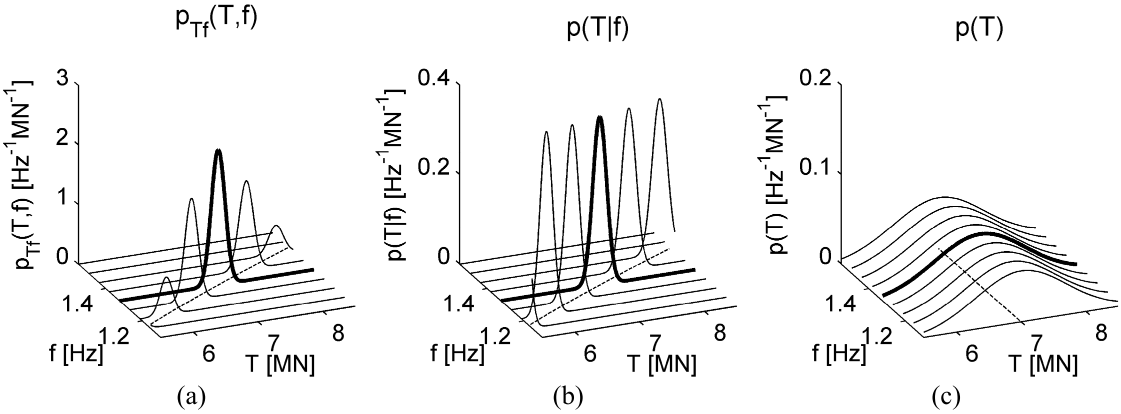

Figure 1: Joint distribution (a), likelihood function (b) and prior distribution (c), of example 1

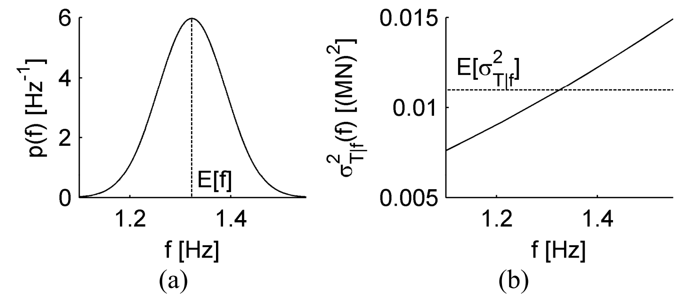

Figure 2: Evidence (a) and variance (b), of example 1

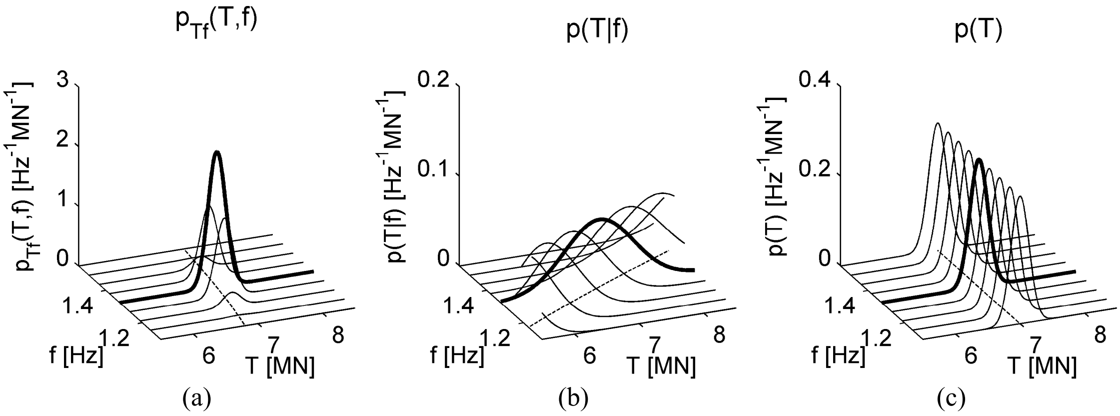

Figure 3: Joint distribution (a), likelihood function (b) and prior distribution (c), of example 2

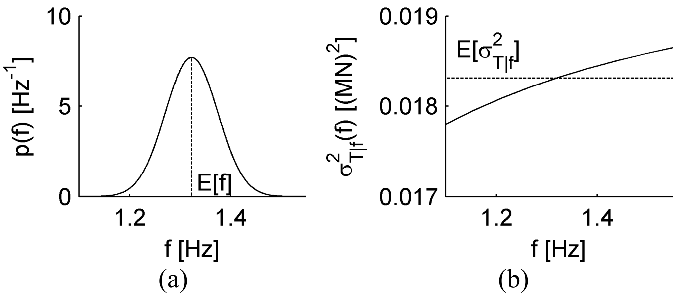

Figure 4: Evidence (a) and variance (b), of example 2