

# Kent Academic Repository

## Full text document (pdf)

### Citation for published version

Eves, Christophe and Wang, Xue and Walker, Stephen G. (2014) Bayesian Information for Sensors. Quality and Reliability Engineering International, 31 (8). pp. 1717-1724. ISSN 0748-8017.

### DOI

<https://doi.org/10.1002/qre.1704>

### Link to record in KAR

<https://kar.kent.ac.uk/58176/>

### Document Version

Pre-print

#### Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

#### Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

#### Enquiries

For any further enquiries regarding the licence status of this document, please contact:

[researchsupport@kent.ac.uk](mailto:researchsupport@kent.ac.uk)

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

# Bayesian Information for Sensors

Christopher Eves and Xue Wang

*School of Mathematics, Statistics and Actuarial Science,  
University of Kent at Canterbury, Canterbury, Kent, UK, CT2 7NF*

Stephen G. Walker

*Department of Mathematics and Division of Statistics and Scientific Computation,  
University of Texas at Austin, Texas, USA 78705*

**Abstract** An engineering company<sup>1</sup> manufacturing high precision sensors had accumulated huge historical databases of information on a type of sensors which had been tested. The aim of the company was not to use this historical data to improve estimation of future individual sensor parameters, but rather to use it to reduce the number of measurements needed per sensor, guaranteeing a required level of accuracy. In the paper we show how this can be done, using Bayesian ideas, and introduce the novel theory for linear regression models which determines how the reduction in individual sensor measurements can be achieved.

Specifically, for estimating parameters of closely related sensors, an estimate can be thought of as comprising a global component, i.e. the mean of all the sensors, and a local component, which is a shift from the mean. The historical data can, in a Bayesian framework, provide the global component and hence all that is needed from an individual sensor is the local component. In non-Bayesian estimation methods, both components are required and hence many measurements are needed. On the other hand, with Bayesian methods, only the local fit is needed and hence fewer measurements per sensor are required. We provide the supporting theory and demonstrate on a real-life application with real data.

**Key words:** Bayesian Model Fitting; Estimation Accuracy; Measurement Reduction; Resonator Density Sensor;

---

<sup>1</sup>The work in this paper was supported by a Knowledge Transfer Project funded by the Technology Strategy Board, UK, grant no. KTP008285. For confidential reasons the name and type of sensor being manufactured can not be disclosed.

**1. Introduction.** Sensors play a crucial role in modern industrial systems as well as consumer devices. The sensors of interest in this paper are associated with industrial control systems where high precision is the dominant concern. The manufacturing of such sensors unavoidably involves variations, such as material variances. It is therefore a necessary task to characterize each sensor via parameter estimation to achieve a high precision or accuracy. As Thornton (1999) noted, even for the most robust product design, manufacturing variations can often cause problems on product quality. In order to maximize production quality, control or monitoring of the key characteristics or critical parameters is essential.

For a sensor, in particular one with high precision, the result of variations in manufacturing is that the output variable will deviate from the true value. For example, the output of a pressure sensor in a pipeline may change undesirably under a varying environmental temperature, even if the pipeline pressure does not change. For example, Baker (2004) discussed the importance of understanding, controlling and reducing the variability for high-quality measuring devices. He noted that slight changes in manufacturing will require calibration of each instrument and suggested that reducing variation and ensuring good housekeeping can lead to high-quality products.

Hence, each sensor needs to be individually characterised. Characterisation generally involves the process of testing sensors under specified conditions in comparison with known standards or references. The characteristics of each individual sensor can then be specifically determined. A typical example is the characterization of a resonance fluid density sensor (Bayliss and Downes, 2003) where the output is influenced by fluid viscosity and temperature. In order to characterize sensors, classical linear regression models have been widely used. For example, Bayliss and Downes (2003) described polynomial regression, and Hong (1999), for a resonant rate sensor, used a piecewise least-squares regression model.

Historical data provides the result of the calibration process, in the form of parameter estimates, for  $n$  sensors which are essentially identical, save for production error, to the current sensor. The parameter estimate for the current sensor should therefore be a per-

turbation of the mean of the  $n$  parameters estimates of the historical sensors. The idea is to use this information to replace information needed to be gathered on the current sensor, without sacrificing accuracy of the estimate. The reason for this is clear; reducing the number of measurements required per sensor could drastically reduce the calibration period. Hence, a methodology to replace one source of information with another is sought.

We summarise the historical data in the form of sample means and sample variances; a mean and a variance for each sensor parameter value, and the number of parameters per sensor will be denoted by  $p$ . Call this sample mean  $\mu = (\mu_1, \dots, \mu_p)$  and sample variance  $\Sigma = (\Sigma_1, \dots, \Sigma_p)$ . One might anticipate, therefore, that the parameter for the current sensor under investigation, i.e., the  $(n + 1)$ st, is, before measurements on it are taken, approximately a sample from the normal distribution with mean  $\mu$  and variance  $\Sigma$ . If it is possible to use this information, then it could turn out that less tests per sensor are needed in order to obtain the same level of accuracy in estimating the parameter for this  $(n + 1)$ st sensor. In other words, the information in related sensors, can be substituted for some of the  $N$  tests on the sensor and so for some  $0 < q < 1$  we may only need  $N_q = qN$  tests per sensor. This is important in a characterisation procedure when manufacturing sensors for it makes the process faster and hence ultimately saves costs.

The formal procedure to incorporate the outcomes of the parameter estimates from  $n$  sensors into the  $(n + 1)$ st is via the Bayes rule. Briefly, here, the parameter  $\beta$  for the  $(n + 1)$ st sensor is, before characterization takes place for the sensor, assumed to come from the normal distribution with mean  $\mu$  and variance  $\Sigma$ , which in Bayesian parlance is known as the prior distribution. This prior is combined with the  $N_q$  tests for the  $(n + 1)$ st sensor to derive a posterior distribution for  $\beta$ , and the estimate for  $\beta$ , say  $\hat{\beta}_B$ , can then be taken as the posterior mean value. The theory for this will be expanded on later in the paper.

Our novel achievement is to determine what  $q$  should be so that the accuracy of  $\hat{\beta}_B$  is identical, or no worse, than the accuracy of the estimate that would have been obtained via sensor tests only, i.e. the non-Bayesian approach. We will write the ordinary least squares

estimate as  $\widehat{\beta}_O$ .

The thinking is quite straightforward; the information from previous related sensor estimates takes the place for a certain number of tests for the next sensor. Some of these tests will only be establishing the value of  $\mu$  as the approximate location of  $\beta$ , then other tests are finding the shift from  $\mu$  to get  $\beta$ . It is these latter tests we wish to keep for the new sensor while making redundant the readings which identify  $\mu$ , since this is already known.

Describing the layout of the paper: In Section 2 we detail the industrial background to the problem and the current estimation process. Section 3 provides the background to the Bayesian approach and some theory indicating how we can achieve equivalent estimation accuracy while reducing the number of measurements. Section 4 presents the industrial case study which was carried out and shows that indeed there is the motivation for reducing the measurements per sensor while not compromising accuracy. The consequence is a reduction in both costs and time for the characterisation or calibration process. Finally, Section 5 concludes with a brief discussion.

**2. Background.** Here we provide further background to the sensor in question and also describe current practice for estimation of the parameters.

The sensor in question is an in-line device and measures the density of liquid passing through a pipe; and specifically for the devices this current application is concerned with are known as the resonator density sensor (or vibratory density sensor).

To elaborate; resonator sensors rely on mechanical resonance to sense physical variables. These include a wide range of sensors developed to sense density, viscosity, force or pressure and more (Langdon, 1985). The primary sensing device for a resonator sensor is a resonator and a commonly exploited signal from this resonance sensing device is its resonance frequency which depends on the mass and stiffness involved in a particular vibration mode. A fluid density sensor is so designed that the resonance frequency is directly related

to fluid density. An example of a density sensor is shown in Figure 1, where two identical U-shaped pipe sections are attached to a common manifold. The U-sections vibrate opposite to each other in the cantilever mode. For sensing fluid density, the fluid can be introduced into contact with the density sensor either off-line (Bayliss and Downes, 2003) or on-line (Headrick, 2008; Zheng et al., 2012).

For a vibrating tube density sensor (Lagourette et al., 1992), the relationship between the density and the resonance frequency is

$$\rho(T, P) = \frac{K(T, P)}{4\pi^2 V(T, P)} \cdot \frac{1}{f^2} - \frac{M}{V(T, P)}, \quad (1)$$

where  $\rho(T, P)$  is the density, which depends on the temperature  $T$  and pressure  $P$ ;  $f$  is the natural resonance frequency, and  $M$  is an unknown but fixed mass. It is known that  $K(T, P)$  is related to the stiffness of the resonance sensing device and  $V(T, P)$  is the internal volume, the change of which is generally much smaller than the stiffness change. For example, for stainless steel, the linear thermal expansion rate is nearly 28 times smaller than the rate of Young's modulus, which is a measure of the stiffness of an elastic material, change due to temperature. Thus assuming a constant internal volume, (1) becomes

$$\rho(T, P) = \frac{K(T, P)}{4\pi^2 V} \cdot \frac{1}{f^2} - \frac{M}{V}.$$

This paper is concerned with a resonator sensor as described by Headrick (Headrick, 2008). For such a sensor, as Headrick (Headrick, 2008) noted, pressure is only one of the causes to create stresses on the sensing element. It is also known that strain sensing devices can be readily attached to the resonator sensor to sense the deformation due to pressure or other forces (Cimalla et al., 2007). If a conventional strain gauge is attached to the sensing element to sense deformation, its resistance change  $\Delta R$  can provide a direct input variable

in addition to temperature. Thus, the above equation can be simplified to

$$\rho = \frac{A}{f^2}C(\Delta T, \Delta R) - B, \quad (2)$$

where  $A$  and  $B$  are known constants since they can be obtained by two reference fluids at a known condition when  $\Delta T = 0$  and  $\Delta R = 0$ , and  $C(\Delta T, \Delta R)$  is the compensation or correction function.

In order to determine  $\rho$ , the correction function  $C$  needs to be estimated and this is done by varying the sensing condition ( $\Delta T$  and  $\Delta R$ ) for a fluid with well-known density characteristics. Water is normally used due to its availability and its density  $\bar{\rho}$  calculated according to temperature and pressure is used as a reference value (see IAPWS-IF97). Hence, rewriting (2), we have

$$Y = \frac{\bar{\rho} + B}{A/f^2} = C(\Delta T, \Delta R). \quad (3)$$

Data exists for this model; i.e.  $(Y_i, (\Delta T, \Delta R)_i)$ , since the reference density values can be determined by flowing water through the density sensor. Hence  $Y$  can be observed under these experimental conditions and  $\Delta T$  and  $\Delta R$  can be regarded as input variables. Additionally, a linear model which has been demonstrated to be appropriate by the company, and widely used in the industry, is given by

$$C(\Delta T, \Delta R) = \beta_1 + \beta_2\Delta T + \beta_3(\Delta T)^2 + \beta_4\Delta T \Delta R + \beta_5(\Delta R)^2.$$

Therefore, this problem can be simplified as follows: given input pairs

$$(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N),$$

with  $x_i \in \mathbb{R}^p$  are the combination of  $\Delta T$  and  $\Delta R$ , the response  $y_i$  is related to the explana-

tory  $x_i$  as follows:

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_p x_{ip} + \epsilon_i, \quad \text{for } i = 1, \dots, N, \quad (4)$$

where the  $(\epsilon_i)$  are considered to be independent error terms and follow a normal distribution, i.e.  $\epsilon_i \sim N(0, \sigma^2)$ , where  $\sigma^2$  can be estimated and therefore treated as known in this paper. The ordinary least squares estimate of  $\beta = (\beta_1, \dots, \beta_p)$  and  $\sigma^2$  are, respectively, given by

$$\widehat{\beta}_O = (X'X)^{-1}X'y \quad (5)$$

and

$$\widehat{\sigma}^2 = \frac{(y - X\widehat{\beta})'(y - X\widehat{\beta})}{N - p},$$

where  $X$ ,  $\mathbf{y}$  are the matrix forms of the  $(x_i)$  and  $(y_i)$ , respectively,  $N$  is the number of observations per sensor and  $p$  is the dimension of the parameter  $\beta$ . Specifically,

$$X = \begin{pmatrix} x_{11} = 1 & \cdots & x_{1p} \\ & \cdots & \\ x_{N1} = 1 & \cdots & x_{Np} \end{pmatrix}$$

and

$$\mathbf{y}' = (y_1, \dots, y_N).$$

Ordinary least squares is widely regarded as the dominant statistical technique for estimating  $\beta$ .

**3. The Bayesian Model.** Here we describe the Bayesian approach to estimation; see also Bernardo and Smith (1994). The defining feature for the Bayesian model is the prior distribution for  $\beta$ . This is taken, as is common in linear regression models, to be a normal



distribution and based on the assumption that the  $\beta$  for the new sensor is a random perturbation from  $\mu$ , where  $\mu$  is the sample mean of the estimates from the first  $n$  sensors. And  $\Sigma$  will be taken as the diagonal matrix with diagonal terms the sample variances from the historical data. Specifically, if  $(\hat{\beta}_1, \dots, \hat{\beta}_n)$  are the parameter estimates from the first  $n$  sensors, even if they are obtained via OLS, we compute

$$\mu = n^{-1} \sum_{i=1}^n \hat{\beta}_i \quad (6)$$

and  $\Sigma = \text{diag}(\Sigma_j)$ , where

$$\Sigma_j = (n-1)^{-1} \sum_{i=1}^n \left( \hat{\beta}_{ij} - \mu_j \right)^2. \quad (7)$$

As a result, we can define the prior information as  $\pi(\beta) = N(\mu, \Sigma)$  which can be written as

$$\pi(\beta) \propto \exp \left\{ -\frac{1}{2} (\beta - \mu)' \Sigma^{-1} (\beta - \mu) \right\}. \quad (8)$$

The likelihood function is from the linear regression model (4). Since  $\sigma^2$  is estimated highly accurately, rather than assign a prior as is usual for Bayesian inference, we instead fix it at  $\hat{\sigma}^2$ , though we shall write it as simply  $\sigma^2$ ; and write  $\lambda = \sigma^{-2}$ . It is reasonable to fix  $\sigma$  since it is estimated with high accuracy and, moreover, it is extremely small (of the order  $\times 10^{-7}$ ).

Consequently, the data model can be written as

$$f(\mathbf{y} | X, \beta) \propto \exp \left\{ -\frac{\lambda}{2} (\mathbf{y} - X\beta)' (\mathbf{y} - X\beta) \right\}. \quad (9)$$

Next we show how to learn about  $\beta$  via the posterior distribution.

The Bayes update combines the prior and the likelihood, equations (8) and (9), respectively, to produce the posterior distribution. Thus we can write the posterior, obtained by

multiplying the likelihood and prior together, in the following way:

$$\pi(\beta | X, \mathbf{y}) \propto \exp \left\{ -\frac{\lambda}{2} (\mathbf{y} - X\beta)' (\mathbf{y} - X\beta) \right\} \exp \left\{ -\frac{1}{2} (\beta - \mu)' \Sigma^{-1} (\beta - \mu) \right\}.$$

Combining the two terms and collecting up the  $\beta$  terms, the posterior distribution can be identified as the normal distribution, given by

$$\pi(\beta | \text{data}) = N \left( (\lambda X'X + \Sigma^{-1})^{-1} (\lambda X'\mathbf{y} + \Sigma^{-1}\mu), (\lambda X'X + \Sigma^{-1})^{-1} \right). \quad (10)$$

The posterior mean is given by

$$\hat{\beta}_B = (\lambda X'X + \Sigma^{-1})^{-1} (\lambda X'\mathbf{y} + \Sigma^{-1}\mu) \quad (11)$$

and this will act as the estimate of  $\beta$ .

Now we show how we can, using the Bayesian model, reduce the number of observations for the new sensor, from  $N$  to  $N_q$ , while retaining the same level of accuracy for the estimate of  $\beta$ . In the following we define

$$p(\hat{\beta}) = \|\text{cov}(\hat{\beta})\|$$

as the precision (with lower values being preferable) of an estimator  $\hat{\beta}$ , which is the trace of the covariance matrix of the estimator. We assume without loss of generality that  $X'X$ , obtained via transformation, is a diagonal matrix so for some set of positive numbers  $(d_j)_{j=1}^p$  we have  $X'X = \text{diag}(d_1, \dots, d_p)$ .

**THEOREM.** There exists a subset  $A_1$  of  $A = \{1, \dots, N\}$  such that, if

$$\min_{j=1, \dots, p} \{d_j \Sigma_j\} > 4\sigma^2,$$

then

$$p(\widehat{\beta}_O) \geq p(\widehat{\beta}_B),$$

where  $\widehat{\beta}_O$  is the OLS estimator from the set  $A$  and  $\widehat{\beta}_B$  is the Bayes estimator from the reduced set  $A_1$ .

*Proof.* Consider the model

$$\mathbf{y} = X\beta + \sigma\epsilon$$

where, under  $A$ ,  $\mathbf{y}$  is  $N \times 1$ ,  $X$  is  $N \times p$  and  $\beta$  is  $p \times 1$ . The  $\epsilon$  is normal with mean 0 and variance-covariance matrix  $I_N$ . The  $X'X$  is a diagonal matrix  $(d_1, \dots, d_p)$  where

$$d_j = \sum_{i=1}^N x_{ij}^2.$$

The ordinary least squares estimator of  $\beta_O$  is

$$\widehat{\beta}_O = (X'X)^{-1}X'\mathbf{y}$$

while the accuracy of this estimator is measured by the trace of the variance-covariance matrix; i.e.

$$\|\text{Cov}(\widehat{\beta}_O)\| = \sigma^2 \sum_{j=1}^p \frac{1}{d_j}.$$

This represents the magnitude, or is a measure of the accuracy of the estimate.

On the other hand, under the reduced observations based on  $A_1$ , we have

$$\mathbf{y}_1 = X_1\beta + \sigma\epsilon,$$

where, under  $A_1$ ,  $\mathbf{y}_1$  is  $N_q \times 1$ ,  $X_1$  is  $N_q \times p$  and  $\beta$  remains  $p \times 1$ . The  $\epsilon$  is now normal with mean 0 and variance-covariance matrix  $I_{N_q}$ . Now  $X_1'X_1$  is the diagonal matrix  $(d_{11}, \dots, d_{1p})$

where

$$d_{1j} = \sum_{i \in A_1} x_{ij}^2.$$

Let us take  $X' = [X'_1, X'_2]$ , so we now have

$$d_j = d_{1j} + d_{2j} = \sum_{i \in A_1} x_{ij}^2 + \sum_{i \in A_2} x_{ij}^2.$$

Then, if we use only the observations from  $A_1$  and use a Bayes estimator for  $\beta$ , with prior  $N(\mu, \Sigma)$ , then

$$\widehat{\beta}_B = (X'_1 X_1 + \sigma^2 \Sigma^{-1})^{-1} (X'_1 \mathbf{y}_1 + \sigma^2 \Sigma^{-1} \mu)$$

and

$$\|\text{Cov}(\widehat{\beta}_B)\| = \sigma^2 \sum_{j=1}^p \frac{d_{1j}}{(d_{1j} + \sigma^2 \Sigma_j^{-1})^2}.$$

We now investigate when these two measures of uncertainty are equal. We can achieve this by setting

$$\frac{1}{d_{1j} + d_{2j}} = \frac{d_{1j}}{(d_{1j} + \sigma^2 \Sigma_j^{-1})^2};$$

that is,

$$d_{2j} = 2\sigma^2 \Sigma_j^{-1} + \sigma^4 \Sigma_j^{-2} / d_{1j}. \quad (12)$$

Hence, we find the sets  $A_1$  and  $A_2$ , with  $A_1$  as small as possible, and the dimension of  $A_1$  will be  $N_q$ , but with (12) conditions holding true. Since the  $d_j$  are fixed, we can solve for  $d_{1j}$  to get

$$d_{1j} = d_j/2 - \sigma^2 \Sigma_j^{-1} + \sqrt{d_j^2/4 - \sigma^2 \Sigma_j^{-1} d_j} \quad (13)$$

and hence we need

$$d_j > 4\sigma^2 \Sigma_j^{-1}.$$

In summary, we find the smallest set  $A_1$  for which (13) holds. This completes the proof.

Specifically, in our example, we have a very small  $\sigma$ . Actually, if

$$\sigma^2 \Sigma_j^{-1} < d_j/4$$

for all  $j$  we can show immediately that

$$\sum_{j=1}^p \frac{d_{1j}}{(d_{1j} + d_j/4)^2} \leq \sum_{j=1}^p \frac{1}{d_j}$$

since

$$d_{1j}d_j \leq d_{1j}^2 + d_{1j}d_j/2 + d_j^2/16 \quad \text{because} \quad (d_{1j} - d_j/4)^2 \geq 0.$$

In summary, we can match the accuracy of the least squares estimator of  $\beta$  by using a Bayes estimate using a reduced set of observations based on the covariate values  $(x_i : i \in A_1)$  whereby, for all  $j = 1, \dots, p$ , we have

$$d_{1j} = \sum_{i \in A_1} x_{ij}^2 = d_j/2 - \sigma^2 \Sigma_j^{-1} + \sqrt{d_j^2/4 - \sigma^2 \Sigma_j^{-1} d_j},$$

and provided  $d_j > 4\sigma^2 \Sigma_j^{-1}$ .

As a way of illustration, suppose that  $4\sigma^2 \Sigma_j^{-1} = \pi d_j$  for some  $0 < \pi < 1$ . Then

$$d_{1j} = d_j/2 - \pi d_j/4 + \frac{1}{2} \sqrt{d_j^2 - \pi d_j^2}.$$

So, if  $\pi = 3/4$ , for example, then

$$d_{1j} = \frac{9}{16} d_j,$$

which allows the reduction sought, allowing  $A_2$  to be non-empty. So in this toy example, we take  $A_1$  such that

$$\frac{\sum_{i \in A_1} x_{ij}^2}{\sum_{i \in [n] - A_1} x_{ij}^2} = \frac{9}{7}$$

where  $[n] = \{1, \dots, n\}$ .

**4. Application.** The methodology proposed in Section 3 demonstrates the potential for a reduction in the testing and calibration, i.e. finding the individual parameter values, of a new sensor in terms of the number of measurements made on it. An industrial case study was therefore proposed between a leading manufacturer of industrial sensors and a team of Bayesian statisticians from the University of Kent. The aim is to demonstrate the Bayesian approach and how it can be used to reduce the number of measurements made per sensor, reducing both costs and time for the calibration process.

The measurements which were to be removed from the calibration process were targeted at the mid-range temperatures from the three temperatures currently used. The calibration involves pumping water through the sensor and measurements made at three temperatures; low, medium and high. For each temperature there are three pressure settings and hence 9 settings in total. For better repeatability, a number of readings, usually 2 or 3, are taken at each of the 9 settings.

In the current practice, the temperature of the water going through the sensor does so in the following order; low, followed by medium, then high temperature, followed by medium and back to low temperature. This pattern with 2 measurements made at each condition is illustrated in Figure 2a. It is the measurements from the medium temperature which were proposed to be removed, and the parameters then estimated via the Bayesian approach, which will be compared with the traditional least squares approach, currently implemented by the company, using all the measurements from all of the temperatures. In total there are 30 measurements and in the specific sensor which was looked at, the Bayesian approach would reduce this to 18, as shown in Figure 2b. Each temperature setting in a calibration takes approximately 2 hours and the total time, including the time to install and remove sensors from the testing facility, takes 7 hours. The time for the calibration process is dependent on the actual shape and size of the sensor. Also, the quality of the least squares estimators does vary. However, for the case study reported here, the focus was on one of the sensors which provided good estimators and took 7 hours for completion of the calibration

process.

The accuracy of the estimators is easily checked for each sensor, the role of which is to compare the sensor output with the standard density for water at a known temperature and pressure. This is the means by which we can also evaluate and compare the Bayesian method with the least squares approach.

With 7 hours as a typical time for processing a single sensor, only two runs per day can be accommodated on each testing facility. For example, an operator starts at 6 a.m. to start the test and the test cycle will finish at 1 p.m. Thus, without a night shift, only 2 batches of sensors per day can be tested. With increased demand for the products, the company is looking for opportunities to increase its manufacturing capacity.

The application of the Bayesian method to this case study is, therefore, mainly aimed at increasing the manufacturing capacity, by reducing the calibration time per sensor. Additionally, and it goes without saying, the costs will also be significantly reduced. Ideally, the theoretical capacity should be increased by 50% so that three test runs can be accommodated per day on each testing facility. Since each test cycle includes three different temperature levels, the goal to eliminate one temperature level in this case has the potential to reduce the calibration time from 7 hours to 5 hours. This can comfortably allow three test runs per day for each testing facility.

Implementation of the proposed Bayesian method started from the study of historical datasets for each sensor type and their ordinary least squares (OLS) regression estimators. This would form the basis for the construction of the prior distribution for the sensor under current investigation. Each historical dataset is based on the full measurement conditions (i.e. 30 measurements as shown in Figure 2a) and the OLS method works well under the required high precision demanded by the company. It was found that the estimate of  $\beta$  via OLS closely follows a 5-dimensional multivariate normal distribution. Thus, the parameter value of the new sensor under current investigation will have its parameter value as coming from this normal distribution to a very high level of probability. With this knowledge,

the measurements from the sensor are to find out exactly where in this distribution the parameter values lie. On the other hand, without this information, one needs to first re-learn where the parameter values are located in a broad sense. Hence, for OLS which does not use the prior information, more readings are needed. The supporting theory for this claim was presented in Section 3.

For the calibration or characterization of new sensors, we are comparing the performance of the OLS estimator and the proposed Bayesian estimation method using the different test conditions, as show in Figure 2. The OLS estimator uses the full test conditions (i.e. 30 measurements) and the proposed Bayesian method uses a reduced dataset, where the tests at the middle level of temperature were removed, which resulted in only 18 test conditions.

Figure 3 shows the differences between the predicted values from the two methods. Using the parameter estimates, i.e.  $\hat{\beta}_O$  and  $\hat{\beta}_B$ , the predicted outcome at each of the input values can be computed and the difference with the true outcome at these values computed. These differences are plotted in Figure 3. The circled line gives this discrepancy for the OLS method, while the dotted line is obtained by the Bayesian method. As can be seen, the proposed Bayesian method using the reduced dataset preforms at least as well as the OLS method; since both methods are extremely close and lie well within the required upper and lower limits demanded by the industry standards. Ultimately, the concern of the company is that the summary of the calibration process allows the sensor to pass the test; and in this respect the two methods provided identical conclusions to be reached. In fact, on over 400 sensors, the pass rate was equivalent at 99.8%.

**5. Conclusions.** It has been shown that a Bayesian method using reduced measurements on an industrial precision sensor, which incorporates historical information on the estimates of similar sensors, can successfully characterise each sensor. This model has the added financial benefits which arise from removing an entire test condition as well as reducing the overall data collection time. The time saving assists the industrial company by increasing



the overall production capacity.

It should be made clear that this is a real life industrial problem and the company have made it apparent that a procedure of this nature will be implemented in the near future in the production line. That is, the sensor parameter values will be estimated using the Bayesian method. As a new sensor is calibrated and passed, so its value will be used to refine the prior distribution for the next sensor and so on. For example, if the mean and the variance–covariance matrix of the prior distribution are obtained from the first  $n$  sensors via ordinary least squares as in (6) and (7) and the  $(n + 1)$ th sensor is calibrated, we have the estimate of  $\beta$ ,  $\widehat{\beta}_B$  as in (11). The mean and the variance–covariance matrix of the prior distribution to be used for the next new sensor will be updated in the following way:

$$\mu' = (n + 1)^{-1} \left( n\mu + \widehat{\beta} \right)$$

and  $\Sigma' = \text{diag}(\Sigma'_1, \dots, \Sigma'_p)$  where

$$\Sigma'_j = n^{-1} \sum_{i=1}^n \left( \widehat{\beta}_j - \mu'_j \right)^2,$$

where the information of the  $(n + 1)$ st sensor is used towards the prior.

The key idea of the proposed Bayesian method is that the prior is taken to be a normal distribution and the mean and the variance–covariance matrix can be determined via simple methods using historical data. While Bayesian methods are not universally accepted, the choice of prior being the argument, in this illustration we have not only demonstrated that the construction of the prior is obvious, but have also shown how its has practical and beneficial consequences.

## References

- [1] Baker RC. Variation in Flowmeter Manufacture: Some Observations and Lessons. *Journal of Engineering Manufacture* 2004; **218**: 961–975. DOI: 10.1243/0954405041486082
- [2] Bayliss DJ, Downes SP. An Evaluation of a Paar DMA 5000 Density Meter. Technical report 2003. National Physical Laboratory.
- [3] Bernardo JM, Smith AFM. *Bayesian Theory* (1st ed.), Wiley Series in Probability and Statistics ,1994.
- [4] Bishop CM, Tipping ME. Bayesian Regression and Classification. *Advances in Learning Theory: Methods, Models and Applications* 2003; **190**.
- [5] Cimalla V, Niebelschutz F, Tonisch K, Foerster Ch, Brueckner K, Cimalla I, Friedrich T, Pezoldt J, Stephan R, Hein M, Ambacher O. Nanoelectromechanical Devices for Sensing Applications. *Sensors and Actuators B: Chemical* 2007; **126**: 24–43. DOI:10.1016/j.snb.2006.10.049
- [6] Clarke GM, Cooke D. (1998), *A Basic Course in Statistics* (4th ed.), Arnold, 1998.
- [7] Headrick DC. High Pressure Resonant Vibrating-Tube Densitometer. Patent Application 2008; US 2008/0115577 A1.
- [8] Hong SK. Compensation of Nonlinear Thermal Bias Drift of Resonant Rate Sensor Using Fuzzy Logic. *Sensors and Actuators* 1999; **78**: 143–148. DOI: 10.1016/S0924-4247(99)00241-1
- [9] Lagourette B, Boned C, Saint-Guirons H, Xans P, Zhou H. Densimeter Calibration Method Versus Temperature and Pressure. *Measurement Science and Technology* 1992; **3**: 699–703. DOI:10.1088/0957-0233/3/8/002

- [10] Langdon RM. Resonator Sensors - a Review. *Journal of Physics E: Scientific Instruments* 1985; **18**: 103–115. DOI:10.1088/0022-3735/18/2/002
- [11] Thornton AK. A Mathematical Framework for the Key Characteristic Process. *Research in Engineering Design* 1999; **11**, 145–157. DOI: 10.1007/s001630050011
- [12] Zheng D, Shi J, Fan S. Design and Theoretical Analysis of a Resonant Sensor for Liquid Density Measurement. *Sensors* 2012; **12**: 7905–7916. DOI: 10.3390/s120607905

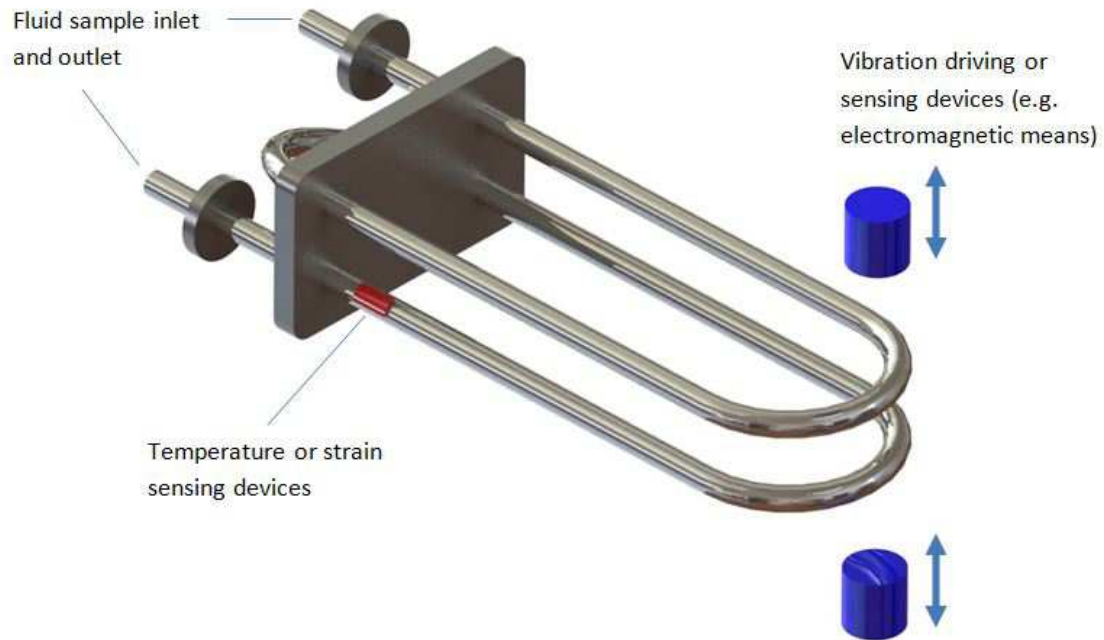


Figure 1: A Vibrating pipe liquid density transducer

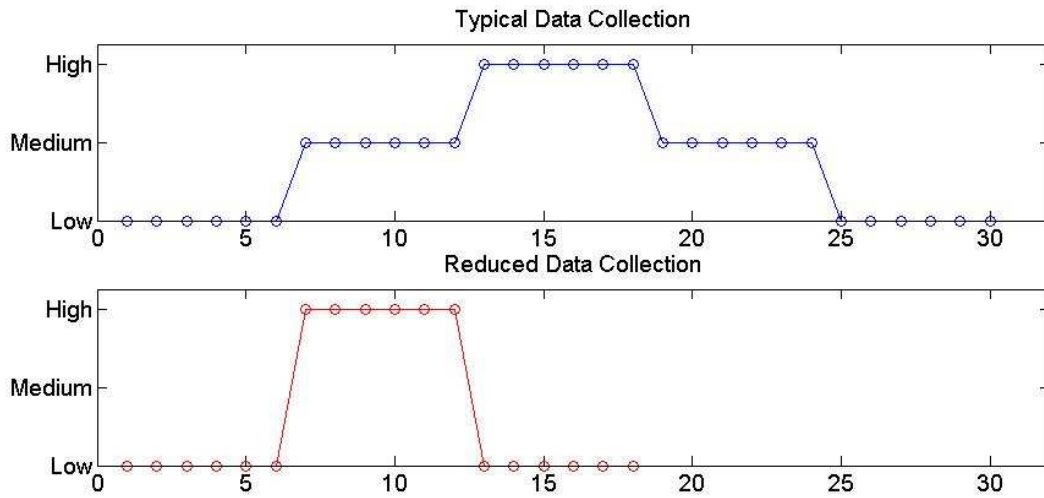


Figure 2: The data collection procedures, top plot (a) is the full test procedure including 30 test conditions for OLS; bottom (b) is the test procedure without medium temperature including 18 test conditions for the proposed Bayesian procedure.

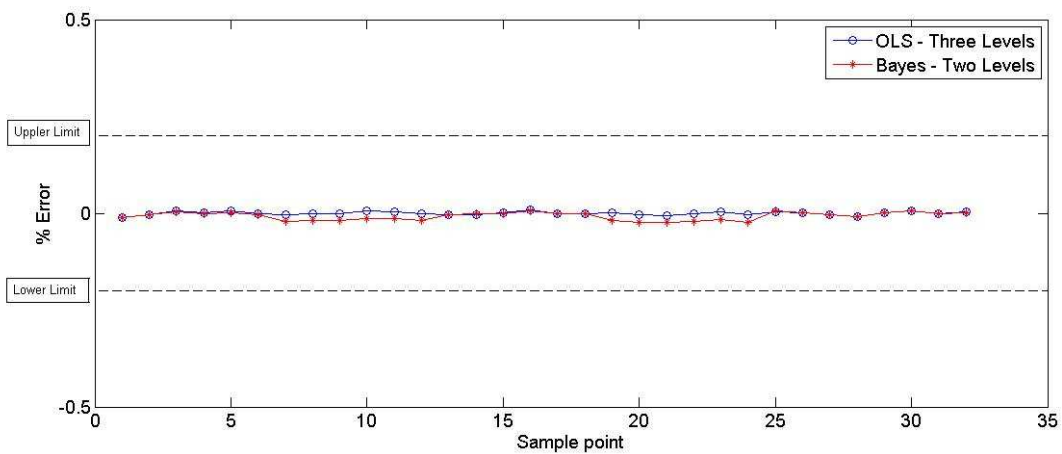


Figure 3: Residuals obtained for a sensor applied OLS and Bayesian methods with the full test conditions and the reduced test conditions, respectively.

**Christopher Eves** has a Masters degree in Statistics from the University of Kent, UK. He was employed as a statistical engineer during the period in which the research was being undertaken.

**Xue Wang** has a degree and PhD in statistics. She is currently lecturer School of Mathematics, Statistics & Actuarial Sciences, University of Kent. Her research interests include wavelet estimation, copulas and Bayesian nonparametric regression.

**Stephen Walker** has a degree in Mathematics and a PhD in Statistics. He is currently Professor in the Department of Mathematics, University of Texas at Austin. His research interests include Bayesian nonparametric methods, sampling approaches to inference and asymptotics for Bayesian models.