LABORATÓRIO NACIONAL
DE ENGENHARIA CIVIL

AN INNOVATIONS APPROACH TO FAULT DETECTION IN SENSOR NETWORKS

A case study

Metodologias para processamento, redução e gestão hierarquizada de dados provenientes da observação automática de estruturas

Lisboa • dezembro de 2013

I&D INSTRUMENTAÇÃO CIENTÍFICA

RELATÓRIO 419/**2013 – CIC/NSEM**

Título

AN INNOVATIONS APPROACH TO FAULT DETECTION IN SENSOR NETWORKS
A case study

Autoria

CENTRO DE INSTRUMENTAÇÃO CIENTÍFICA

João dos Reis
Bolseiro, Núcleo de Sistemas Eletrotécnicos e Mecânicos

José de Almeida Garrett
Investigador-Coordenador (aposentado)

Carlos Oliveira Costa
Investigador Principal, Diretor do Centro de Instrumentação Científica

AV DO BRASIL 101 • 1700-066 LISBOA
e-mail: lnec@lnec.pt
www.lnec.pt

Relatório 419/2013

Proc. 0903/112/17808

DETEÇÃO DE FALHAS EM REDES DE SENSORES BASEADA EM INOVAÇÕES

Resumo

As redes de sensores têm um largo espectro de aplicações, estando algumas implementadas em estruturas críticas, como é o caso das redes de monitorização estrutural. Contudo os sensores, devido à sua natureza e às condições ambientais em que operam, estão sujeitos a falhas que em ultima instância podem comprometer a qualidade da informação indispensável para um controlo de segurança efectivo. Neste relatório é apresentado um caso-estudo relativo à deteção de falhas em redes de sensores baseada em inovações. É estudada a possível aplicação, num contexto de monitorização estrutural, da deteção de falhas baseada em modelos utilizando inovações a uma rede de sensores. Este relatório faz uma breve introdução à detecção e isolamento de falhas (FDI) baseada em modelos e uma concisa revisão do seu estado da arte. Seguidamente é introduzido o conceito de FDI baseado em dados no contexto da validação de sensores. Tendo-se introduzido o necessário contexto é descrito o método utilizado para detecção de falhas em sensores usado neste caso-estudo. Seguidamente são apresentados os resultados obtidos e tecidas as considerações finais.

Palavras-chave: Sensor, Detecção de Falhas, Filtro de Kalman

AN INNOVATIONS APPROACH TO FAULT DETECTION IN SENSOR NETWORKS

Abstract

Sensor Networks are widespread across a large spectrum of applications. Some are even deployed in critical infrastructures, as the case of Structure Health Monitoring networks. However the sensors, due to its nature and to the harsh environmental conditions that they withstand, are subject to faults that may compromise the information quality that is indispensable for an effective safety control. In this report we present an Innovations Approach to Fault Detection in Sensor Networks case study. We study the possible application of the model based innovations approach to fault detection in sensor networks in a SHM context. This report gives a brief introduction to the model-based Fault Detection and Isolation (FDI) and a concise review of its state of the art. Next we introduce a data based FDI approach in the context of sensor validation. Given the necessary background we describe the method for sensor fault detection used in the case study. We then present the obtained results and final considerations.

Keywords: Sensor, Fault Detection, Kalman Filter

Index

Figure Index

Table Index

1. Introduction

In the domain of Civil Engineering it is normally designated as Structure Health Monitoring (SHM) the set of actions aimed at the detection and diagnosis of abnormal situations during the exploration of major civil engineering works of art in order to ensure safety and reduce maintenance and inspection costs. For that purpose, it is necessary the installation of a large number of sensors in the monitored works of art, in a robust and autonomous way, according to the established observation plan. However the sensors, due to their nature and to the harsh environmental conditions in which they operate, are subject to faults that ultimately may compromise the quality of the information essential for an effective safety control.

In this report we present an Innovations Approach to Fault Detection in Sensor Networks case study. We keep in mind the possible application of the model based innovations approach to fault detection in sensor networks studied in a SHM context. It is hoped that this study can contribute to the future development of innovative measurement Fault Tolerant Sensor Networks applied to Civil Engineering Structures; which may have potential application in the Structure Health Monitoring systems.

In a metrological sense, the sensor is a device used in the measuring process that is directly affected by the mensuranda and according to a predetermined law generates a signal related to its value. In the context of this report a broader sense of the term sensor is used, comprising the sensor itself as well as all the other elements in the measuring chain.

In this report we first make a brief description of model-based Fault Detection and Isolation (FDI) and a concise review of it’s state of the art. Next we introduce data based FDI approach in the context of sensor validation. Given the necessary background we describe the method for sensor fault detection used in the present case study. We then present the obtained results and make some final considerations.

2. Model based FDI

In the context of an industrial application, a fault is perceived as a non-permitted deviation of a characteristic property that leads to failure of the system or manufacturing facility to fulfill the purpose for which it was designed [1]. Although some effort has been made by the scientific community to establish a common terminology ([1]-[4]), the peculiarities of the involved multidisciplinary topics often lead to terminologies that are not unique.

Currently it is accepted that Model based Fault Detection and Isolation (FDI) consists of two steps [5],[6]:

- Generation of residuals (Innovations);
- Decision-making (including evaluation of the residuals).

In most publications it can be seen that the majority of the approaches are based on mathematical models of the system. Such models can be of two types [7]:

- Models based on first principles models;
- Input-output (data-driven) models.

While the model based on first principles is obtained through differential equations that represent the physical behavior of the system components, the input-output (also designated as data-driven) model is constructed using system identification techniques [8]-[11].

Although in general, analytical models based on first principles allow a greater depth in diagnosis, they usually require a difficult and laborious modeling, especially for non-linear cases.

Alternatively the input-output model is a powerful tool for dealing with the problem of modeling as well as to serve fault diagnosis [7]. In addition to the conventional techniques for identification, mainly for non-linear cases, methods to infer these models have been developed through: neural networks [12],[13]; fuzzy clustering [14]-[16]; immune networks [17]-[19]; Relevance Vector Machines (RVM) [20]; Pattern Recognition [21]; minimum mean square error (MMSE) [22]; and hierarchical Bayesian space-time (HBST) modeling [23].

3. Sensor Validation

When applied to sensors, fault detection and diagnosis (FDI) is usually designated as sensor validation [24]. However, although less frequently, another name is also common in the literature regarding this topic, and it is referred to as signal validation [21],[25].

Almost all the techniques of fault detection and diagnosis (FDI) described in the literature can be applied to sensor validation [24].

Although the fault detection can not usually depend on the physical redundancy of system sensors, due to cost and inefficiency inherent in the replication of the entire sensor system, tools can be designed to explore the redundancy of physical sensors in parts of the system where it exists [24].

As in fault detection and diagnosis (FDI) at the process level (Figure 1), in the validation of sensors one can compare the results of the system sensors measurements with mathematical models results that describe the static and dynamic relationships between the measured sensor data. This procedure is supported in the fault detection based on models techniques and, in the event of a fault, makes it possible to provide an estimate (during a finite time window) of the missing measurements [24].

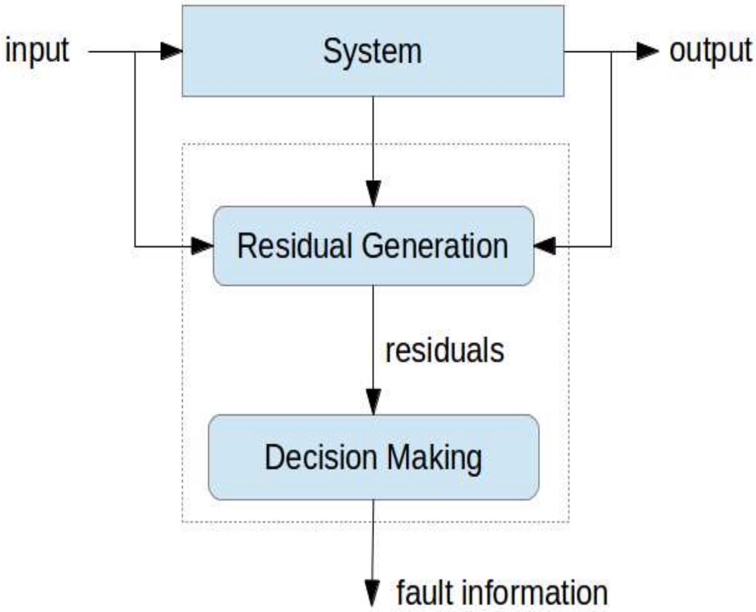

Figure 1 Model Based FDI

However, using this approach, care must be taken to distinguish between errors in the sensors and faults in the process or control system. Since a failure in the process may result in abnormal readings from the sensors, the developed algorithms may report a faulty sensor. The use of a higher layer in the diagnostic system must consider this situation [26].

In the literature various approaches exist for performing data-based fault detection, which can be divided basically into three main groups [27]:

- Classification-based fault detection: different system conditions representing faulty cases define different fault classes due to their appearance in the corresponding recorded data (fault patterns); also the faulty-free case represents one class; these classes are learned by training data and are applied whenever a new data point needs to be assigned to a class. The big disadvantage of this approach is that all kind of faults and their corresponding appearance in the data need to be known a-priori and therefore new faults cannot be detected often (only in the case that its pattern is luckily similar to an already known pattern of another fault). This disadvantage can be somewhat overturned using one class classification [28], or novelty detection methods [29] .

- Signal-based fault detection in intelligent sensors: can be seen as a single channel check approach where dynamic sensor data is analyzed with respect to the occurrence of peaks, drifts or other unnatural behaviors in their corresponding signal curves. As commonly no interactions in form of redundancy or correlation analysis between other sensors are taken into account, no wide-spread system failures can be detected within this approach.

- Model-based fault detection: multi-dimensional models or some of their parameters are trained from simulated, historic or online measured data and used as reference situation characterizing functional dependencies between measurement variables for the faulty-free case. The drawback for this approach is that if systematic failures occur in the training data (when no simulated data is available, of course), wrong models are trained, which get useless for fault detection. Besides, a fault isolation strategy has to be appended in order to identify the faulty variables amongst faulty-free ones, all integrated in complex high-dimensional models.

4. Method

General Method

Taking as base the general method proposed in [1] we will consider the following procedure for model-based sensor fault detection (Figure 2).

Development of a model

A mathematical model is developed for the system, based on physical information and statistical data. This model can be static or dynamic, linear or non-linear, continuous or discrete and deterministic or stochastic. The input and the output variables of the system are clearly defined and all the relevant parameters are identified. The model describes the behavior of the system under normal operating conditions. It also specifies the statistics of the measurement noise in the output variables.

Generation of a residual signal or the innovation process

The residual signal or the innovation process is defined as the difference between the actual system output and the expected output based on the model and the previous output data. The latter is generated directly by the model if the system is deterministic or by a statistical filter if the system is stochastic, i.e. subject to random inputs and variations. This difference is called the innovation process since it represents the new information brought by the latest observation. Under normal conditions, the error signal is "small" and corresponds to random fluctuations in the output since all the systematic trends are eliminated by the model. However, under faulty conditions, the error signal is "large" and contains systematic trends due to the fact that the model no longer represents the physical system adequately.

Statistics of the residual signal under normal conditions

In deterministic systems, the random fluctuations in the residual signal are due to the measurement noise in the output variables. Their statistics are obtained as part of the system description in the model development step. In stochastic systems, the statistics of the error signal are obtained from the filter which is used to predict the output of the system. For linear dynamic systems with Gaussian random inputs, a Kalman filter [30] generates both the residual signal and its statistics. For these systems, it is known that under normal conditions, the residual signal or the innovation process is a zero mean Gaussian white noise process [31], [32].

Outlier Detection via limit value checking

Given the statistics of the residual signal under normal conditions, a univariate statistical approach to limit sensing can be used to determine the thresholds for each generated residual; these thresholds define the boundary limits, and a violation of these limits would indicate an outlier in the actual system output. This approach is typically employed using a Shewart control chart [33].

Sensor fault detection via hypothesis testing

The problem of sensor fault detection is formulated as a Hypothesis Testing [34] problem, by considering as the null hypothesis, the normal operation of the system. The actual residual signal from the system is tested against this hypothesis at a certain level of significance. For example, if the system is described by a set of linear differential equations and a Kalman filter is used to generate the innovation process, the null hypothesis consists of testing the innovation process for zero mean, whiteness and a given covariance.

Isolation of the sensor fault

If a system fault is detected, the current model used to describe the behavior of the system may no longer be adequate. In order to diagnose the fault it may be necessary to develop a new model for the system. Since a failure in the process may result in abnormal readings from the sensors, the developed algorithms may report a faulty sensor. As stated before, to clarify this situation a higher layer in the diagnostic system must be considered. These subsequent procedures are beyond the scope of this report.

Linear dynamic systems

In this case study we are going to apply the general approach, outlined above, to a sensor network system describable by a set of discrete linear differential equations for sake of simplicity. The approach can also be used in continuous-time linear systems or carried over to nonlinear dynamic systems.

We propose to extend the Mehra and Peschon [1] “Innovations approach to Fault Detection and Diagnosis in Dynamic Systems” method (see Annex I) by adding outlier detection via limit value checking procedure, and including a normality test in the sensor fault detection via an hypothesis testing step.

Outlier Detection via limit value checking

Given the statics of the Standardized Innovation $\eta(t+1)$, generated by a Kalman Filter [30], the upper and lower thresholds on the Shewart Chart can be set. Therefore at the 0.27 per cent significance level, the measurement $z_i(t+1)$ is classified as outlier whenever

$$|\eta_i(t+1)| > 3 \tag{1}$$

Normality Tests

Several methods can be applied for detecting departures from normality [36]. Using the frequencist's inference framework, we opted to use the statistical hypothesis testing method. In this method the data is tested against the null hypothesis, that is: that the data is normally distributed. We propose to use the Anderson-Darling test and the Cramér–von Mises criterion. For details on this tests see Annex II.

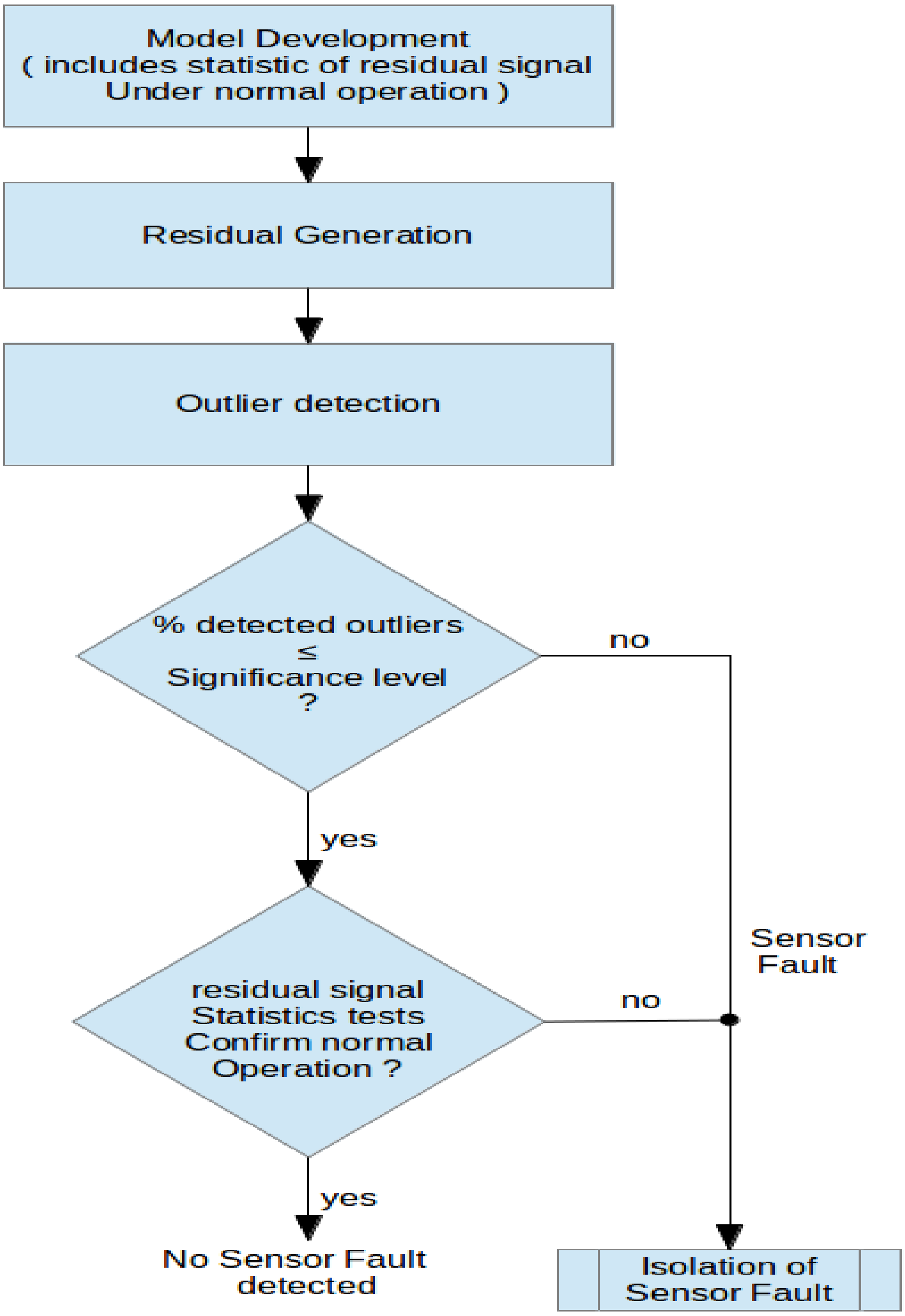

Figure 2 General Method for model-based sensor fault detection

5. Practical case study

The proposed methods were applied to a real world temperature sensor reading. The measurement data came from a sensor network installed on a bridge (Figure 3). The measurements were sampled at 500 Hz (global acquisition frequency of the sensor network system). Since no previous test was made to this data, it is unknown if the sensor data has faults or not.

It is assumed in this chapter that the reader has some kind of knowledge about state space discrete-time linear dynamic systems, Kalman filters and the normality tests used (Anderson-Darling test , Cramér–von Mises criterion), if that is not the case it is advisable that the Annexes Section should be read prior to reading of the remaining text.

For the intensities of Plant noise $\xi(t)$ and Measurement noise $\theta(t)$ were considered the following values:

$$\Xi(t) = 8.6703\mathrm{E}\text{-}16 \tag{2}$$

$$\Theta(t) = 2.729\mathrm{E}\text{-}8 \tag{3}$$

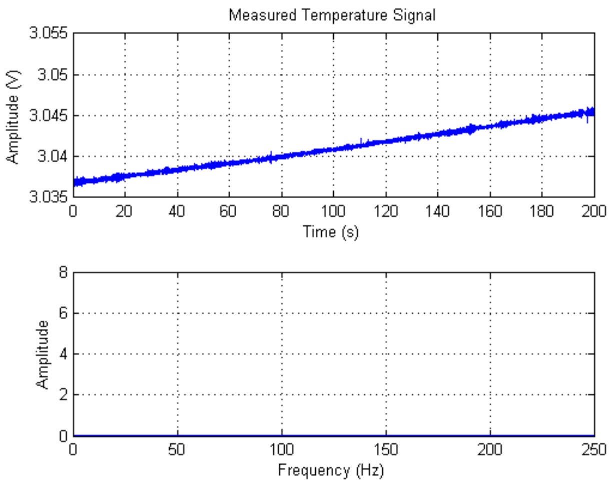

Figure 3 Real World Temperature sensor data

The dynamics of the system was described by an autoregressive process model:

$$\begin{cases} x(t+1) = x(t) + v(t) \\ v(t+1) = v(t) + \xi(t) \end{cases} \tag{4}$$

Where,

$x(t)$ is the amplitude of the temperature signal in volts;

$v(t)$ is the difference between the temperature signal in two consecutive time instants. It is assumed, for simplification purposes, that the sampling interval is 1 (one) second;

The measurement model considered was:

$$Z(t+1) = x(t+1) + \theta(t+1) \tag{5}$$

For the initialization step of the Kalman filter was considered:

$$x(0) = \begin{bmatrix} x(0) \\ \Xi \end{bmatrix}, \quad \Sigma_0 = \begin{bmatrix} \Theta & 0 \\ 0 & \Xi \end{bmatrix} \tag{6}$$

The calculated steady state filter gain was:

$$H = \begin{bmatrix} 0.0187 \\ 1.7657E-4 \end{bmatrix} \tag{7}$$

And the covariance matrixes:

$$\begin{aligned} \Sigma_p &= 1E-9 \begin{bmatrix} 0.5202 & 0.0049 \\ 0.0049 & 0.0001 \end{bmatrix} \\ \Sigma &= 1E-9 \begin{bmatrix} 0.5104 & 0.0048 \\ 0.0048 & 0.001 \end{bmatrix} \\ S(t+1) &= 2.7810E-8 \end{aligned} \tag{8}$$

In Figure 4 the result of the estimated output from the Kalman filter can be seen.

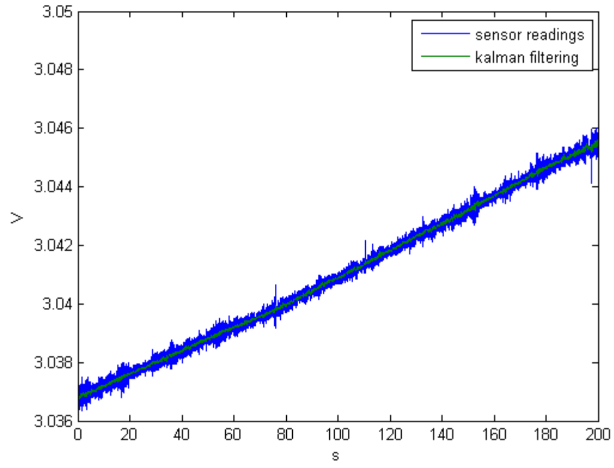

Figure 4 kalman filter estimation

Standardizing the innovation sequence and applying the threshold (1), 225 outliers were detected (Figure 5) which is within the significance level since we have 0.22 per cent outliers (10^5 samples).

If we look at the histogram of the innovation sequence (Figure 6) and to the normal probability plot (Figure 7) we can see that the outliers appear on the tails of the distribution, i.e. the distribution seems normal, although with departures from normality on the tails. If the outliers are removed or substituted by the previous value the distribution becomes asymmetrical and the nonlinearities on the normal probability plot accentuate on the tails (Figure 8).

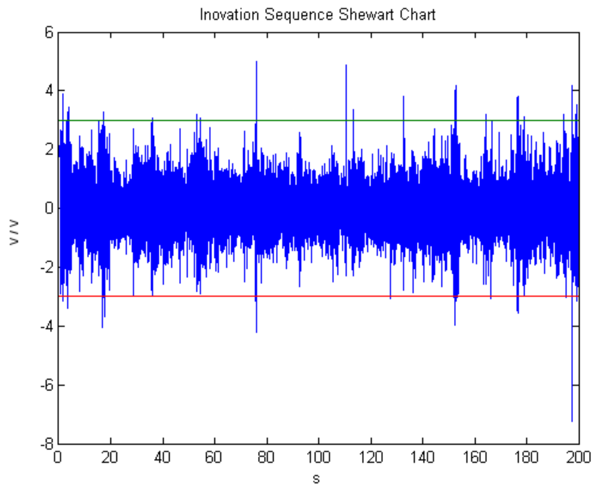

Figure 5 Innovation Sequence Shewart Chart

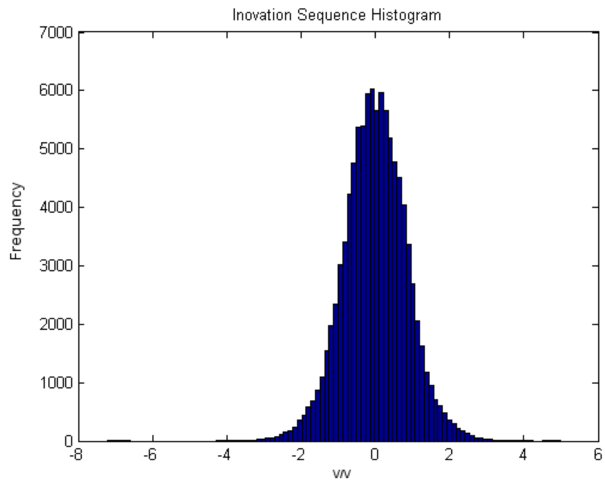

Figure 6 Histogram of the Innovation Sequence

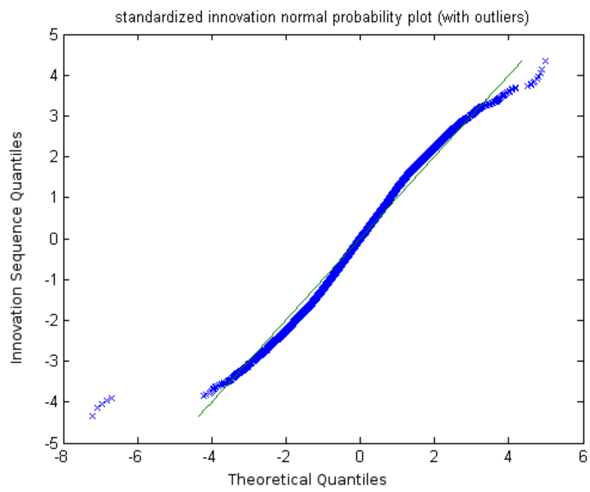

Figure 7 Normal probability plot of the innovation sequence (with outliers)

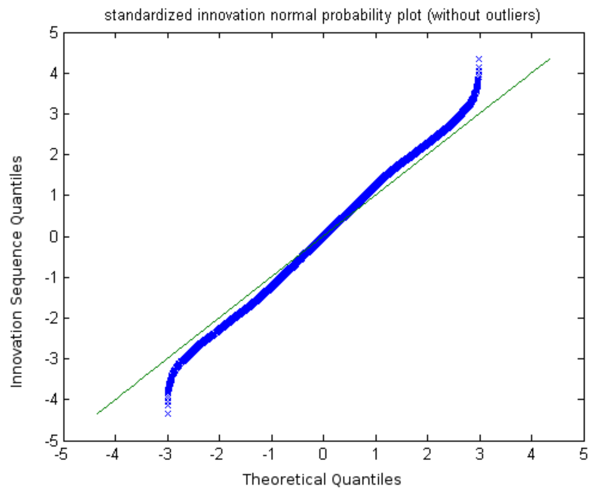

Figure 8 Normal probability plot of the innovation sequence (without outliers)

The null hypothesis is rejected in all of the normality tests (Table 1). The best results were obtained for the Cramér–von Mises criterion with the parameters estimated from the data.

Calculating the autocorrelation for the innovation sequence for the whiteness test (Figure 9), the null hypothesis is also rejected.

Table 1 - normality tests results

NORMALITY TEST	DATA WITH OUTLIERS	DATA WITHOUT OUTLIERS	NORMAL DISTRIBUTION	5% SIGNIF.
(Anderson-Darling) A^2	3704.9	3738	0.4825	2.492
(modified Anderson-Darling, mean and variance unknown) A^{*2}	112.99	19.106	0.5477	0.787
(modified Cramér–von Mises, mean and variance both known) W_1^{*2}	566.79	570.50	0.0976	0,461
(modified Cramér–von Mises, mean and variance unknown) W_2^{*2}	16.575	2.7471	0.0438	0,126

The Normal Distribution column represents the application of the tests to a synthetic dataset generated with normal distribution and the same number of samples as the tested innovation sequence.

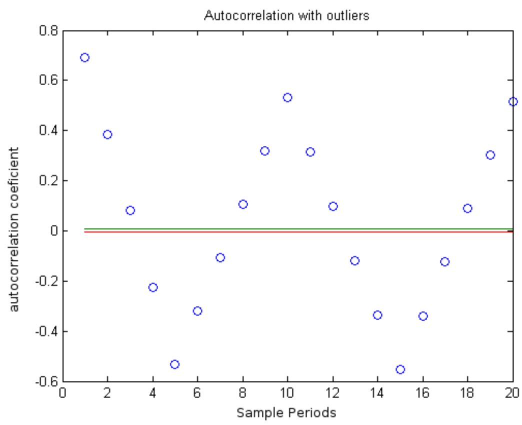

Figure 9 Autocorrelation of the innovation sequence

Mean and covariance tests assume that the innovation sequence is white. Therefore, the test sequence should be stopped. Analyzing the autocorrelation results, we can see that as expected short lags have a correlation since we are using a high sample frequency for the dynamic of the problem; what is unusual is the periodic pattern with a 20 ms period.

If we look thoroughly at the frequency spectrum of the measurement data (Figure 10) we found that low amplitude noise peaks are present on 50Hz, 150 Hz and 250 Hz.

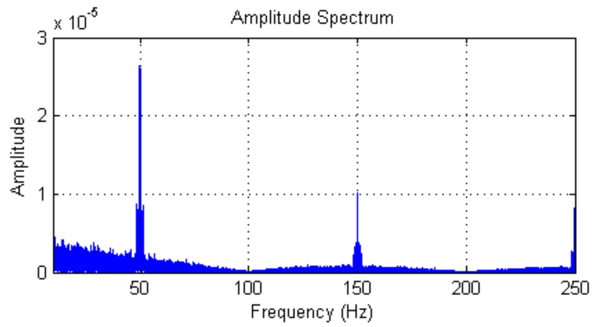

Figure 10 Amplitude Spectrum of the measurement data

It is suspected that the 50 Hz noise comes from a train power catenary located near the location of the sensor. The sensor measurements were captured immediately after the sensor network had been installed and at that time the section from where the data was obtained was poorly grounded. Since the measured noise amplitude is within the sensor accuracy ($\pm 3\ mV$) this does not affect in practice the accuracy of the sensor measurements.

We assumed for the model of the system that the measurement noise was white Gaussian noise, what in the end did not hold true. If the system ground can not be improved to eliminate this interference in the sensor readings, the model should be updated and colored measurement noise should be considered. To transform the colored sensor noise into the standard formulation of the Kalman filter the concept of prewhitening must be considered. Using this concept the colored sensor noise is represented as the output of a fictitious linear time variant dynamic system driven by pure white noise. System identification techniques can be used to identify the correct model for the fictitious prewhitening system. System states must be augmented to use the standard formulation of the Kalman filter [35]. The subsequent procedures are beyond the scope of this report.

6. Conclusion

A method for detecting faults in sensor networks based in the innovation sequence of a Kalman filter was tested. The tested method was applied to a real world data measurement with successful results, since it could identify that the tested signal sensor data had a (unknown) fault resulting from the interference of a 50 Hz non linear noise. Although the detected noise amplitude was inside the sensor measurement accuracy, the sensor Gaussian white noise assumption for the system did not hold true.

Care must be taken in the cleaning (or substitution) of the outliers since that operation can impair the sensor sample probability distribution as showed in chapter 5.

The normality test steps that were added to the method gave an early warning that a fault was present on the system, although the most informative hint for the identification of the fault did came from the autocorrelation test.

Subsequent work should consider multivariate statistical tests.

Lisboa, Laboratório Nacional de Engenharia Civil, dezembro de 2013

VISTOS

O Chefe do Núcleo de
Sistemas Eletrotécnicos e Mecânicos

João C. P. Palma

O Diretor do
Centro de Instrumentação Científica

Carlos Oliveira Costa

AUTORIA

João dos Reis
Lic. Eng.ª Eletrotécnica,
Lic. Eng.ª Informática
Bolseiro de Doutoramento

José de Almeida Garrett
Eng.º Eletrotécnico, Especialista
Investigador-coordenador
(aposentado)

C. Oliveira Costa
Eng.º Eletrotécnico, Especialista
Investigador Principal

References

[1] Mehra, R.K., Peschon, J. – **An Innovations Approach to Fault Detection and Diagnosis in Dynamic Systems**. U.K., Automatica, Vol. 7, pg 637-640, Pergamon Press, 1971

[2] Isermann, R. – **Experiences with process fault detection via parameter estimation**.S.G. Tzafestas, M.G.Singh, G. Schidt, eds.System Fault Diagnostics, Reliability & Related Knowledge-based Approaches. Dordrecht: D. Reidel Press.1987, p.3-33

[3] van Schrick, D. **- Remarks on terminology in the field of supervision, fault detection and diagnosis**: proceedings. of the IFAC Symposium On Fault Detection,Supervision and Safety for Technical Processes: SAFEPROCESS'97. University of Hull: Pergamon, 1997, p. 959-964

[4] Omdahl T **- Reliability, availability and maintainability (RAM) dictionary**. Milwaukee: ASQC QualityPress, 1988

[5] Chow, E.Y.; Wikksky, A.S. - **Issues in the development of a general algorithm for reliable failure detection**. Albuquerque, NM: Proc. Of the 19th Conference on Decision & Control, 1980

[6] Chow, E.Y.; Wikksky, A.S. - **Analytical redundancy and the design of robust detection systems**: IEEE Trans. Automat. Contr., 1984, AC-29(7), p.603-614

[7] Luh, Guan-Chun; Cheng, Wei-Chong **- Immune model-based fault diagnosis**: Mathematics and Computers in Simulation, Elsevier, 2005, vol. 67(6), p.515–539

[8] Frank, P.M. - **Fault diagnosis in dynamic system using analytical and knowledge based redundancy: a survey and some new results**. Automatica, 1990, 26(3), p.459-474

[9] Schneider , H.; Frank, P.M. - **Observer-based supervision and fault detection in robots using nonlinear and fuzzy logic residual evaluation**. IEEE Trans. Control Syst. Technol., 1996, vol. 4 (3), p.274–282

[10] Patton, R.J.;Chen, J. - **Observer-based fault detection and isolation: robustness and applications**. Control Eng. Pract., 1996, vol. 5 (5), p. 671–682

[11] Balle, P. - **Fuzzy-model-based parity equations for fault isolation**. Control Eng. Pract., 1999, vol. 7 (2), p. 261–270

[12] Chen, Y.M.; Lee, M.L. - **Neural networks-based scheme for system failure detection and diagnosis**. Math. Comput. Simul., 2002, vol. 58, p. 101–109

[13] Thomas, P. ; Lefebvre, D. - **Fault detection and isolation in non-linear systems by using oversized neural networks**. Math. Comput. Simul., 2002, vol. 60, p. 181–192

[14] Ronen, M.; Shabtai,Y.; Guterman, H. - **Hybrid model building methodology using unsupervised fuzzy clustering and supervised neural networks**. Biotechnol. Bioeng., 2002, vol. 77 (4), p. 420-429

[15] Pakhira, M.K.; Bandyopadhyay, S.; Maulik, U. - **Validity index for crisp and fuzzy clusters. Pattern Recog.**, 2004, vol. 37 (3), p. 487–501

[16] Yang, T.N.; Wang, S.D. - **Competitive algorithms for the clustering of noisy data. Fuzzy Sets Syst.**, 2004, vol. 141 (2), p. 281–299

[17] Costa Branco, P.J.; Dente, J.A.;Vilela Mendes, R. - **Using immunology principles for fault detection**. IEEE Trans. Ind. Electron., 2003, vol. 50 (2), p. 362–373

[18] Gonzalez, F.; Dasgupta, D.; Kozma, R. - **Combining negative selection and classification techniques for anomaly detection**. Hawaii: Proceedings of the IEEE Congress on Evolutionary Computation, 12–17 May 2002, p. 705–710

[19] Ishiguro, A.; Watanabe, Y.; Uchikawa, Y. - **Fault diagnosis of plant systems using immune networks**. Las Vegas, NV, USA: Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, 1994, p. 34–42

[20] Conde, Erick F. - **Environmental Sensor Anomaly Detection Using Learning Machines**. Logan, USA: Utah State University, 2011. Master Tesis

[21] Fantoni, P.F.; Hoffmann, M:I.;Shankar, R.; Davis, E.L. - **On-Line Monitoring of Instrument Channel Performance in Nuclear Power Plant Using PEANO**: Progress in Nuclear Energy. London: Pergamon, v. 43 (1–4), 2003
[22] Kullaa, J. – **Sensor validation using minimum mean square error estimation**. Mechanical Systems and Signal Processing 24 144-1457: Elsivier, 2010
[23] Ni, Kevin; Pottie, Greg - **Sensor network data fault detection with maximum a posteriori selection and bayesian modelling**. Transactions on Sensor Networks (TOSN) Volume 8 Issue 3, Article No. 23, July 2012 : ACM New York, NY, USA, 2012
[24] Fortuna, Luigi [*et al.*] - Soft **Sensors for Monitoring and Control of Industrial Processes**. London: Springer-Verlag, 2007.(Advances in industrial control). Ch. 9
[25] Jabloński, Adam; Barszcz, Tomasz; Bielecka, Marzena – **Automatic validation of vibration signals in wind farm distributed monitoring systems**. Measurement: Elsevier, 2011, 44 , p.1954-1967.
[26] Ibargüengoytia, Pablo H.; Sucar, Luis Enrique; Vadera, Sunil **– Real Time Intelligent Sensor Validation**. IEEE Transactions on Power Systems, 2001, vol. 16(4), p. 770-775
[27] Lughofer, E.; Klement , E. P.; Lujan, J. M. ; Guardiola, C. - **Model-based Fault Detection in Multi-Sensor Measurement Systems**. Proceedings IEEE IS 2004, Varna, Bulgaria, 2004
[28] Johannes , D. M. - **One-class classification**. Netherlands : Advanced School for Computing and Imaging, 2001 .Phd. Tesis
[29] Markou, M.; Singh ,S. – **Novelty detection: a review—part 1: statistical approaches**. Signal Processing 83 (2003) 2481 – 2497: Elsevier, 2003
[30] Kalman, R. E. - **New methods and results in linear prediction and filtering theory**. Proc. Syrmp. on Engineering Applications of Random Function Theory and Probability: John Wiley, New York, 1961
[31] Mehra, R. K. **- On the identification of variances and adaptive Kalman filtering**. IEEE Trans. Aut. Control, 1970
[32] Kailath, T. - **An innovations approach to least-squares estimation**, Part I, IEEE Trans. Aut. Control AC-13, 646-655, 1968
[33] Chiang, L.; Russell E.; Braatz, R. – **Fault Detection and Diagnosis in Industrial Systems**. Springer-Verlag, 2001
[34] Kendall, M. G.; Stuart, A. - **The Advanced Theory of Statistics, Vol, 2**: Hafner, New York, 1961.
[35] Bryson, A. E.; Ho, Y. C. - **Applied Optimal Control**: Blaisdell, Waltham, Mass, 1969
[36] Stephens, M. A. - **EDF Statistics for Goodness of Fit and Some Comparisons**. Journal of the American Statistical Association 69: 730-737, 1974

Annexes

ANNEX I
Fault Detection in Linear Dynamic Sensor Network Systems

Fault Detection in Linear Dynamic Sensor Network Systems

We describe in this annex the application of Mehra and Peschon "Innovations approach to Fault Detection and Diagnosis in Dynamic Systems" method [1] to a sensor network system that can be describable by a set of linear difference equations. Continuous-time sensor network linear systems can be treated in the same way (See Fig. 1). Static systems can be regarded as special cases of the dynamic systems. The approach can also be carried over to nonlinear dynamic systems.

Consider a discrete-time linear dynamic system whose model is:

State Dynamics:

$$x(t+1) = A(t)x(t) + L(t)\xi(t) \tag{1}$$

Measurements:

$$z(t+1) = C(t+1)x(t+1) + \theta(t+1) \tag{2}$$

With time index $t = 0,1,2,...$

Where

$x(t)$ is a $\Re^n$ state vector (stochastic sequence non-white)

$z(t)$ is a $\Re^r$ measurement vector

$\xi(t)$ is a $\Re^p$ white plant noise

$\theta(t)$ is a $\Re^r$ white measurement noise

$A(t)$ is as $\Re^{n\times n}$ state-transition matrix

$L(t)$ is as $\Re^{n\times p}$ noise distribution matrix

and

$C(t)$ is a $\Re^{r\times n}$ output matrix.

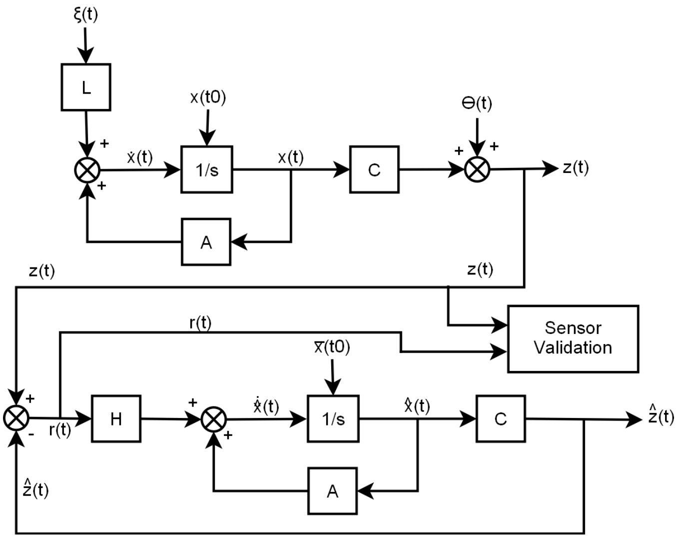

Figure 1 Innovation approach to sensor fault detection

Plant noise $\xi(t)$ and Measurement noise $\theta(t)$ are Gaussian discrete white noise, with mean and covariance:

$$E[\xi(t)] = 0, \quad \text{cov}[\xi(t); \xi(\tau)] = \Xi(t)\delta_{t\tau} \tag{3}$$

$$\Xi(t) = \Xi^T(t) \geq 0 \tag{4}$$

$$E[\theta(t)] = 0, \quad \text{cov}[\theta(t); \theta(\tau)] = \Theta(t)\delta_{t\tau} \tag{5}$$

$$\Theta(t) = \Theta^T(t) > 0 \tag{6}$$

$$\text{cov}[\theta(t); \xi(\tau)] = 0 \tag{7}$$

Where $\delta_{t\tau}$ denotes the Kronecker delta

$$\delta_{t\tau} = \begin{cases} 1, & t = \tau \\ 0, & t \neq \tau \end{cases}$$

and $E[\cdot]$ denotes the expectation operator , and $\text{cov}[\cdot;\cdot]$ the covariance operator.

The initial state $x(0)$ is also assumed to be random. The distribution of the state variables is Gaussian with mean and covariance:

$$E[x(0)] = \bar{x}(0), \quad \text{cov}[x(0); x(0)] = \Sigma_0 \tag{8}$$

$$\Sigma_0 = \Sigma_0^T \geq 0 \tag{9}$$

$$\operatorname{cov}[x(0);\xi(t)] = 0\ ,\quad \operatorname{cov}[x(0);\theta(t)] = 0 \tag{10}$$

The more general case of correlated plant noise and correlated measurement noise can be reduced to the above case by augmenting the state vector $x(t)$ [5].

The sensor validation method is performed as follows:

1. Development of the model

This consists of identifying matrices $A(t)$, $L(t)$, $\Xi(t)$, $C(t)$, $\Theta(t)$ and the order n of the system under normal operating conditions. This identification is mostly done by using a combination of physical information and statistical data on the system. The various methods for system identification and model validation are useful at this stage.

2. Generation of the innovation sequence

The innovation sequence $r(t+1)$ is defined as:

$$r(t+1) = z(t+1) - \hat{z}(t+1\,|\,t) \tag{11}$$

where $\hat{z}(t+1\,|\,t)$ denotes the unbiased minimum variance estimate of $z(t+1)$ given the sequence of past measurements up to (t), i.e. based on the set $\{z(1), z(2), \ldots, z(t)\}$. If it is assumed that all the system parameters and statistics are known exactly, the innovation sequence can be generated by a Kalman filter of the following form [1]:

Off-line Calculations

Initialization (t=0):

$$\Sigma(0\,|\,0) = \operatorname{cov}[x(0); x(0)] \tag{12}$$

Predict Cycle:

$$\Sigma(t+1\,|\,t) = A(t)\Sigma(t\,|\,t)A^T(t) + L(t)\Xi(t)L^T(t) \tag{13}$$

Update Cycle:

$$\begin{aligned}\Sigma(t+1\,|\,t+1) = \Sigma(t+1\,|\,t) - \Sigma(t+1\,|\,t)C^T(t+1)\left[C(t+1)\Sigma(t+1\,|\,t)C^T(t+1) + \Theta(t+1)\right]^{-1}\cdot \\ \cdot C(t+1)\Sigma(t+1\,|\,t)\end{aligned} \tag{14}$$

where

$\Sigma(t\,|\,\tau)$ is the error covariance of $\hat{x}(t\,|\,\tau)$;

$\hat{x}(t\,|\,\tau)$ is the unbiased minimum variance estimate of $\hat{x}(t)$ given the measurements up to time τ.

Filter Gain Matrix

$$H(t+1) = \Sigma(t+1\,|\,t+1)C^T(t+1)\Theta^{-1}(t+1) \tag{15}$$

where

$H(t+1)$ is $\Re^{n\times r}$ Kalman gain matrix

On-Line Calculations

Initialization (t=0):

$$\hat{x}(0\,|\,0) = \mathrm{E}[x(0)] \tag{16}$$

Predict Cycle:

$$\hat{x}(t+1\,|\,t) = A(t)\hat{x}(t\,|\,t) \tag{17}$$

Update Cycle:

$$r(t+1) = z(t+1) - C(t+1)\hat{x}(t+1\,|\,t) \tag{18}$$

$$\hat{x}(t+1\,|\,t+1) = \hat{x}(t+1\,|\,t) + H(t+1)r(t+1) \tag{19}$$

3. Statistics of the innovation sequence

It is well known that the innovation sequence $r(t+1)$ is a zero mean Gaussian white noise sequence

$$\mathrm{E}[r(t+1)] = 0 \tag{20}$$

with covariance

$$\begin{aligned} S(t+1) &\equiv \mathrm{cov}[r(t+1); r(t+1)] \\ &= C(t+1)\Sigma(t+1\,|\,t)C^T(t+1) + \Theta(t+1) \end{aligned} \tag{21}$$

$$\mathrm{cov}[r(t); r(\tau)] = S(t)\delta_{t\tau} \tag{22}$$

For testing purposes, it is more convenient to consider the Standardized Innovation Sequence:

$$\eta(t+1) = \left(C(t+1)\Sigma(t+1\,|\,t)C^T(t+1) + \Theta(t+1)\right)^{-\frac{1}{2}} r(t+1) \tag{23}$$

where $(\cdot)^{-\frac{1}{2}}$ denotes the square root of the inverse of a matrix. Then

$$\mathrm{cov}[\eta(t); \eta(\tau)] = \mathrm{E}[\eta(t)\eta^T(\tau)] = I\delta_{t\tau} \quad (24)$$

where I denotes the identity matrix.

4. Sensor fault detection via hypothesis testing

Different kinds of faults can develop in the system. Some of these are:

- bias errors in instruments,
- noisy instruments,
- change in system parameters,
- change in level of input noise,

- change in the structure of the system, etc.

All these faults make the standardized innovation $\eta(t+1)$, depart from their zero mean, unit variance and whiteness properties. Therefore it is useful to perform the following statistical tests:

4.1. Tests of whiteness

The most important property of the innovation sequence is whiteness or independence at different time instants. Most of the tests of independence are based on the autocorrelation function c_k of a stationary process for lag k = 1, 2, … as follows:

$$c_k = \mathrm{E}\left[(\eta_i - \overline{\eta})(\eta_{i-k} - \overline{\eta})^T\right] \tag{24}$$

where $\overline{\eta}$ denotes the mean of η_i.

c_k is often estimated as

$$\hat{c}_k = \frac{1}{N}\sum_{i=k}^{N}\left(\eta_i - \hat{\overline{\eta}}\right)\left(\eta_{i-k} - \hat{\overline{\eta}}\right)^T \tag{25}$$

where $\hat{\overline{\eta}}$ denotes the sample mean

$$\hat{\overline{\eta}} = \frac{1}{N}\sum_{i=1}^{N}\eta_i \tag{26}$$

It can be shown the $\hat{c}_k$ is an asymptotically unbiased and consistent estimate of c_k [3]. Under the null hypothesis, $\hat{c}_k$, k=l, 2, … are asymptotically independent and normal with zero mean and covariance of I/N. Thus they can be regarded as samples from the same normal distribution and must lie in the band $\pm 1.96/\sqrt{n}$ more than 95 per cent of the times for the null hypothesis [3].

4.2. Tests of mean

These tests check whether the observed innovation sequence is zero mean or not. The mean of the innovation sequence is estimated by (26). Under the null hypothesis, $\hat{\overline{\eta}}$ has a Gaussian distribution with zero mean and covariance

$$\mathrm{E}\left[\hat{\overline{\eta}}\hat{\overline{\eta}}^T\right] = I/N \tag{27}$$

Therefore at the 5 per cent significance level, the null hypothesis is rejected whenever

$$\left|\hat{\overline{\eta}}\right| > 1.96\, I/\sqrt{N} \tag{28}$$

4.3. Tests of covariance

The covariance of the innovation sequence is estimated as

$$\hat{c}_0 = \frac{1}{N}\sum_{i=1}^{N}\left(\eta_i - \hat{\overline{\eta}}\right)\left(\eta_i - \hat{\overline{\eta}}\right)^T \tag{29}$$

Under the the null hypothesis, $\hat{c}_0$ has a WISHART Distribution [4]. The trace of $\hat{c}_0$ has a Chi-Square distribution with (N-l)r degrees of freedom. Thus $\hat{c}_0$ can be tested for its null hypothesis with covariance equal to an identity matrix.

Both the tests of mean and covariance assume that the innovation sequence is white. Therefore, it is important to test the innovation sequence for whiteness first, especially using tests which are invariant with respect to the mean and covariance of the distribution.

5. References

[1] Mehra, R.K., Peschon, J. – **An Innovations Approach to Fault Detection and Diagnosis in Dynamic Systems**. U.K., Automatica, Vol. 7, pg 637-640, Pergamon Press, 1971

[2] Kalman, R. E. - **New methods and results in linear prediction and filtering theory**. Proc. Syrmp. on Engineering Applications of Random Function Theory and Probability: John Wiley, New York, 1961

[3] Jenkin , G. M. ; Watts, D. G. **- Spectral Analysis and its Applications**. Holden Day, San Francisco, 1968

[4] Anderson, T. W. - An Introduction to Multivariate Statistical Analysis. John Wiley, New York , 1958

[5] Bryson, A. E.; Ho, Y. C. **- Applied Optimal Control**: Blaisdell, Waltham, Mass, 1969

ANNEX II
Normality Tests

Normality Tests

We describe in this annex two implementations of the statistical hypothesis testing method (wich is included in the frequencist inference framework): the Anderson-Darling test and the Cramér–von Mises criterion. The data is tested in this method against the null hypothesis, that is: that it's normally distributed.

1. Anderson-Darling test

The A^2 Empirical distribution function statistics may be used with small sample sizes 5 ≤ n ≤ 25 [1]. Large sample sizes may reject the assumption of normality with only slight imperfection. The computation differs based on what is known about the distribution. For the proposed test, two of the four possible cases are considered:

- Case 1: The mean and the variance are both known.
- Case 2: Both the mean and the variance are unknown.

Although the parameters are known, the second case is considered because in [1] it is claimed that the test becomes better when the parameters are computed from the data, even if they are known.

In Case 2 the parameters can be estimated as

$$\hat{\bar{r}} = \frac{1}{N}\sum_{i=1}^{N} r_i \tag{1}$$

$$\hat{c}_{r_0} = \frac{1}{N-1}\sum_{i=1}^{N}\left(r_i - \hat{\bar{r}}_i\right)\left(r_i - \hat{\bar{r}}_i\right)^T \tag{2}$$

where $\hat{\bar{r}}$ denotes the innovation sample mean, $\hat{c}_0$ the innovation sample covariance, and the i index corresponds to the equivalent t+1 value. The values of the innovation sequence are then standardized according to the estimated parameters

$$\hat{\eta}(t+1) = \hat{c}_{r_0}^{\;-\frac{1}{2}}\left(r(t+1) - \hat{\bar{r}}\right) \tag{3}$$

The standardized Innovation Sequence $\eta(t+1)$ in case 1 or $\hat{\eta}(t+1)$ in case 2 is then sorted from low to high.

A^2 is then calculated by

$$A^2 = -N - \frac{1}{N}\sum_{i=1}^{N}\begin{bmatrix}(2i-1)(\ln(\Phi(\eta_i))) + \cdot \\ \cdot\,(2(N-i)+1)\ln(1-\Phi(\eta_i))\end{bmatrix} \tag{4}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

A modified statistic is calculated for Case 2:

$$A^{*2} = A^2\left(1+\frac{4}{N}-\frac{25}{N^2}\right) \tag{5}$$

At the 5 per cent significance level, the null hypothesis is rejected whenever

$$A^2 > 2.492 \tag{6}$$

$$A^{*2} > 0.787 \tag{7}$$

2. Cramér–von Mises criterion

As in the Anderson-Darling test the two cases of known and unknown parameters are considered. The test also follows the same procedure for the standardization using (1), (2), (3) and then performing the calculations on the sorted standardized Innovation Sequence.

W^2 is then calculated by

$$W^2 = \frac{1}{12N} + \sum_{i=1}^{N}\left(\Phi(\eta_i) - \frac{2i-1}{2N}\right)^2 \tag{8}$$

A modified statistic is then calculated for Case 1

$$W_1^{*2} = \left(W^2 - \frac{0.4}{N} + \frac{0.6}{N^2}\right)\left(1.0 + \frac{1.0}{N}\right) \tag{9}$$

and Case 2

$$W_2^{*2} = W^2\left(1 + \frac{0.5}{N}\right) \tag{10}$$

Then, at the 5 per cent significance level, the null hypothesis is rejected whenever

$$W_1^{*2} > 0.461 \tag{11}$$

$$W_2^{*2} > 0.126 \tag{12}$$

3. References

[1] Stephens, M. A. - **EDF Statistics for Goodness of Fit and Some Comparisons**. Journal of the American Statistical Association 69: 730-737, 1974