IDENTIFYING AN OBSERVABLE PROCESS WITH ONE OF SEVERAL SIMULATION MODELS VIA UMPI TEST

Nicholas A. Nechval, Konstantin N. Nechval, Edgars K. Vasermanis, Kristine Rozite
Department of Mathematical Statistics, University of Latvia, Raina Blvd 19, Riga LV-1019, Latvia

Keywords: Observable process, UMPI test, Identification

Abstract: In this paper, for identifying an observable process with one of several simulation models, a uniformly most powerful invariant (UMPI) test is developed from the generalized maximum likelihood ratio (GMLR). This test can be considered as a result of a new approach to solving the Behrens-Fisher problem when covariance matrices of multivariate normal populations (compared with respect to their means) are different and unknown. The test is based on invariant statistic whose distribution, under the null hypothesis, does not depend on the unknown (nuisance) parameters.

1 INTRODUCTION

Computational modeling has become an important tool for building and testing theories in Cognitive Science during the last years. The area of its applications includes, in particular, business process simulation, resource management, knowledge management systems, operations research, economics, optimization, stochastic models, logic programming, operation and production management, supply chain management, work flow management, total quality management, logistics, risk analysis, scheduling, forecasting, cost benefit analysis, economic revitalization, financial models, accounting, policy issues, regulatory impact analysis, etc. One of the most important steps in the development of a simulation model is recognition of the simulation model, which is an accurate representation of the process being studied. This procedure consists of two basic stages: (i) establishing the form of an adequate simulation model for the process under study and then (ii) estimating precisely the values of its parameters.

In developing strategies for the design of experiments for parameter estimation, it is customarily assumed that the correct form of the model is known. However, experimenters often do not have just one model known to be correct but have instead m>1 rival models to consider as possible explanations of the process being investigated. It is natural for model users to devise rules so as to identify an observable process with one of several distinct models, collected for simulation, which accurately represents the process, especially when decisions involving expensive resources are made on the basis of the results of the model.

Substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model is usually referred to as model validation and is the definition used in this paper.

Validation is defined in this paper following a classic simulation textbook (Law and Kelton, 1991, p. 299): "*Validation* is concerned with determining whether the conceptual simulation model (as opposed to the computer program) is an accurate representation of the system under study". Hence, validation cannot result in a perfect model: the perfect model would be the real system itself. Instead, the model should be 'good enough', which depends on the goals of the model. Validation is a central aspect to the responsible application of models to scientific and managerial problems. The importance of validation to those who construct and use models is well recognized. General discussions on validation of simulation models can be found in all textbooks on simulation. Examples are Banks and Carson (1984), Law and Kelton (1991), and Pegden, Shannon, and Sadowski (1990). A well-known article is Sargent (1991). Recent survey article is Kleijnen (1995), including 61 references.

Statistical hypothesis testing (Naylor and Finger, 1967), as distinguished from graphical or descriptive techniques, offers a framework that is particularly

Nechval N., Nechval K., Vasermanis E. and Rozite K. (2004).
IDENTIFYING AN OBSERVABLE PROCESS WITH ONE OF SEVERAL SIMULATION MODELS VIA UMPI TEST.
In *Proceedings of the First International Conference on Informatics in Control, Automation and Robotics*, pages 152-159
DOI: 10.5220/0001133101520159

attractive for model validation. A test would compare a sample of observations taken from the target population against a sample of predictions taken from the model. Not surprisingly, a number of statistical tools have been applied to validation problems. For example, Freese (1960) introduced an accuracy test based on the standard χ^2 tests. Ottosson and Håkanson (1997) used R^2 and compared with so-called highest-possible R^2, which are predictions from common units (parallel time-compatible sets). Jans-Hammermeister and McGill (1997) used an F-statistic-based lack of fit test. Landsberg et al. (2003) used R^2 and relative mean bias. Bartelink (1998) graphed field data and predictions with confidence intervals. Finally, Alewell and Manderscheid (1998) used R^2 and normalized mean absolute error (NMAE).

In practice, simulations are usually validated by considering not one but several output measures (e.g., expected waiting time, expected queue length, etc.). In this case, one could in principle validate the simulation for each output measure individually, as discussed previously. However, these output measures will in general be dependent. In some cases, it may be possible to model this dependence explicitly – e.g., using a multivariate normal distribution. The aim of this study was to develop and use criteria, which permit an objective comparison of different models to the observed field data and to each other. A given model, which describes a specific system significantly better, will be declared the 'valid' model while the other will be rejected. The term 'valid' is used here in a sense that any model that could not be proven invalid would be a valid model for the system.

Real plants are, in general, time-varying for various reasons, such as plant operating point changes, component aging, equipment wear, heat and material transfer degradation effects.

In this paper, we propose an effective technique for validation of simulation models (static or dynamic), performing the UMPI test for comparison of a real process data set and data sets of several simulation models.

2 TESTING THE VALIDITY OF A SIMULATION MODEL

Suppose that we desire to validate a kth multivariate stationary response simulation model of an observable process, which has p response variables. Let $x_{ij}(k)$ and y_{ij} be the ith observation of the jth response variable of the kth model and the process under study, respectively. It is assumed that all observation vectors, $\mathbf{x}_i(k)=(x_{i1}(k), \ldots, x_{ip}(k))'$, $\mathbf{y}_i=(y_{i1}, \ldots, y_{ip})'$, $i=1(1)n$, are independent of each other, where n is a number of paired observations. Let $\mathbf{z}_i(k)=\mathbf{x}_i(k)-\mathbf{y}_i$, $i=1(1)n$, be paired comparisons leading to a series of vector differences. Thus, for testing the validity of a simulation model of a real, observable process, it can be obtained and used a sample of n independent observation vectors $\mathbf{Z}(k)=(\mathbf{z}_1(k), \ldots ,\mathbf{z}_n(k))$. Each sample $\mathbf{Z}(k)$, $k\in\{1, \ldots, m\}$, is declared to be realization of a specific stochastic process with unknown parameters.

In this paper, for testing the validity of the kth simulation model of a real, observable process, we propose a statistical approach that is based on the generalized maximum likelihood ratio. In using statistical hypothesis testing to test the validity of a simulation model under a given experimental frame and for an acceptable range of accuracy consistent with the intended application of the model, we have the following hypotheses:

$H_0(k)$: the kth model is valid for the acceptable range of accuracy under a given experimental frame;
$H_1(k)$: the kth model is invalid for the acceptable range of accuracy under a given experimental frame. (1)

There are two possibilities for making a wrong decision in statistical hypothesis testing. The first one, type I error, is accepting the alternative hypothesis $H_1(k)$ when the null hypothesis $H_0(k)$ is actually true, and the second one, type II error, is accepting the null hypothesis when the alternative hypothesis is actually true. In model validation, the first type of wrong decision corresponds to rejecting the validity of the model when it is actually valid, and the second type of wrong decision corresponds to accepting the validity of the model when it is actually invalid. The probability of making the first type of wrong decision will be called model builder's risk ($\alpha(k)$) and the probability of making the second type of wrong decision will be called model user's risk ($\beta(k)$). Thus, for fixed n, the problem is to construct a test, which consists of testing the null hypothesis

$$H_0(k)\text{: } \mathbf{z}_i(k) \sim N_p(\mathbf{0},\mathbf{Q}(k)), \quad \forall i = 1(1)n, \tag{2}$$

where $\mathbf{Q}(k)$ is a positive definite covariance matrix, versus the alternative

$$H_1(k)\text{: } \mathbf{z}_i(k) \sim N_p(\mathbf{a}(k),\mathbf{Q}(k)), \quad \forall i = 1(1)n, \tag{3}$$

where $\mathbf{a}(k)=(a_1(k), \ldots ,a_p(k))'\neq(0, \ldots ,0)'$ is a mean vector. The parameters $\mathbf{Q}(k)$ and $\mathbf{a}(k)$ are unknown.

It will be noted that the result of Theorem 1 given below can be used to obtain test for the hypothesis of the form H_0: $\mathbf{z}_i(k)$ follows

$N_p(\mathbf{a}(k),\mathbf{Q}(k))$ versus H_a: $\mathbf{z}_i(k)$ does not follow $N_p(\mathbf{a}(k),\mathbf{Q}(k))$, $\forall i=1(1)n$. The general strategy is to apply the probability integral transforms of w_k, $\forall k=p+2(1)n$, to obtain a set of i.i.d. $U(0,1)$ random variables under H_0 (Nechval, 1998b). Under H_a this set of random variables will, in general, not be i.i.d. $U(0,1)$. Any statistic, which measures a distance from uniformity in the transformed sample (say, a Kolmogorov-Smirnov statistic), can be used as a test statistic.

Theorem 1 (*Characterization of the Multivariate Normality*)**.** Let $\mathbf{z}_i(k)$, $i=1(1)n$, be n independent p-multivariate random variables ($n\geq p+2$) with common mean $\mathbf{a}(k)$ and covariance matrix (positive definite) $\mathbf{Q}(k)$. Let $w_r(k)$, $r=p+2, \ldots, n$, be defined by

$$w_r(k) = \frac{r-(p+1)}{p}\frac{r-1}{r}$$

$$\times \left(\mathbf{z}_r(k)-\overline{\mathbf{z}}_{r-1}(k)\right)'\mathbf{S}_{r-1}^{-1}(k)\left(\mathbf{z}_r(k)-\overline{\mathbf{z}}_{r-1}(k)\right)$$

$$= \frac{r-(p+1)}{p}\left(\frac{|\mathbf{S}_r(k)|}{|\mathbf{S}_{r-1}(k)|}-1\right), \quad r=p+2,\ldots,n, \quad (4)$$

where

$$\overline{\mathbf{z}}_{r-1}(k) = \sum_{i=1}^{r-1}\mathbf{z}_i(k)/(r-1), \quad (5)$$

$$\mathbf{S}_{r-1}(k) = \sum_{i=1}^{r-1}(\mathbf{z}_i(k)-\overline{\mathbf{z}}_{r-1}(k))(\mathbf{z}_i(k)-\overline{\mathbf{z}}_{r-1}(k))', \quad (6)$$

then the $\mathbf{z}_i(k)$ ($i=1, \ldots, n$) are $N_p(\mathbf{a}(k),\mathbf{Q}(k))$ if and only if $w_{p+2}(k), \ldots, w_n(k)$ are independently distributed according to the central F distribution with p and $1, 2, \ldots, n-(p+1)$ degrees of freedom, respectively.

Proof. The proof is similar to that of the characterization theorems (Nechval et al., 1998a, 2000) and so it is omitted here. □

3 GMLR STATISTIC

In order to distinguish the two hypotheses ($H_0(k)$ and $H_1(k)$), a generalized maximum likelihood ratio (GMLR) statistic is used. The GMLR principle is best described by a likelihood ratio defined on a sample space Z with a parameter set Θ, where the probability density function of the sample data is maximized over all unknown parameters, separately for each of the two hypotheses. The maximizing parameter values are, by definition, the maximum likelihood estimators of these parameters; hence the maximized probability functions are obtained by replacing the unknown parameters by their maximum likelihood estimators. Under $H_0(k)$, the ratio of these maxima is a $\mathbf{Q}(k)$-free statistic. This is shown in the following.

Let the complete parameter space for $\theta(k)=(\mathbf{a}(k),\mathbf{Q}(k))$ be $\Theta=\{(\mathbf{a}(k),\mathbf{Q}(k))$: $\mathbf{a}(k)\in\mathbf{R}^p$, $\mathbf{Q}(k)\in\mathcal{Q}_p\}$, where $\mathcal{Q}_p$ is a set of positive definite covariance matrices, and let the restricted parameter space for $\theta(k)$, specified by the $H_0(k)$ hypothesis, be $\Theta_0=\{(\mathbf{a}(k),\mathbf{Q}(k))$: $\mathbf{a}(k)=\mathbf{0}$, $\mathbf{Q}(k)\in\mathcal{Q}_p\}$. Then one possible statistic for testing $H_0(k)$: $\theta(k)\in\Theta_0$ versus $H_1(k)$: $\theta(k)\in\Theta_1$, where $\Theta_1=\Theta-\Theta_0$, is given by the generalized maximum likelihood ratio

$$LR = \frac{\max\limits_{\theta(k)\in\Theta_1} \mathrm{L}_{H_1(k)}(\mathbf{Z}(k);\theta(k))}{\max\limits_{\theta(k)\in\Theta_0} \mathrm{L}_{H_0(k)}(\mathbf{Z}(k);\theta(k))}. \quad (7)$$

Under $H_0(k)$, the joint likelihood for $\mathbf{Z}(k)$ is given by

$$L_{H_0(k)}(\mathbf{Z}(k);\theta(k)) = (2\pi)^{-np/2}|\mathbf{Q}(k)|^{-n/2}$$

$$\times \exp\left(-\sum_{i=1}^{n}\mathbf{z}_i'(k)[\mathbf{Q}(k)]^{-1}\mathbf{z}_i(k)/2\right). \quad (8)$$

Under $H_1(k)$, the joint likelihood for $\mathbf{Z}(k)$ is given by

$$L_{H_1(k)}(\mathbf{Z}(k);\theta(k)) = (2\pi)^{-np/2}|\mathbf{Q}(k)|^{-n/2}$$

$$\times \exp\left(-\sum_{i=1}^{n}(\mathbf{z}_i(k)-\mathbf{a}(k))'[\mathbf{Q}(k)]^{-1}(\mathbf{z}_i(k)-\mathbf{a}(k))/2\right). \quad (9)$$

It can be shown that

$$\max_{\theta(k)\in\Theta_0} L_{H_0(k)}(\mathbf{Z}(k);\theta(k))$$

$$= (2\pi)^{-np/2}\left|\hat{\mathbf{Q}}_0(k)\right|^{-n/2}\exp(-np/2) \quad (10)$$

and

$$\max_{\theta(k)\in\Theta_1} L_{H_1(k)}(\mathbf{Z}(k);\theta(k))$$

$$= (2\pi)^{-np/2}\left|\hat{\mathbf{Q}}_1(k)\right|^{-n/2}\exp(-np/2), \quad (11)$$

where

$$\hat{\mathbf{Q}}_0(k) = \mathbf{Z}(k)\mathbf{Z}'(k)/n, \quad (12)$$

$$\hat{\mathbf{Q}}_1(k) = (\mathbf{Z}(k)-\hat{\mathbf{a}}(k)\mathbf{u}')(\mathbf{Z}(k)-\hat{\mathbf{a}}(k)\mathbf{u}')'/n, \quad (13)$$

and $\hat{\mathbf{a}}$(k)=$\mathbf{Z}$(k)$\mathbf{u}$/$\mathbf{u}'\mathbf{u}$ are the well-known maximum likelihood estimators of the unknown parameters $\mathbf{Q}$(k) and $\mathbf{a}$(k) under the hypotheses H_0(k) and H_1(k), respectively, $\mathbf{u}$=(1,...,1)′ is the n-dimensional column vector of units. A substitution of (10) and (11) into (7) yields

$$LR = \left|\hat{\mathbf{Q}}_0(k)\right|^{n/2}\left|\hat{\mathbf{Q}}_1(k)\right|^{-n/2}. \tag{14}$$

Taking the (n/2)th root, this likelihood ratio is evidently equivalent to

$$LR_\bullet = \left|\hat{\mathbf{Q}}_0(k)\right|\left|\hat{\mathbf{Q}}_1(k)\right|^{-1}$$

$$= \left|\mathbf{Z}(k)\mathbf{Z}'(k)\right| / \left|\mathbf{Z}(k)\mathbf{Z}'(k) - (\mathbf{Z}(k)\mathbf{u})(\mathbf{Z}(k)\mathbf{u})'/\mathbf{u}'\mathbf{u}\right|. \tag{15}$$

Now the likelihood ratio in (15) can be considerably simplified by factoring out the determinant of the $p \times p$ matrix $\mathbf{Z}(k)\mathbf{Z}'(k)$ in the denominator to obtain this ratio in the form

$$LR_\bullet = \frac{\left|\mathbf{Z}(k)\mathbf{Z}'(k)\right|}{\left[\left|\mathbf{Z}(k)\mathbf{Z}'(k)\right|\left(1 - \frac{(\mathbf{Z}(k)\mathbf{u})'[\mathbf{Z}(k)\mathbf{Z}'(k)]^{-1}(\mathbf{Z}(k)\mathbf{u})}{\mathbf{u}'\mathbf{u}}\right)\right]}$$

$$= 1\Big/\left(1 - (\mathbf{Z}(k)\mathbf{u})'[\mathbf{Z}(k)\mathbf{Z}'(k)]^{-1}(\mathbf{Z}(k)\mathbf{u})/n\right). \tag{16}$$

This equation follows from a well-known determinant identity. Clearly (16) is equivalent finally to the statistic

$$v_n(k) = \left(\frac{n-p}{p}\right)(LR_\bullet - 1)$$

$$= \left(\frac{n-p}{p}\right) n\hat{\mathbf{a}}'(k)[\mathbf{T}(k)]^{-1}\hat{\mathbf{a}}(k), \tag{17}$$

where $\mathbf{T}(k) = n\hat{\mathbf{Q}}_1(k)$. It is known that $(\hat{\mathbf{a}}(\mathrm{k}), \mathbf{T}(\mathrm{k}))$ is a complete sufficient statistic for the parameter $\theta(k)$=($\mathbf{a}(k)$,$\mathbf{Q}(k)$). Thus, the problem has been reduced to consideration of the sufficient statistic $(\hat{\mathbf{a}}(k), \mathbf{T}(k))$. It can be shown that under H_0, $v_n(k)$ is a $\mathbf{Q}(k)$-free statistic which has the property that its distribution does not depend on the actual covariance matrix $\mathbf{Q}(k)$. This is given by the following theorem.

Theorem 2 (*PDF of the Statistic $v_n(k)$*). Under $H_1(k)$, the statistic $v_n(k)$ is subject to a noncentral F-distribution with p and n–p degrees of freedom, the probability density function of which is

$$f_{H_1(k)}(v_n(k); n, q)$$

$$= \left[B\left(\frac{p}{2}, \frac{n-p}{2}\right)\right]^{-1}\left[\frac{p}{n-p}\right]^{\frac{p}{2}} v_n(k)^{\frac{p}{2}-1}$$

$$\times \left[1 + \frac{p}{n-p}v_n(k)\right]^{-\frac{n}{2}}$$

$$\times\ \mathrm{e}^{-nq/2}\,{}_1F_1\left(\frac{n}{2}; \frac{p}{2}; \frac{nq(k)}{2}\left[1 + \frac{n-p}{pv_n(k)}\right]^{-1}\right),$$

$$0 < v_n(k) < \infty. \tag{18}$$

where ${}_1F_1(b;c;x)$ is the confluent hypergeometric function, $q(k)$=$\mathbf{a}'(k)[\mathbf{Q}(k)]^{-1}\mathbf{a}(k)$ is a noncentrality parameter. Under $H_0(k)$, when $q(k)$=0, (18) reduces to a standard F-distribution with p and n–p degrees of freedom,

$$f_{H_0(k)}(v_n(k); n) = \left[B\left(\frac{p}{2}, \frac{n-p}{2}\right)\right]^{-1}$$

$$\times\left[\frac{p}{n-p}\right]^{\frac{p}{2}} v_n(\mathrm{k})^{\frac{p}{2}-1}\left[1 + \frac{p}{n-p}v_n(k)\right]^{-\frac{n}{2}},$$

$$0 < v_n(k) < \infty. \tag{19}$$

Proof. The proof follows by applying Theorem 1 (Nechval, 1997a, 1999) and being straightforward is omitted. □

4 GMLR TEST

The GMLR test of $H_0(k)$ versus $H_1(k)$, based on $v_n(k)$, is given by

$$v_n(k)\begin{cases} \geq h(k), & \text{then } H_1(k), \\ < h(k), & \text{then } H_0(k), \end{cases} \tag{20}$$

and can be written in the form

$$d(v_n(k)) = \begin{cases} 1, & \text{if } v_n(k) \geq h(k) \quad (H_1(k)), \\ 0, & \text{if } v_n(k) < h(k) \quad (H_0(k)), \end{cases} \tag{21}$$

where $h(k)>0$ is a threshold of the test which is uniquely determined for a prescribed level of significance $\alpha(k)$ so that

$$\sup_{\theta(k)\in\Theta_0} E_\theta\{d(v_n(k))\} = \alpha(k). \tag{22}$$

When the parameter $\theta(k)=(\mathbf{a}(k),\mathbf{Q}(k))$ is unknown, it is well known that no the uniformly most powerful (UMP) test exists for testing $H_0(k)$ versus $H_1(k)$ (Nechval, 1997b). However, it can be shown that the test (20) is UMPI for a natural group of transformations on the space of observations. Here the following theorem holds.

Theorem 3 (*UMPI Test*). For testing the hypothesis $H_0(k)$: $q(k)=0$ versus the alternative $H_1(k)$: $q(k)>0$, the test given by (20) is UMPI.

Proof. The proof is similar to that of Nechval (1997b) and so it is omitted here. □

5 ROBUSTNESS PROPERTY

In what follows, as one more optimality of the v_n-test, a robustness property can be studied in the following set-up. Let $\mathbf{Z}(k)=(\mathbf{z}_1(k), ..., \mathbf{z}_n(k))'$ be an $n \times p$ random matrix with a PDF φ, let $\mathcal{C}_{np}$ be the class of PDF's on $\mathbf{R}^{np}$ with respect to Lebesque measure $d\mathbf{Z}(k)$, and let $\mathcal{H}$ be the set of nonincreasing convex functions from $[0,\infty)$ into $[0,\infty)$. We assume $n \geq p+1$. For $\mathbf{a}(k)\in \mathbf{R}^p$ and $\mathbf{Q}(k)\in \mathcal{Q}_p$, define a class of PDF's on $\mathbf{R}^{np}$ as follows:

$$\mathcal{C}_{np}(\mathbf{a}(k),\mathbf{Q}(k))$$

$$= \left\{ \begin{array}{l} f \in \mathcal{C}_{np} : f(\mathbf{Z}(k);\mathbf{a}(k),\mathbf{Q}(k)) = |\mathbf{Q}(k)|^{-n/2} \\ \times \eta\left(\sum_{i=1}^{n} (\mathbf{z}_i(k)-\mathbf{a}(k))'[\mathbf{Q}(k)]^{-1}(\mathbf{z}_i(k)-\mathbf{a}(k)) \right), \\ \eta \in \mathcal{H} \end{array} \right\}. \tag{23}$$

In this model, it can be considered the following testing problem:

$$H_0(k) : \varphi \in \mathcal{C}_{np}(\mathbf{0},\mathbf{Q}(k)), \mathbf{Q}(k) \in \mathcal{Q}_p \tag{24}$$

versus

$$H_1(k) : \varphi \in \mathcal{C}_{np}(\mathbf{a}(k),\mathbf{Q}(k)), \mathbf{a}(k) \neq \mathbf{0}, \mathbf{Q}(k) \in \mathcal{Q}_p, \tag{25}$$

and shown that v_n-test is UMPI. Clearly if ($\mathbf{z}_1(k)$, ..., $\mathbf{z}_n(k)$) is a random sample of $\mathbf{z}_i(k)\sim N_p(\mathbf{a}(k),\mathbf{Q}(k))$, $i=1(1)n$, or $\mathbf{Z}(k)\sim N_{np}(\mathbf{u}\mathbf{a}'(k),\mathbf{I}_n\otimes\mathbf{Q}(k))$, where $\mathbf{u}=(1, ..., 1)'\in \mathbf{R}^n$, the PDF φ of $\mathbf{Z}(k)$ belongs to $\mathcal{C}_{np}(\mathbf{a}(k),\mathbf{Q}(k))$. Further if $f(\mathbf{Z}(k);\mathbf{a}(k),\mathbf{Q}(k))$ belongs to $\mathcal{C}_{np}(\mathbf{a}(k),\mathbf{Q}(k))$, then

$$g_*(\mathbf{Z}(k);\mathbf{a}(k),\mathbf{Q}(k)) = \int_0^\infty f(\mathbf{Z}(k);\mathbf{a}(k),r\mathbf{Q}(k))dG_*(r) \tag{26}$$

also belongs to $\mathcal{C}_{np}(\mathbf{a}(k),\mathbf{Q}(k))$ where G_* is a distribution function on $(0,\infty)$, and so $\mathcal{C}_{np}(\mathbf{a}(k),\mathbf{Q}(k))$ contains the (np-dimensional) multivariate t-distribution, the multivariate Cauchy distribution, the contaminated normal distribution, etc. Here the following theorem holds.

Theorem 4 (*Robustness Property*). For the problem (24)-(25), $v_n(k)$-test is UMPI and the null distribution of $v_n(k)$ is F-distribution with p and n-p degrees of freedom.

Proof. The proof is similar to that of Nechval [1997b] and so it is omitted here. □

In other words, for any $\mathbf{Q}(k)\in\mathcal{Q}_p$ and any $\varphi\in\mathcal{C}_{np}(\mathbf{0},\mathbf{Q}(k))$, the null distribution of $v_n(k)$ is exactly the same as that when $\mathbf{Z}(k)\sim N_{np}(\mathbf{0},\mathbf{I}_n\otimes\mathbf{Q}(k))$, that is, the distribution of $v_n(k)$ under $H_0(k)$ is the F-distribution with p and $n-p$ degrees of freedom. In this sense, the $v_n(k)$-test is robust against departures from normality.

6 RISK MINIMIZATION

For fixed n, in terms of the above probability density functions in (18) and (19), the probability of making the first type of wrong decision (model builder's risk ($\alpha(k)$)) is found by

$$\alpha(k)[h(k);n] = \int_{h(k)}^\infty f_{H_0(k)}(v_n(k);n)dv_n(k) \tag{27}$$

and the probability of making the second type of wrong decision (model user's risk ($\beta(k)$)) by

$$\beta(k)[h(k);n,q(k)] = \int_0^{h(k)} f_{H_1(k)}(v_n(k);n,q(k))dv_n(k). \tag{28}$$

This implies that the model is a perfect representation of the process with respect to its mean behavior. Any value of **a**(k) will result in a value for $q(k)$ that is greater than zero. As the value of **a**(k) increases, the value of $q(k)$ will also increase. Hence, the noncentrality parameter $q(k)$ is the validity measure for the above test (20). Let us assume that for the purpose for which the simulation model is intended, the acceptable range of accuracy (or the amount of agreement between the model and the process) can be stated as $0\le q(k)\le q^{\bullet}(k)$, where $q^{\bullet}(k)$ is the largest permissible value. In the statistical validation of simulation models, for preassigned $n=n^{\bullet}$ ($n^{\bullet}>p$) determined by a data collection budget, if we let $w_{\alpha(k)}$ and $w_{\beta(k)}$ be the unit weight (cost) of the model builder's risk ($\alpha(k)$) and the model user's risk ($\beta(k)$), then the optimal threshold of test, $h^*(k)$, can be found by solving the following optimization problem:

Minimize:

$$R[h(k);n^{\bullet},q^{\bullet}(k)] = w_{\alpha(k)}\alpha(k)[h(k);n^{\bullet}]$$

$$+ w_{\beta(k)}\beta(k)[h(k);n^{\bullet},q^{\bullet}(k)] \tag{29}$$

Subject to:

$$h(k)\in(0,1), \tag{30}$$

where $R[h(k);n^{\bullet},q^{\bullet}(k)]$ is a risk representing the weighted sum of the model builder's risk and the model user's risk. It can be shown that $h^*(k)$ satisfies the equation

$$w_{\alpha(k)}f_{H_0(k)}(h^*(k);n^{\bullet})$$

$$= w_{\beta(k)}f_{H_1(k)}(h^*(k);n^{\bullet},q^{\bullet}(k)). \tag{31}$$

In the statistical validation of simulation models, the model user's risk is more important that the model builder's risk, so that $w_{\alpha(k)}\le w_{\beta(k)}$.

For instance, let us assume that p=10, $n^{\bullet}$=40, $q^{\bullet}(k)$=0.5, and $w_{\alpha(k)}=w_{\beta(k)}$=1. It follows from (31) that the optimal threshold $h^*(k)$ is equal to 0.365.

If the sample size of observations, n, is not bounded above, then the optimal value n^* of n can be defined as

$$n^* = \inf n:$$

$$\left(\begin{array}{l}\alpha(k)[h^*(k);n]+\beta(k)[h^*(k);n,q^{\bullet}(k)]\le r^{\bullet}(k),\\[2ex] h^*(k)=\arg\min\limits_{h(k)\in(0,1)} R[h(k);n,q^{\bullet}(k)]\end{array}\right), \tag{32}$$

where $r^{\bullet}(k)$ is a preassigned value of the sum of the kth model builder's risk and the kth model user's risk.

7 PROCESS IDENTIFICATION

Let us assume that there is available a sample of measurements of size n from each simulation model. The elements of a sample from the kth model are realizations of p-dimensional random variables $\mathbf{x}_i(k)$, i=1(1)n, for each $k\in\{1, \ldots, m\}$. We are investigating an observable process on the basis of the corresponding sample of size n of p-dimensional measurements $\mathbf{y}_i=(y_{i1}, \ldots ,y_{ip})'$, i=1(1)n. We postulate that this process can be identified with one of the m simulation models but we do not know with which one. The problem is to identify the observable process with one of the m specified simulation models. When there is the possibility that the observable process cannot be identified with one of the m specified simulation models, it is desirable to recognize this case.

Let $\mathbf{y}_i$ and $\mathbf{x}_i(k)$ be the ith observation of the process and kth model variable, $k\in\{1, \ldots, m\}$, respectively. It is assumed that all observation vectors, $\mathbf{y}_i=(y_{i1}, \ldots, y_{ip})'$, $\mathbf{x}_i(k)=(x_{i1}(k), \ldots, x_{ip}(k))'$, i=1(1)n, are independent of each other, where n is a number of paired observations. Let $\mathbf{z}_i(k)=\mathbf{x}_i(k)-\mathbf{y}_i$, i=1(1)n, be paired comparisons leading to a series of vector differences. Thus, for identifying the observable process with one of the m specified simulation models, it can be obtained and used samples of n independent observation vectors $\mathbf{Z}(k)=(\mathbf{z}_1(k), \ldots ,\mathbf{z}_n(k))$, k=1(1)m. It is assumed that under $H_0(k)$, $\mathbf{z}_i(k)\sim N_p(\mathbf{0},\mathbf{Q}(k))$, $\forall i$=1(1)n, where $\mathbf{Q}(k)$ is a positive definite covariance matrix. Under $H_1(k)$, $\mathbf{z}_i(k)\sim N_p(\mathbf{a}(k),\mathbf{Q}(k))$, $\forall i$=1(1)n, where $\mathbf{a}(k)=(a_1(k), \ldots, a_p(k))'\neq(0, \ldots ,0)'$ is a mean vector. The parameters $\mathbf{a}(k)$ and $\mathbf{Q}(k)$, $\forall k$=1(1)m, are unknown. For fixed n, the problem is to identify the observable process with one of the m specified simulation models. If the observable process cannot be identified with one of the m specified simulation models, it is desirable to recognize this case.

The test of $H_0(k)$ versus $H_1(k)$, based on the GMLR statistic $v_n(k)$, is given by (20). Thus, if

$v_n(k) \geq h(k)$ then the kth simulation model is eliminated from further consideration.

If $(m-1)$ simulation models are so eliminated, then the remaining model (say, kth) is the one with which the observable process may be identified.

If all simulation models are eliminated from further consideration, we decide that the observable process cannot be identified with one of the m specified simulation models.

If the set of simulation models not yet eliminated has more than one element, then we declare that the observable process may be identified with simulation model k^* if

$$k^* = \arg\max_{k \in D} (h(k) - v_n(k)), \tag{33}$$

where D is the set of simulation models not yet eliminated by the above test.

8 APPLICATION OF THE TEST

This section discusses an application of the above test to the following problem. An airline company operates more than one route. It has available more than one type of airplanes. Each type has its relevant capacity and costs of operation. The demand on each route is known only in the form of the sample data, and the question asked is: which aircraft should be allocated to which route in order to minimize the total cost (performance index) of operation? This latter involves two kinds of costs: the costs connected with running and servicing an airplane, and the costs incurred whenever a passenger is denied transportation because of lack of seating capacity. (This latter cost is "opportunity" cost.) We define and illustrate the use of the loss function, the cost structure of which is piecewise linear. Within the context of this performance index, we assume that a distribution function of the passenger demand on each route is known. Thus, we develop our discussion of the allocation problem in the presence of completely specified demand distributions. We formulate this problem in a probabilistic setting.

Let A_1, ..., A_g be the set of airplanes which company utilize to satisfy the passenger demand for transportation en routes 1, ..., h. It is assumed that the company operates m routes which are of different lengths, and consequently, different profitabilities. Let $f_{ij}^{(k)}(s)$ represent the probability density function of the passenger demand S for transportation en route j (j=1, ..., h) at the ith stage ($i \in 1, \ldots, n$) for the kth simulation model ($k \in \{1, \ldots, m\}$). It is required to minimize the expected total cost of operation (the performance index)

$$J_i(\mathbf{U}_i) = \sum_{j=1}^{h} \left[\sum_{r=1}^{g} w_{rij} u_{rij} + c_j \int_{Q_{ij}}^{\infty} (s - Q_{ij}) f_{ij}^{(k)}(s) ds \right] \tag{34}$$

subject to

$$\sum_{j=1}^{h} u_{rij} \leq a_{ri}, \quad r = 1, \ldots, g, \tag{35}$$

where

$$Q_{ij} = \sum_{r=1}^{g} u_{rij} q_{rj}, \quad j = 1, \ldots, h, \tag{36}$$

$\mathbf{U}_i = \{u_{rij}\}$ is the $g \times h$ matrix, u_{rij} is the number of units of airplane A_r allocated to the jth route at the ith stage, w_{rij} is the operation costs of airplane A_r for the jth route at the ith stage, c_j is the price of a one-way ticket for air travel en jth route, q_{rj} is the limited seating capacity of airplane A_r for the jth route, a_{ri} is available the number of units of airplane A_r at the ith stage.

Let us assume that $\mathbf{U}_i^* = \{u_{rij}^*\}$ is the optimal solution of the above-stated programming problem. Since information about the passenger demand is not known precisely, this result provides only approximate solution to a real airline system. To depict the real, observable airline system more accurately, the test proposed in this paper, might be employed to validate the results derived from the analytical model (34)-(36). In this case

$$Z_{ij}(k) = X_{ij}(k) - Y_{ij}, \quad j = 1(1)h, \quad \forall i \in \{1, \ldots, n\}, \tag{37}$$

where

$$X_{ij}(k) = c_j \left[\int_0^{Q_{ij}^*} s f_{ij}^{(k)}(s) ds + Q_{ij}^* \int_{Q_{ij}^*}^{\infty} f_{ij}^{(k)}(s) ds \right], \tag{38}$$

is the expected gain (ensured by the service of a passenger demand on the jth route at the ith stage) derived from the analytical model (34)-(36),

$$Q_{ij}^* = \sum_{r=1}^{g} u_{rij}^* q_{rj}, \quad j = 1, \ldots, h, \tag{39}$$

Y_{ij} is the real gain ensured by the service of a passenger demand on the jth route at the ith stage (an observation of the airline system response variable).

Thus, the methodology proposed in this paper allows one to determine whether the analytical model (34)-(36) is appropriate for minimizing the total cost of airline operation.

9 CONCLUSIONS

The main idea of this paper is to find a test statistic whose distribution, under the null hypothesis, does not depend on unknown (nuisance) parameters. This allows one to eliminate the unknown parameters from the problem.

The authors hope that this work will stimulate further investigation using the approach on specific applications to see whether obtained results with it are feasible for realistic applications.

ACKNOWLEDGMENTS

This research was supported in part by Grant No. 02.0918 and Grant No. 01.0031 from the Latvian Council of Science and the National Institute of Mathematics and Informatics of Latvia.

REFERENCES

Alewell, C. and Manderscheid, B., 1998. Use of objective criteria for the assessment of biogeochemical ecosystem models. *Ecol. Modelling*, **105**, 113–124.

Banks, J. and Carson, J.S., 1984. *Discrete-event System Simulation.* NJ : Prentice-Hall, Englewood Cliffs.

Bartelink, H.H., 1998. Radiation interception by forest trees: a simulation study on effects of stand density and foliage clustering on absorption and transmission. *Ecol. Modelling*, **105**, 213–225.

Freese, F., 1960. Testing accuracy. *Forest Sci.*, **6**, 139–145.

Jans-Hammermeister, D.C. and McGill, W.B., 1997. Evaluation of three simulation models used to describe plant residue decomposition in soil. *Ecol. Modelling*, **104**, 1–13.

Kleijnen, J.P.C., 1995. Verification and validation of simulation models. *European Journal of Operational Research*, **82**, 145-162.

Law, A.M. and W.D. Kelton, W.D., 1991. *Simulation Modeling and Analysis*. New York: McGraw-Hill.

Naylor, T.H. and Finger, J.M., 1967. Verification of computer simulation models. *Management Science*, **14**, 92-101.

Nechval, N.A., 1997a. Adaptive CFAR tests for detection of a signal in noise and deflection criterion. In: *Digital Signal Processing for Communication Systems*, T. Wysocki, H. Razavi, & B. Honary, eds. Kluwer Academic Publishers, 177-186.

Nechval, N.A., 1997b. UMPI test for adaptive signal detection. In: *Proc. SPIE 3068: Signal Processing, Sensor Fusion, and Target Recognition VI*, I. Kadar, ed. Orlando, Florida USA, Paper No. 3068-73, 12 pages.

Nechval, N.A. and Nechval, K.N., 1998a. Characterization theorems for selecting the type of underlying distribution. In: *Abstracts of Communications of the 7th Vilnius Conference on Probability Theory and Mathematical Statistics & the 22nd European Meeting of Statisticians* (Vilnius, Lithuania, August 12-18). TEV, 352-353.

Nechval, N.A., 1988b. A general method for constructing automated procedures for testing quickest detection of a change in quality control. *Computers in Industry*, **10**, 1988, 177-183.

Nechval, N.A. and Nechval, K.N., 1999. CFAR test for moving window detection of a signal in noise. In: *Proceedings of the 5th International Symposium on DSP for Communication Systems* (Perth-Scarborough, Australia, February 1-4). IEEE, 134-141.

Nechval, N.A., Nechval, K.N., and Vasermanis, E.K., 2000. Technique of testing for two-phase regressions. In: *Proceedings of the Second International Conference on Simulation, Gaming, Training and Business Process Reengineering in Operations*, (Riga, Latvia, September 21-23). RTU, 129-133.

Ottosson, F. and Håkanson, L., 1997. Presentation and analysis of a model simulating the pH response of lake liming. *Ecol. Modelling*, **104**, 89–111.

Pegden, C.P., Shannon, R.E., and Sadowski, R.P., 1990. *Introduction to Simulation using SIMAN*. New York: McGraw-Hill.

Sargent, R.G., 1991. Simulation model verification and validation. In: *Proceedings of the 1991 Winter Simulation Conference,* 37-47.