SMTDA 2010: Stochastic Modeling Techniques and Data Analysis
International Conference, *Chania, Crete, Greece, 8 - 11 June 2010*

Modelling of high speed railroad geometry ageing as a discrete-continuous process.

L. Quiroga, E. Schnieder
Institute for Traffic Safety and Automation Technologies
Technische Universität Braunschweig
Langer Kamp 8, 38106 Braunschweig
Tel: +49 531 391 3313
Fax: +49 531 391 5197
E-Mail: {quiroga|schnieder@iva.ing.tu-bs.de}

Key words: Railway track, deterioration, forecasting, hybrid systems

Travelling safely and comfortably on high speed railway lines requires excellent conditions of the whole railway infrastructure in general and of the railway track geometry in particular. The maintenance process required to achieve such excellent conditions is largely complex and expensive, demanding an increased amount of both human and technical resources. In this framework, choosing the right maintenance strategy is a very important issue. A reliable forecast of the railway geometry ageing process is indispensable for an optimised planning and scheduling of maintenance activities. For this reason the French railway operator SNCF has been measuring periodically the geometrical characteristics of its high speed network since its commissioning, i.e. for more than 20 years now. In this paper a hybrid system model to obtain such a forecasting is presented. The proposed method uses a "grey-box" approach: a model structure and its constraints are specified basing on previous knowledge of the process to be identified, then the set of parameter values which best fits the signal measurements is searched. As previous knowledge indicates that the process is non linear the parameters are searched by means of the Levenberg-Marquardt (LM) algorithm, an iterative technique that finds a local minimum of a function expressed as the sum of squares of nonlinear functions. Finally, the method is applied on real data of a French high speed TGV line and its results compared with those of benchmark approaches.

1 Introduction

Measuring and keeping railway geometry under control are fundamental tasks of the railway infrastructure maintenance process. Railway geometry is representative of the travelling comfort and the derailing risk, so if its deviation exceeds a certain limit value, the travelling speed on that sector must be reduced. Therefore, railway geometry is both a measure of travelling quality and safety.
For these reasons the French railway operator SNCF has been measuring periodically the geometrical characteristics of its high speed network since its commissioning, i.e. for more than 20 years now.

Figure 1 shows the measurements of the longitudinal levelling (in French *Nivellement Longitudinal*, NL) for a 1 Km track sector for the last 20 years. The NL parameter is the longitudinal mean deviation of rails respect to the ideal position, and it is considered representative of the general railway geometry deterioration [Meyer-Hirmer]. By default the deterioration grade increases with time, reflecting the track geometry deterioration. Thus decrements take place only when some maintenance activity is performed. In figure 1 two different types of maintenance activities are included: tamping (red bars) and grinding (grey bars). Their heights represent the fraction of the railway sector affected by the maintenance activity. Tamping yields a visually more obvious effect, generating a sudden drop in NL. On

the other hand, grinding yields a more subtle, still according to expert knowledge significant effect.

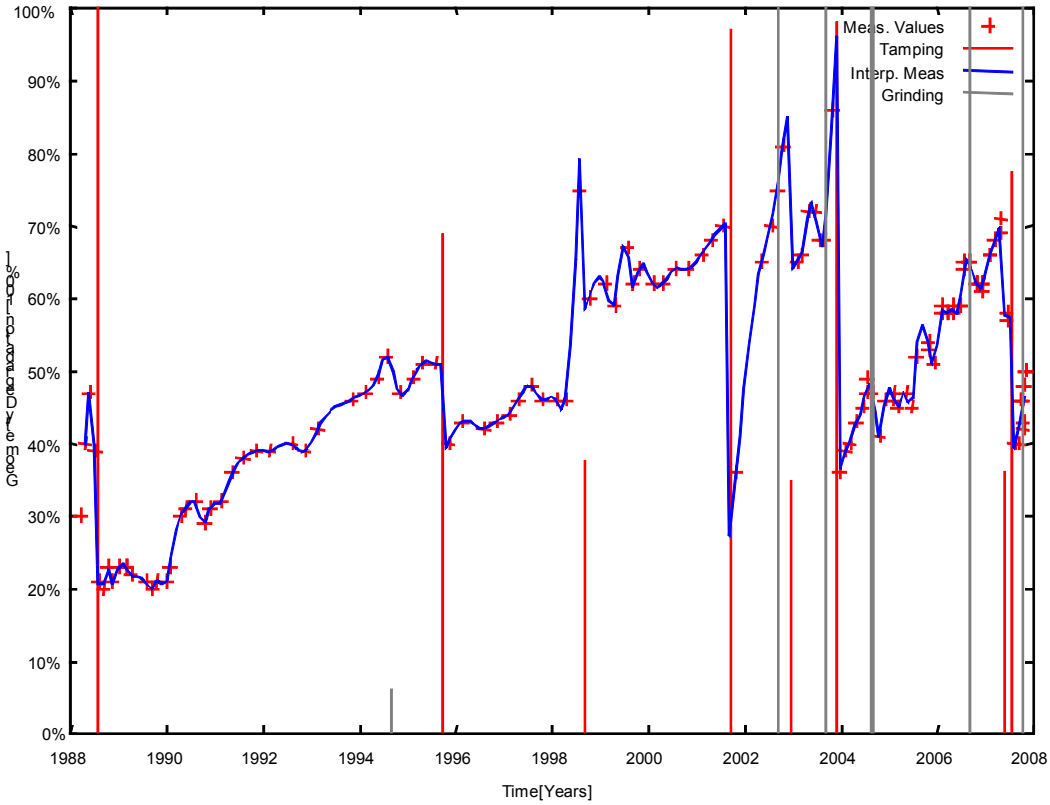


Figure 1: Course of longitudinal levelling degradation for a railroad sector.

Due to high logistic costs constraints, most track geometry maintenance activities need to be planned up to one year in advance. This is why a precise forecasting of the railway geometry is a key factor for effective maintenance activities planning, helping to answer the fundamental questions about **when** and **where** maintenance will be needed.
In this work we present some benchmark approaches to railway track geometry degradation and restoration forecasting, and we introduce an alternative discrete-continuous (hybrid) approach. In section 5 the these models are applied on real data and the obtained results are analysed and compared.

2 Problem definition

Railway geometry is measured periodically by means of special measuring coaches equipped with mechanical and/or electrical sensors. As it can be observed in figure 1, the periodicity of the measuring runs has been irregular since line commissioning, so the first problem for forecasting railway geometry deviation is the irregular sampling rate. To overcome this, we interpolate the measured points using splines, and then resample with the sampling rate of the last years. This is a compromise solution minimising information loss in the last measurement years and keeping the addition of artificial measurements in the first years at an acceptable level.

The resampled data is then used to tune a series of models expressed as grey boxes, i.e. models with variable parameters for which optimal values are to be found by means of an

optimisation process (see sections 3 and 4). The performances of the different grey boxes are then compared and the most suitable models are chosen.

Figure 2 shows a schematic overview of the complete forecasting model selection process hereby described.

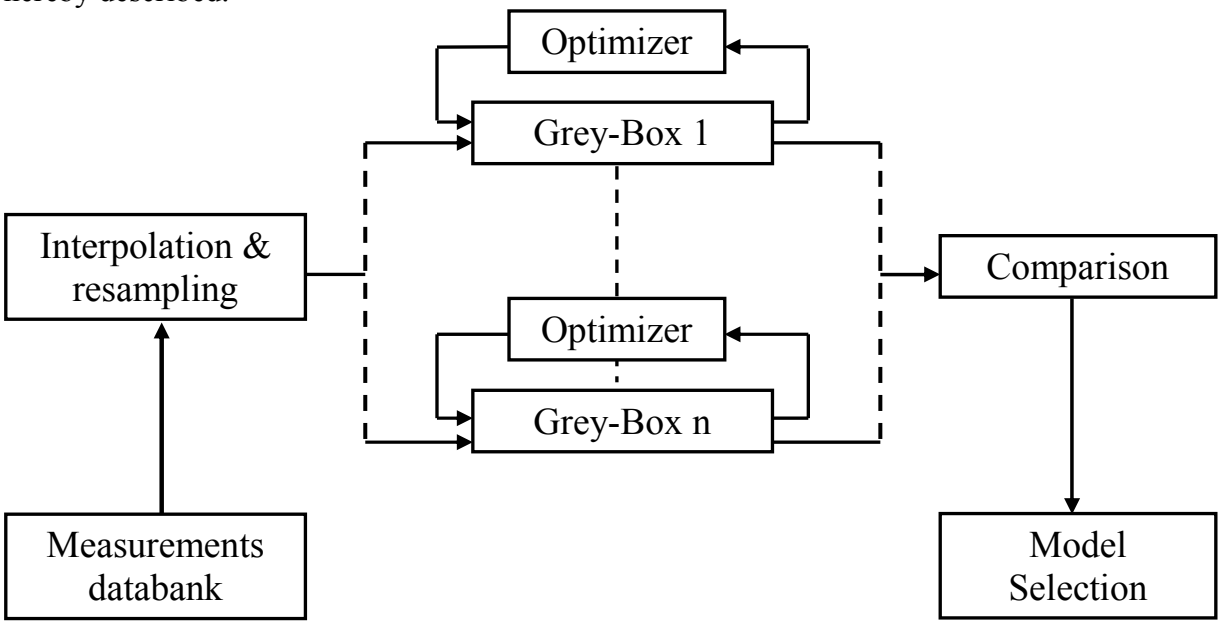


Figure 2: Schematic overview of the forecasting selection method

3 Applied models

In this section the implemented grey-box models are described. We consider railway deterioration being subject to two processes: deterioration and restoration (maintenance). Different models are proposed for each of them and then put together in grey-box models.

3.1 Restoration model

In contrast to the deterioration process, restoration process models are not numerous.
In this work we consider the non linear model proposed by [Miwa], described by equation 3.1

$$NL_{t+1} - NL_t = \beta_0 \cdot NL_t + \beta_1 \qquad (3.1)$$

In order to be able to apply this model on real cases, where sometimes track sectors are only partially restored (tamping is applied on only a fraction of the sector), we define

$$u_t = \frac{\text{length of tamped track at time } t}{\text{total sector length}}$$

as a signal with value 0 when no tamping takes place, and a value between 0 and 1 when tamping takes place in the sector, according to the fraction of the sector tamped. Then

$$NL_{t+1} = \beta_0 NL_t u_t + \beta_1 u_t + NL_t \qquad (3.2)$$

is used to forecast NL after tamping.

3.2 Deterioration models

The railway track geometry deterioration process has been deeply investigated in the last 30 years. As stated in [Sadeghi & Askarinejad 08], there are two main types of approaches: the *engineering* approach and the *statistical* approach.

The *engineering* approach aims to asses the mechanical properties of track degradation, providing a good understanding of how track responds to vehicle loading. In general, technical references agree that degradation depends on traffic intensity, travelling speed and axle load. Most of them aim to finding an equation describing deterioration rate as a function of these three variables. For an overview of the different formulas proposed see [Ubalde et al. 05].

On the other hand, the *statistical* approach aims to analysing observations of actual geometry deviations. The latter is then considered the dependant variable and the explanatory variables can be for example accumulated axle load or simply time. The aim of this work is not to define degradation rate as a function of known variables, but to forecast degradation at l steps in time, assuming that all other variables affecting degradation remain unchanged. This is in general a valid assumption for the aimed forecasting periods, namely up to 1 year. Therefore, this work belongs to the *statistical* approach.
We consider four different models: double exponential smoothing, generic, autoregressive and hybrid. In section 5 an analysis of their performances is presented.

3.2.1 Exponential smoothing

In [Miwa et al. 00] a double exponential smoothing based approach is proposed for modelling railway track geometry deterioration. Exponential smoothing is widely used for forecasting time series in the field of econometrics, and was developed by [Brown 62]. Double exponential smoothing assumes an only locally constant linear model of first order (locally constant linear trend), thus giving more weight to recent observations. At time n the parameters $\beta_{0,n}$ and $\beta_{1,n}$ are determined by minimizing

$$\sum_{j=0}^{n-1} \omega^j [z_{n-j} - (\beta_{0,n} - \beta_{1,n} j)]^2$$

The constant $\omega\,(|\omega| < 1)$ is a discount factor that discounts past observations exponentially.

At time n the l-step ahead forecasts $\hat{Z}(n+l)$ are calculated using

$$\hat{Z}(n+l) = \hat{\beta}_{0,n} + \hat{\beta}_{1,n} l \qquad (3.2)$$

For the model to also successfully explain the effects of degrading and restoration, it is necessary to combine the models expressed in equations equations 3.1 and 3.2. Figure 3 describes this combination. If the jumps caused by tamping are added to the original signal, a new signal is obtained which represents the effects of degradation only. On the other side these jumps seen as impulses make up another signal which represents the restoration process. Applying equation 3.1 on the restoration signal (here it is off course assumed that the dates of future tamping activities are known) and 3.2 on the degradation signal we obtain forecasts for both processes. Finally by combining again both signals, a forecast for the combined process is obtained.

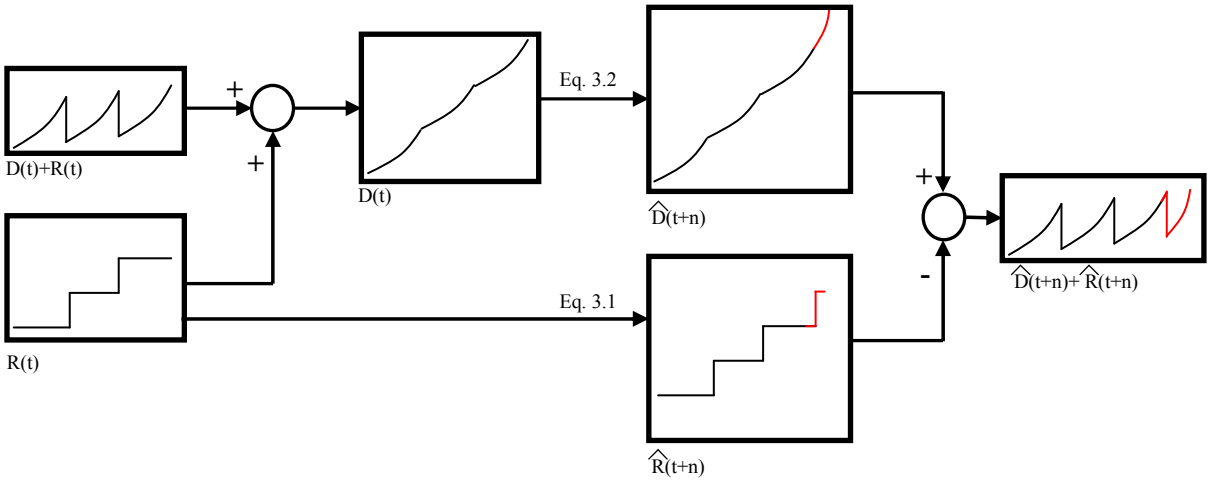


Figure 3: Procedure to decouple degradation and restoration processes

3.2.2 Polynomial

In [Jovanovic] a "generic/universal deterioration model" for railway track geometry is proposed. In this model deterioration is assumed to be a sum of processes running in parallel. Maintenance activities "reset" one or more of these processes. In this work we consider 2 parallel deterioration processes: one represents the deterioration which can not be corrected by means of tamping, but only by ballast renewal, and the other represents deterioration which is effectively corrected by tamping. According to this model, track geometry condition NL(t) can be described as

$$NL(t, t_{slt}) = \alpha_1 + \alpha_2 t + \alpha_3 t_{slt} + \alpha_4 t_{slt}^2$$

where t_{slt} is the elapsed time since the last tamping activity and $\alpha_1 ... \alpha_4$ are parameters for which adequate values must be found. This model explains both deterioration and restoration.

3.2.3 Autoregressive model

In [Hamid & Gross 81] the use of ARMA (autoregressive moving average) models for railway track geometry deterioration is proposed. To apply this model, we consider NL_t as a discrete signal sampled at regular intervals. Furthermore we define

$$U_t = \sum_{n=0}^{t} u_t$$

as the cumulated tamping.
Then we can define the ARIMA model used in this work as

$$NL_t = \alpha_1 NL_{t-1} + \alpha_2 NL_{t-2} + \alpha_3 U_{t-1} + \alpha_4 U_{t-2} + \alpha_5 u_t NL_{t-1} + \alpha_6$$

which is an AR model of second order plus the non-linear restoration model of equation 3.1.

3.2.4 Hybrid model

By observing track geometry degradation curves like the one of figure 1 and the performances of the forecasting model presented in 3.2.1, 3.2.2 and 3.2.3, some remarks can be asserted:

- The gradient of the curve (degradation speed) between two tamping activities remains constant or at least locally constant.
- The gradient of the curve very often changes abruptly after tamping.
- The forecasting models presented so far seem to work satisfactorily for the first years, but after some tamping activities become inappropriate.

[Lunze 06] argues that "...state jumps are the basic hybrid phenomenon that cannot be represented and analysed by methods elaborated either in continuous or discrete systems theory...". Having said this, we propose a hybrid model considering tamping activities as state jumps.
For modelling the degradation process, double exponential smoothing is used. But in contrast to the model presented in 3.2.1, after each tamping activity the measurements before tamping are forgotten, and a new set of parameters is identified. *This leads to a model adapting itself very quickly after each tamping, and very slowly between tampings*.

The algorithm for obtaining the *l*-step forecast $\hat{Z}(n+l)$ at time *n* can be described as follows:

1- Let M={m_0, ..., m_n} be all available measurements since the last tamping activity and T={t_0,...,t_n} their associated times, excluding those who are too near tamping activities, i.e. excluding all measurements $m_i : |t_i - t_{nt}| < \text{SETTLING\_TIME}$, where m_i is a given measurement, t_i its associated time and t_{nt} is the time of the tamping activity nearest to that measurement.
2- **If** size(M)>MIN_CYCLE_SIZE **then** apply double exponential smoothing with smoothing coefficient OMEGA to find the estimations $\hat{\beta}_{0,n}$ and $\hat{\beta}_{1,n}$, and go to step 5, **else** go to step 3.
3- **If** the current tamping cycle is not the first one **then** take the estimations $\hat{\beta}_{1,n}$ from the tamping cycle immediately before the current one, and $\hat{\beta}_{0,n}$ such that $m_i = \hat{\beta}_{0,n} + \hat{\beta}_{1,n} t_i$ holds, where m_i is the last measured value and t_i its associated time, and go to step 6, **else** go to step 4.
4- Take $\hat{\beta}_{1,n} = \text{INIT\_}\beta 1$ and $\hat{\beta}_{0,n}$ such that $m_i = \hat{\beta}_{0,n} + \hat{\beta}_{1,n} t_i$ holds.
5- Calculate $\hat{Z}(n+l) = \hat{\beta}_{0,n} + \hat{\beta}_{1,n} l$ as the forecast at *l* steps
6- **If** a tamping activity is planned within the next *l* steps, i.e. at step *n*+*k*, with *k*<*l*, **then** update the forecast $\hat{Z}(n+l)$ by subtracting to it the expected tamping effect given by $\Delta NL = (\hat{Z}(n+k) - \text{NL\_AFTER\_TAMPING})u_t$

The algorithm uses a series of parameters, for which an explanation is next given.
SETTLING_TIME: In approx. the first 2 months after and/or before tamping takes place, the behaviour of the track geometry can be strongly non linear [Meier-Hirmer], so the measurements taking place in this lapse are discarded.
MIN_CYCLE_SIZE: In order to perform a linear regression with a reasonable confidence, a minimal number of samples is required. In this case we take MIN_CYCLE_SIZE=5.

OMEGA: As mentioned in section 3.2.1, double exponential smoothing uses a smoothing constant ω to give more weight to recent observations. [Brown] ω recommends choosing ω such that $0.84 < \omega < 0.97$. In our work we take $\omega = 0.95$.
INIT_β0 and INIT_β1: When there are lesser than MIN_CYCLE_SIZE sample since the last tamping available, the line gradient $\hat{\beta}_{1,n}$ is taken from the last tamping cycle (we call *tamping cycle* the time between two consecutive tampings) and $\hat{\beta}_{0,n}$ is taken such that the curve determined by $\hat{\beta}_{0,n}$ and $\hat{\beta}_{1,n}$ contains the last measurement. But if the current tamping cycle is the first one, the only option left is to take an arbitrary initial value $\hat{\beta}_{1,n}$. In this work we use INIT_β1=0.05, i.e. an initial degradation rate of 0.05 mm. per year.
NL_AFTER_TAMPING: this is the typical value of NL after tamping a complete sector. In this work its value is set to 0.35.

4 Case study

In this section we apply the models presented in section 3 to real data measured on a TGV high speed line.
All forecasts are made recursively, i.e., forecasts of NL_{t+l} are made using information in time periods 1, 2, ..., t. For the forecast of NL_{t+l}, parameters are estimated using the data (NL_1, NL_2, ..., NL_t). In all models, parameters are estimated by minimizing the sum of squared residuals of the l-step ahead forecast. As we use a non linear restoration model (see 3.1), for solving these minimization problems [levmar], an implementation of the Levenberg-Marquardt algorithm, is used. The Levenberg-Marquardt (LM) algorithm is an iterative technique that finds a local minimum of a function that is expressed as the sum of squares of nonlinear functions.

Figures 4 to 7 show real measured values on a given track sector of 1 Km and the 1-year-ahead forecasts of the four models presented in 3.2,. Figure 4 shows the forecasts using the AR model. In general the forecast errors increase abruptly after each tamping, and after a lapse of between 1 and 4 years, the forecasts converge again to the measured values. In the case of exponential smoothing (figure 5), the observed behaviour is similar in the sense that error increases after tamping, but the forecast converge much more quickly, so the average error is significantly lesser. The generic model of figure 6 seems to fail to identify the nature of the degradation and restoration processes, being its performance very poor. Finally, the hybrid model shows an improved behaviour in the first year after each tamping, and after that its forecasts are very similar to those of the exponential smoothing model.

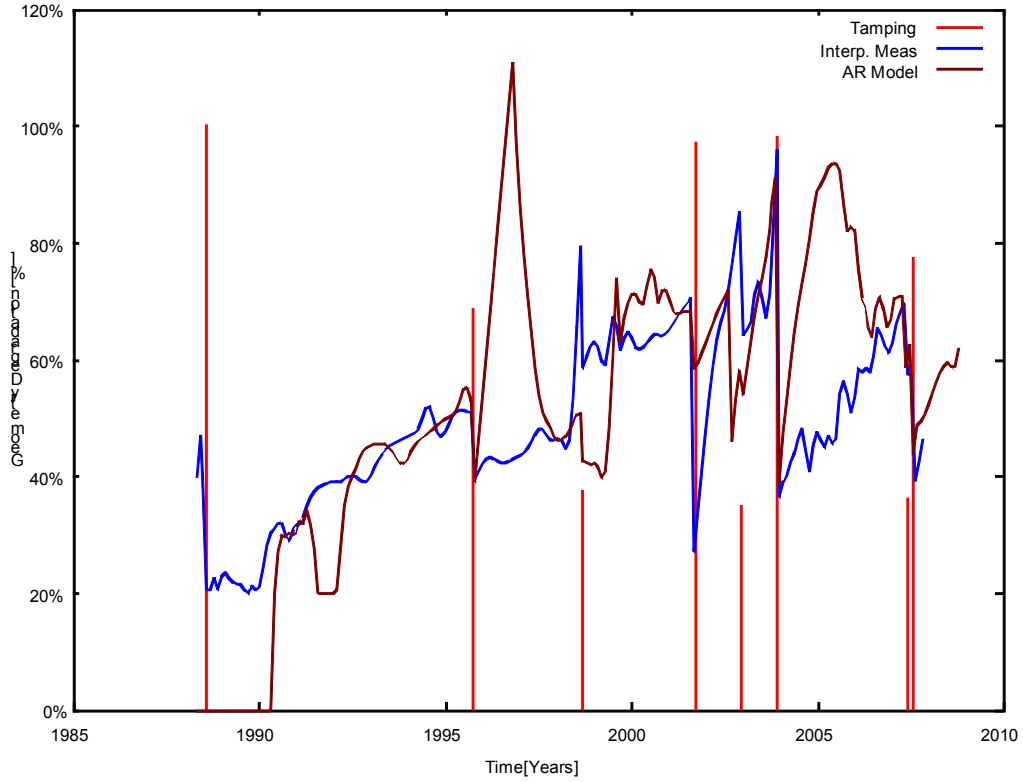


Figure 4: Measured data, tampings and 1-year-ahead forecasting using an AR model

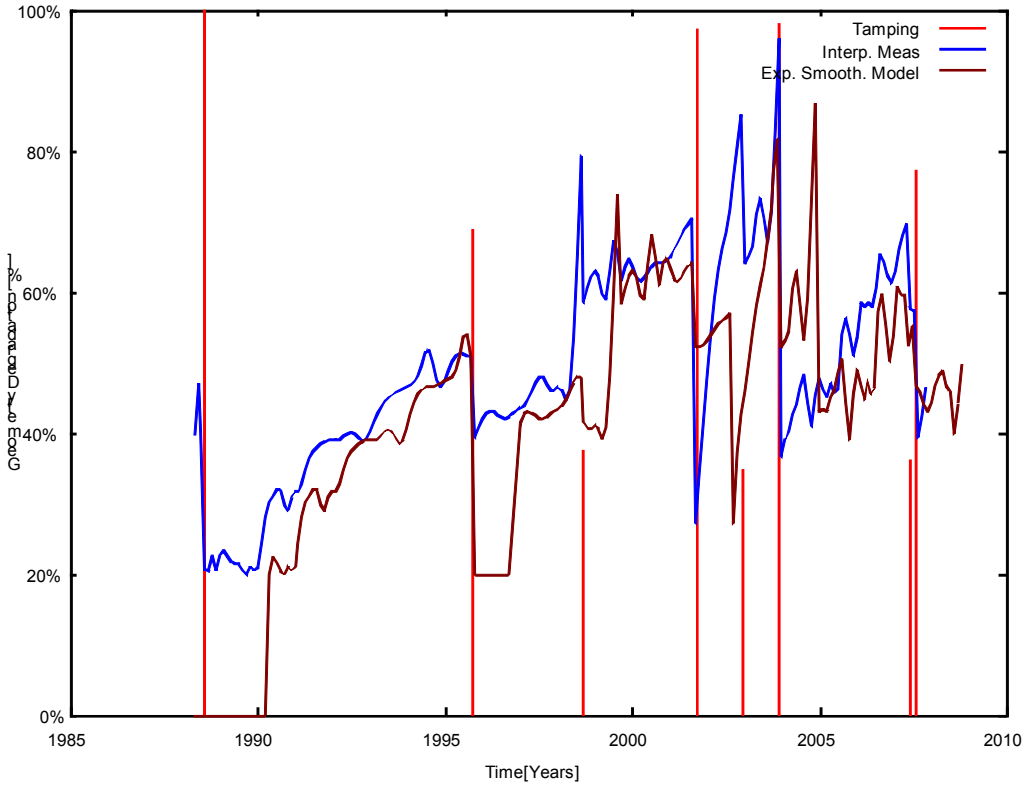


Figure 5: Measured data, tampings and 1-year-ahead forecasting using an exponential smoothing model

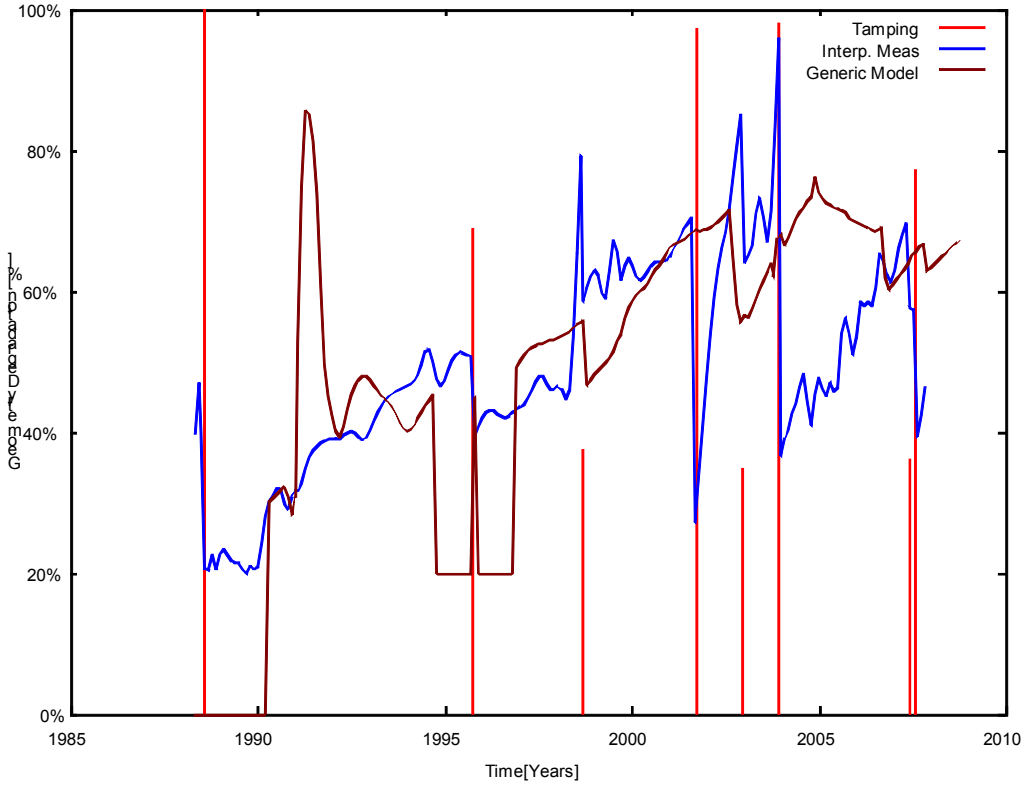


Figure 6: Measured data, tampings and 1-year-ahead forecasting using the generic degradation model

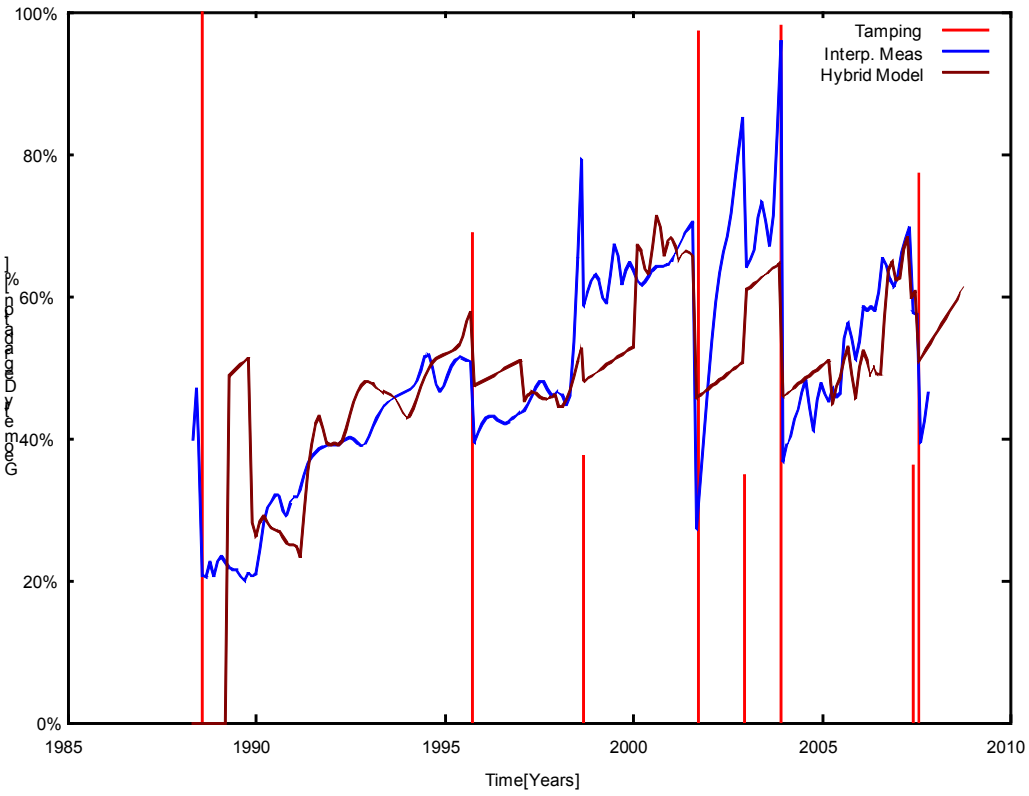


Figure 7: Measured data, tampings and 1-year-ahead forecasting using a hybrid model

Figure 8 compares the error signal $NL_t - \hat{NL}_t$ for the two models with best forecasting performance on this track sector, namely the exponential smoothing model and the hybrid model. Additionally, the times where tamping activities take place are marked with red bars in the same fashion as in figures 4 to 7. Here it becomes evident that error peaks are much smaller for the hybrid system model, and in the time zones where absolute error is under 0.1 both models behave similarly.

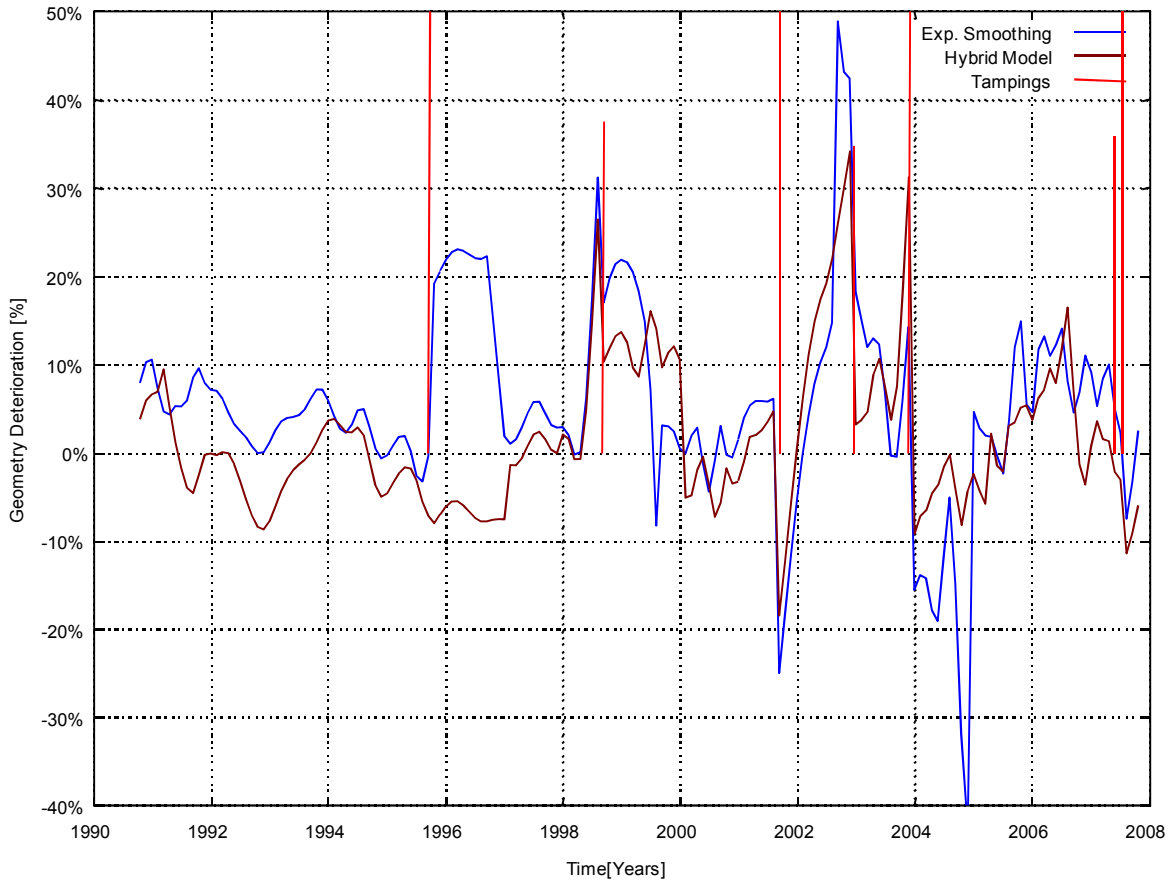


Figure 8: Comparison of the error signals of the hybrid model and the exp. smoothing model

Figures 4 to 8 should deliver a glimpse of how well each of the presented methods performs on railway track geometry. But degradation and restoration of track geometry can be very different from one sector to another. Therefore, in order to asses the prediction potential and robustness of all 4 methods, we apply them on real data of 200 sectors of 1 Km length each. We record for each model and each sector the mean absolute percentage error (MAPE) and mean square error (MSE) defined as

$$MAPE_i = \frac{1}{N_i}\sum_{n=0}^{N_i-1}\left|\frac{NL_{it} - \hat{NL}_{it}}{NL_{it}}\right| \qquad \text{and} \qquad MSE_i = \frac{\sum_{n=0}^{N_i-1}\left(NL_{it} - \hat{NL}_{it}\right)^2}{N_i}$$

where NL_{it} is the value measured on sector *i* at time *t*, $\hat{NL}_{it}$ its 1-year-ahead forecast and N_i is the number of samples available for sector *i*.

In table 1 the estimated mean value $\hat{\mu}$ and variance $\hat{\sigma}$ for both MAPE and MSE are expressed for each of the presented methods, plus the naive forecast $\hat{NL}_{t+l} = NL_t$. The means and variances are calculated as

$$\hat{\mu}(X) = \frac{\sum_{n=1}^{200} X_i}{200} \quad \text{and} \quad \hat{\sigma}(X) = \frac{\sum_{n=1}^{200} (X_i - \overline{X})^2}{200 - 1}$$

where X in this case are MAPE and MSE.
The last two rows of the table shows for each model the number of sectors for which it was better than all other methods, in the sense of MAPE and MSE, respectively.

Table 1. Overview of forecasting results on 200 1-Km-long track sectors with all 5 models

	Hybrid	AR	Generic	Exp. Smooth.	Naive
$\hat{\mu}(MAPE)$	16.2%	18.7%	45.9%	24.7%	19.6%
$\hat{\sigma}(MAPE)$	2.79×10^{-3}	3.59×10^{-3}	13.1×10^{-3}	4.18×10^{-3}	3.97×10^{-3}
$\hat{\mu}(MSE)$	0.0149	0.0193	0.116	0.0314	0.0225
$\hat{\sigma}(MSE)$	3.85×10^{-4}	3.44×10^{-4}	37.0×10^{-4}	3.93×10^{-4}	3.36×10^{-4}
Sectors Best MAPE	61.5%	16.5%	0%	10%	12%
Sectors Best MSE	66.5%	17%	0%	8%	6.5%

The minimal mean value for both MAPE (16.2%) and MSE (0.0149) are obtained using the hybrid model. The minimal MAPE variance (2.79×10^{-3}) is also achieved with the hybrid model, but in the case of MASE variance, the minimal value (3.44×10^{-4}) corresponds to the autoregressive (AR) model. The hybrid model is also the most appropriate in most of the track sectors, both from the point of view of MAPE (61.5%) and MSE (66.5%).
The generic model and the exponential smoothing models fail to improve the overall performance of the naive model, while the hybrid model achieves a MAPE reduction of 17.3% and an MSE reduction of 33.8% respect to the naive model, and the AR model achieves a MAPE reduction of 4.59% and an MSE reduction of 14.2% respect to the naive model.
The relatively low variances of MAPE and MSE achieved by the hybrid model show that the approach is robust in the sense that it can deliver satisfactory results in sectors with different degradation characteristics.

5 Conclusions

This paper presents an approach to railway track geometry forecasting modelling the degradation-restoration process as a hybrid system. The results obtained after applying this approach on the data collected in the lapse of almost 20 years on a 200 Km high speed railway track are compared with the ones obtained using some benchmark approaches. This comparison shows that the hybrid model in general achieves better results than the benchmark models, due mainly to its increased adaptability after tamping activities.
The approach hereby presented is intended to be used by the tamping scheduling optimization system under development at the Institute for Traffic Safety and Automation Technologies of the Technische Universität Braunschweig in cooperation with the SNCF.

6 References

[Brown 62] Brown, R.G. : Smoothing*, Forecasting and Prediction of Discrete Time Series*. Prentice-Hall, Englewood Cliffs, NJ.
[Hamid & Gross 81] Hamid, A., Gross, A.: Track-quality indices and track degradation models for maintenance-of-way planning. Transportation Research Board, 802 (1981), 2-8.
[Jovanovic 04] Jovanovic, S.: *Railway Track Quality Assessment and related Decision Making*. In: 2004 IEEE International Conference on Systems, Man and Cybernetics.
[Lourakis 10] Lourakis, M.I.A.: levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++. In: http://www.ics.forth.gr/~lourakis/levmar/. Accessed on 1 Jan. 2010.
[Meier-Hirmer 07] Meier-Hirmer, C.: *Modèles et techniques probabilistes pour l´optimisation des stratégies de maintenance. Application au domaine ferroviaire*. PhD, Université de Marne-la-Vallée (2007).
[Miwa et al. 00] Miwa, M., Ishikawa, T., und Oyama, T.: *Modeling the Transition Process of Railway Track Irregularity and Its Application to the Optimal Decision-making for Multiple Tie Tamper Operations*. In: 3rd International Conference Proceedings "Railway Engineering 2000". CD-ROM (2000), London.
[Oyama & Miwa 06] Oyama, T., Miwa, M.: *Mathematical Modeling Analyses for Obtaining an Optimal Railway Track Maintenance Schedule*. In: Japan Journal of Industrial and Applied Mathematics, 23-2 (2006), p. 207-224.
[Ubalde et al. 05] Ubalde, L., López Pita, A., Teixeira, P., Bachiller, A., and Gallego, I.: *Track deterioration in high-speed railways: influence of stochastic parameters*. In: *Proceedins of the Railway Engineering 2005, 8th Internationl conference and exhibition*, London. University of Westminster. (2005)
[Lunze 02] Lunze, J.: *What is a hybrid system?*. In: Engell, S., Frehse, G., Schnieder, E. (Eds.): Modelling, analysis, and Design of Hybrid Systems, Springer (2002), Berlin.
[Sadeghi & Askarinejad 08] Sadeghi, J., Askarinejad, H.: *Development of improved railway track degradation models*. In: Structure and Infrastructure Engineering, Taylor & Francis (2008), London.

SMTDA 2010: Stochastic Modeling Techniques and Data Analysis
International Conference, *Chania, Crete, Greece, 8 - 11 June 2010*

Data Mining Application for Eutrophication Control in Surface Waters

ANPALAKI J. RAGAVAN, Departments of Mathematics and Statistics, and Civil and Environmental Engineering, University of Nevada, Reno, NV 89557 (ragavan@unr.edu)

ABSTRACT

High total phosphorus concentrations(TP) has been found to be the major cause for Eutrophication and the subsequent depletion of dissolved oxygen(DO), enhancement of dissolved organic carbon(DOC), and poor water quality in Truckee River in Nevada. Identifying the exact pattern of relationship among multiple independent variables related to high TP levels is important to implement remediation methods. In this study non-linear mixed (NLMIXED) modeling, time series cross sectional regression, and non-linear least squares (NLLSQ) minimization were used to model the relationship of multiple independent variables to TP as closely as possible. Independent variables included alkalinity, total soluble phosphorus concentrations(STP), stream flow(SF), water pH, water temperature (Temp), DOC, and DO, sampled monthly at the same time of TP (from January 1997 to December 2007) over six monitoring sites (McCarran Bridge(MB), Wordsworth Bridge(WB), Steamboat Creek(SC), Derby Dam(DD), Lockwood(LW), and North Truckee Drain(NTD)) along Truckee River in Nevada. Seasonal variations and man-made intervention in TP were included in the analysis. Fitted NLMIXED model closely predicted observed data explaining 96.7% of total variation (R^2 =0.908). All independent variables influenced TP significantly at 1% significance level. All six sites contributed significantly towards overall TP at 5% significance level ($p<0.0001$). NLLSQ minimization solution (0.0694 mg/L) to TP was much above the observed overall minimum (0.001 mg/L). Although much lower mean TP were observed at sites MB, DD, and LW, compared to SC and NTD, smaller variations in TP made them significant contributors towards overall TP. Non-linearity in the relationship of TP to independent variables significantly influenced the prediction of TP hence should be included in all analyses related to prediction of TP in Truckee River. TP in Truckee River is also subject to significant seasonal fluctuations and man-made interventions. Non-linear programming is a suitable and accurate method to identify possible ranges of TP in Truckee River.

INTRODUCTION

Water quality management involves issues related to municipal, industrial and amenity irrigation practices. Due to increasing population and urbanization in Nevada in the past few years, increased concentrations of total phosphorus (TP) in the Truckee River have been recorded. The increased river diversions have increased agricultural practices which have lead to heavy growths of aquatic weeds and benthic algae, caused by high nutrient loads and low flows. Increased fertilizer use and sewage have modified the natural cycle of phosphorus, the relationships of which to soil use and agricultural, domestic and industrial activities are expected to rise in the future. Subsequently dissolved oxygen (DO) levels in the river have decreased due to plant respiration and decaying biomass. Low DO levels have impaired the river's ability to support populations of Lahontan cutthroat trout, a threatened species, and cui-ui (kwee-wee), a

national endangered species. There is also spatial variability in different catchments in phosphorus loading into Truckee River which imposes tremendous uncertainty in pollution load estimation. Water management practices must be improved in Nevada, to guarantee improved quality of water of sustainable water bodies affected by development of urban and suburban areas. Determination of factors affecting or causing variation of total phosphorus concentrations can provide a robust solution to quantify total phosphorus pollution in urban areas in Nevada.

Total phosphorus concentration (TP) in Truckee River varies both temporally and spatially and has been reported to be a function of several factors such as STP, SF, seasonality (Summer), man made intervention (X1), DOC, DO, Alkalinity (ALK), water pH (pH), and water temperature (Temp). All the above variables are collected at discrete time intervals including TP and therefore form time series. The degree and spatial variation of influence of the different factors on TP need to be predicted to reflect different sources of the phosphorus loading into the river. In addition the relationship of many of the factors to TP in the river has been found to be non-linear (Ragavan, 2008). Non-linear time series programming and modeling is an appropriate approach to analyze such data. Statistical models in which both fixed and random effects enter nonlinearly are becoming increasingly popular. Perhaps the greatest theoretical progress in time series analysis in the last ten years has been in the understanding of testing and modeling for nonlinearity. Nonlinear time series analysis raises the possibility of improving the power of parameter estimation and forecasting techniques. For any time series Y_t that is normal $\rho_k(Y_t^2) = \{\rho_k(Y_t)\}^2$ (where $\rho_k(.)$ denotes the lag k autocorrelation). Any departure from this result indicates a degree of non-linearity.

The major focus of this data mining study is to apply non-linear programming to identify and model the relationship of multiple independent variables to TP in Truckee River, as close as possible, which will enable designers to target and manage TP concentration in the Truckee River accurately as possible to their source of origination. Least squares minimization solution to the fitted non-linear model that leads to the detection of minimum TP levels in the Truckee River was also found to help environmental policy makers and designers to help in developing criteria for phosphorus loading into the river. The relationship between each independent variable included in the model to TP was predicted as accurately as possible before fitting the model. The distribution of TP at the different sampling sites was modeled as a function of the multiple independent variables using the previously identified relationships between dependent and independent variables. Monthly water quality data were obtained from Truckee Meadows Water Reclamation Facility (TMWRF, www.tmwrf.com) for the period from January of 1995 through December of 2007. The data from January of 1997 through December of 2004 were used to fit the non-linear mixed model. The data from January of 2005 through December of 2007 were used to validate the fitted model through forecasting. The data from January 1995 through December 1996 were excluded from model fitting due to missing values. The fitted non-linear mixed model was used as the objective function (TP is the objective variable) to find solution to NLLSQ minimization with respect to TP. The developed model can provide guidance to probable range and type of TP load generated and deposited into the Truckee River.

Study site

The Truckee River can be best described as a river in northern Nevada and northern California, that is140 mi (225 km) long, originates from the mountains, south of Lake Tahoe, flows into the Lake Tahoe at its south end, drains part of the high Sierra Nevada, and empties into Pyramid Lake in the Great Basin (USEPA, 1991). It flows generally northwest through the

mountains to Truckee, California, and then turns sharply to the east and flows into Nevada, past Reno and Sparks and along the northern end of the Carson Range. The river passes through the Reno-Sparks metropolitan area, located in Nevada's Truckee Meadows. A picture of Truckee River taken from down town Reno, Nevada is shown in Figure 1. East of the Truckee Meadows, fourteen ditches remove water for irrigation. The most significant diversion is Derby Dam, where at least 32% of the river's water is diverted annually (Peternel and Laurel, 2005). TMWRF currently maintains 11 continuous monitoring stations within the Truckee water system. These stations are located at: Mogul, SC, MB, NTD, LW, Patrick, Waltham, Tracy, Painted Rock, Wadsworth and Marble Bluff Dam. The Lockwood (LW) monitoring site is currently chosen as the compliance site for assessing total maximum daily loads (TMDL) for phosphorus into the Truckee River because most controllable sources are thought to be upstream. LW monitoring site is located in the lower Truckee River basin 65.6 river miles from Lake Tahoe located down stream, of MB, NTD, and SC monitoring sites and Vista (www.tmwrf.com) (Figure 2; source: http://truckeeriverinfo.org/gallery).

The problem

Truckee River's waters are an important source of drinking and irrigation along its valley and adjacent valleys. As discussed previously, increased urbanization and the prevalence of water diversions have caused a decline in water quality, and the resulting detrimental effects on habitat have brought about the need to restore the river to a more natural condition to improve habitat and the river's overall health. The water is quite clear near Lake Tahoe, but as it descends, the water turns muddy and concentrated in nutrients by the time it passes Reno, Nevada. The California State Water Resources Control Board (State board) has classified the middle reach of the Truckee River as "impaired", under Section 330(d) of the Clean Water Act.

Total phosphorus concentration in Truckee River is affected by spatial as well as temporal variations. Currently there is a need for a consistent, scientifically defensible approach for assigning nutrient criteria for Truckee River water, to control Eutrophication. Until now, exceedances of TP in Truckee River, has been found to be the major cause of Eutrophication (EPA, 2007). Recently, researchers are reporting other variables such as DO, SF, Temp and water pH to affect biomass activity and growth in the river. However, the relationship of these variables to TP in the river has not been studied and not fully understood. All the above factors that affect biomass in the river also directly or indirectly influence TP. This inadequacy of information currently limits the ability of NDEP to revise these values (required for determining, use status for the 303(d), Impaired Waters List (Category 5 of the Integrated Report)) to impose criteria for nutrients (NDEP, 2007). Subsequently implementing the beneficial use criteria is challenged because the beneficial use criteria focuses, on phosphorus (not nitrogen) for Eutrophication control. Phosphorus criteria are in the regulations.

It is beneficial to know the relationship of the multiple independent factors that affect biomass in the river to TP to implement nutrient criteria for phosphorus. The TMDL compliance level for total phosphorus concentration for Truckee River is currently at 0.075 mg/L (214 lb/day) set at the Lockwood (LW) monitoring site. The Eutrophication problem still persists. May be this level of compliance needs be revised in terms of value and location. Currently SC is the major contributor towards overall TP in Truckee River.

The total phosphorus concentration (TP), has been classified by the Environmental Protection Agency (NDEP, 1994) as a -conservative pollutant (conservative pollutants persist in the water segment of the aquatic environment over time remaining essentially constant in

concentrations), hence not expected to be perturbed by seasonal variations or other short term cyclical and non-cyclical variations in the system but varies directly with the volumes of flows of dischargers of the receiving water body. However, it is possible that TP concentration in Truckee River can be affected by seasonal Agricultural practices such as fertilizer use and other cyclical and non-cyclical man made activities. This classification needs revising too in the future. These variations are site specific and a suitable method to evaluate the overall TP level in the River, based on the data collected at the diverse monitoring sites is currently lacking. Optimization with respect to TP can be beneficial here.

Given the criteria of using phosphorus concentrations to predict Eutrophication impairments, it is not known in many cases, if the phosphorus levels are actually impacting the beneficial uses, e.g. aquatic life, recreation, etc. Some algae are a necessary component of the ecosystem, while excessive Algae can impact the beneficial uses in a variety of ways. According to EPA (2000): "Algae are either the direct or indirect cause of most problems related to excessive nutrient enrichment, e.g. algae are directly responsible for excessive, unsightly periphyton mats or surface plankton scums, and may cause high turbidity, and algae are indirectly responsible for diurnal changes in DO and pH". The exact relationship among DO, pH, DOC and TP need be identified.

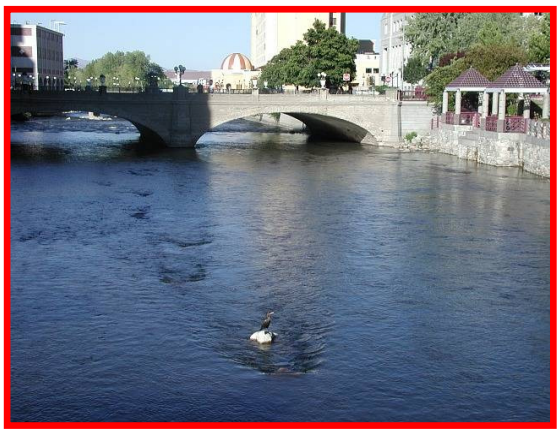


Figure 1: Truckee River in Reno, Nevada

Figure 2: TMWRF Monitoring Stations

Time series cross sectional regression

Time series cross sectional (TSCS) data analysis have become increasingly popular over the last ten years especially in environmental sciences. With TSCS data both cross-sectional and cross-temporal variation can be tested. But this theoretical power comes at the cost of methodological complexity. TSCS data structures invariably violate the standard Gaussian error assumptions underlying classical linear regression models. Hence much effort is required to develop and test models that can handle this pitfall. Early attempts in TSCS data analysis involve running feasible generalized least squares, estimating the variance covariance matrix of the errors using the ordinary least squares (OLS) method (Parks, 1967). A more sophisticated method known as the "panel corrected standard errors" (PCSE) has been the most common techniques for performing TSCS data analysis in the past few years (Parks, 1967). Due to a variety of problems with PCSE (they cannot account for unit effects, costly in terms of degrees of freedom, misspecification of model), the modern approach to modeling TSCS data is Time

Series Cross Sectional Regression (TSCSREG). The TSCSREG approach was used in this study.

Non-linear programming

TP in Truckee River is affected by spatial as well as temporal variations. Non Linear Programming (NLP) is an efficient way to identify the range of possible TP values in the river along several monitoring sites over a long period of time. NLP involves optimizing (minimizing or maximizing) a continuous non-linear objective function $f(x)$ with n independent (decision) variables, $x = (x_1,...,x_n)^T$ subject to constraints. Constraints include be: i) linear and nonlinear, ii) equality and inequality, and iii) lower and upper bound. For example the optimization (minimization) of the objective function $f(x)$ can be expressed as solving: $\min_{x \in R^n} f(x)$ subject to the following constraints:

$$c_i(x) = 0 \qquad i = 1,...,m_e$$
$$c_i(x) \geq 0 \qquad i = m_e+1,...,m$$
$$u_i \geq x_i \geq l_i \qquad i = 1,...,n$$

where c_i's are the constraint functions, and u_i, l_i's are the upper and lower bounds. The above setting can be applied to real world problems to find optimal control values and/or maximum likelihood estimates as solutions to an identified objective equation or relationship. The NLP programming can handle the following problems: i) Non-linear LSQ minimization, ii) quadratic programming, iii) constrained optimization (minimization/maximization), iv) unconstrained optimization (minimization/maximization), and, v) linear complementarities. Optimizing (minimizing or maximizing) the objective function $f(x)$ with the quadratic (non-linear) programming can be described as: $\min(\max) f(x) = \frac{1}{2} x^T G x + g^T x + b$ subject to constraints: $c_i(x) = 0 \quad i = 1,...,m_e$; where $c_i(x)$'s are linear functions. In the above example: $g = (g_1,...g_n)^T$ is a vector and b is a scalar of parameters, and G is a (n x n) symmetric matrix.

The non-linear LSQ programming with the objective function $f(x)$ to be minimized can be described as: $\min f(x) = \frac{1}{2}\{f_1^2(x) + ... + f_n^2(x)\}$, {where $f_1(x),...f_n(x)$ are nonlinear functions of x}, subject to the following constraints: $c_i(x) = 0 \quad i = 1,...,m_e$; where the $c_i(x)$'s are linear functions.

In this study the Non-Linear LSQ (NLLSQ) minimization technique was used to find solutions to a previously identified objective function, among the dependent and independent variables, minimizing the objective variable (TP), subject to non-linear and linear constraints and boundary values. Solution to NLLSQ minimization was obtained using SAS® software using the Quasi-Newton optimization algorithm (it is the only optimization algorithm which supports the use of non-linear constraints), and the Lagrange Multiplier Method (LMM) of solution. Objective function and the constraints were specified algebraically using SAS® programming statements.

Method of least squares

The method of least squares (LSQ) is a method of fitting data in regression analysis. In the LSQ method best fit of a model is obtained with the least value of the sum of squared residuals (SSR). Parameters of the model function are adjusted to best fit a data set. For example, for a simple data set consisting of N data points (data pairs) $(x_i, y_i, i = 1,..., N)$, where x_i is an independent variable and y_i is a dependent variable (whose value is found by observation), the model function has the form: $f(x_i, \beta)$, (where X is a matrix of independent variables and M adjustable parameters are held in the vector β). The LSQ method is the "best" fit when the SSR $= \sum_{i=1}^{N} r_i^2$ is the minimum. The residual (r_i), is the difference between the values of the dependent (y_i) variable and the predicted values from the estimated model as: $r_i = y_i - f(x_i \hat{\beta})$.

Non-linear least squares

One major problem with the application of LSQ methods is that there is no closed form solution with LSQ method to non-linear systems. Instead, numerical algorithms are used to find the value of the parameters (β), that minimize the objective function. Most algorithms involve choosing initial values for the parameters. The parameters are refined iteratively, that is, the values are obtained by successive approximation as:

$$\beta_j^{b+1} = \beta_j^b + \Delta\beta_j \qquad (1)$$

In Eq. [1], b is an iteration number and $\Delta\beta_j$ is the vector of increments known as the shift vector. Many solution algorithms for non-linear LSQ problems require that the Jacobian be calculated. The analytical expressions for the partial derivatives are complicated and impossible to obtain, hence the partial derivatives must be calculated by numerical approximation.

OBJECTIVES

1. To develop a non-linear mixed model to closely predict total phosphorus concentration as a function of multiple independent variables from several monitoring sites in Trucker River in Nevada.
2. To find solution and parameter estimates for the developed non-linear mixed model through non-linear least squares minimization.

LITERATURE REVIEW

Because of the endangered species present and due to the fact that Lake Tahoe Basin comprises the headwaters of the Truckee River, the river has been the focus of several water quality investigations, the most detailed starting in the mid-1980s. Under the direction of the U.S. Environmental Protection Agency, comprehensive dynamic studies have been undertaken to study the impacts of a variety of land use and wastewater management decisions throughout the 3120 square mile Truckee River Basin and also to provide guidance to other U.S. river basins (USEPA, 1991). Analytes mostly addressed include nitrogen, phosphorus, dissolved oxygen, and total dissolved solids. Impacts upon, the receiving waters of Pyramid Lake has also been analyzed (Truckee River Geographic Response Plan, 2005).

Nitrogen and phosphorus are the main nutrients that cause excessive algal growth in Truckee River (EPA, 2000). Elevated phosphorus loads have encouraged the proliferation of aquatic plants and benthic algae. Respiration by these plants and the decay of their associated detritus decreases dissolved oxygen (DO) in the water column, resulting in violations of the DO standard. Violations of the in-stream DO standard have continued in spite of recent nutrient removal enhancements by the TMWRF (NDEP, 1993).

However, the use of nutrient concentrations alone are poor predictors of assessing Eutrophication impacts (Tetra Tech (2005). Dodds and others, (2002) found by examining data from over 600 streams that nutrients concentrations to account for less than half of the variance in the benthic algae biomass. They speculated that other factors, such as flow, light availability, channel conditions, and grazing, were responsible for the remaining variability. In a detailed study of Colorado streams, Lewis, Jr. and McCutchan (2005) found even less of a relationship between nutrient concentrations and benthic biomass, with dissolved inorganic nitrogen accounting for only 15% of the variance. No statistically significant relationship was found between benthic biomass and other nitrogen and phosphorus species.

According to Biggs (2000) 62 percent of the variance in peak biomass was explained by the time since the last flood event. Increased water temperature can increase biological activity, including algae growth (Tetra Tech, 2002). However, Cladophora algae, has been found to die-off at temperatures exceeding 23.5 °C (Dodds and Gudder, 1992). These die-off events can lead to low dissolved oxygen levels as the algae decay. On the other hand, lower temperatures can lead to lower algal biomass. Lewis, Jr. and McCutchan, Jr. (2005) identified an inverse relationship between periphyton biomass and elevation, therefore a positive relationship between biomass and temperature.

METHODS

Data integrity testing

Autocorrelation

The assumption of residual independence when fitting regression models for time series data (data collected at discrete time intervals) requires, that the time ordered error terms display no autocorrelation. Whenever the errors corresponding to observations across time periods are not independent an autocorrelated error structure occurs. Such serially correlated errors (known as autocorrelation) speak about the linear dependence between observations (Box and Jenkins, 1976). Time series data must be corrected for autocorrelated errors before fitting any regression model with it (Parks, 1967). Autocorrelation function (ACF) plots can reveal the presence of autocorrelated errors in a series. ACF plots list the estimated autocorrelation coefficients of the series at each time lag. If the autocorrelation coefficient is statistically significant at a certain time lag it indicates the presence of significant autocorrelation at that time lag. Presence of autocorrelation in a series can also be tested using standard tests procedures such as Chi-Square tests at specific time lags with a previously specified significance level. In this study Chi-Square tests with 5% level of significance were performed at selected time lags (6, 12, 18, & 24) to test the data for the presence of autocorrelation in the series. An AR(1) covariance structure was used to correct the data for autocorreled errors. The AR(1) process predicts the future as an immediate past.

Outliers and influential observations

On the balance of probabilities, an observation beyond, 2.0 standard deviations (SD) from the mean need to be highlighted for follow-up investigation to identify causes such as intervention, and recording errors (Salas and Obeysekera, 1988). Only the most extreme observations (4.0 or more SD from the mean) need to be excluded from the analysis. In this study all time series variables were tested and corrected for the presence of outliers and any influential observations. No observations with values above 4.0 SD from the mean were observed in any of the variables in the data. All observations were included in the analysis. The observations above 2.0 SD from the mean of TP (0.305 mg/L) were recorded as influential and the intervention in response (TP) due to these influential observations were computed and included in the analysis as a dummy variable (X1).

Data non-stationarity

All time series regression models require that the time series modeled is stationary. It reveals whether the variations in a time series are likely to be permanent or temporary. A time series (X_t, t (time)=0, ±1, ±2,….) is said to be stationary if it has statistical properties similar to those of the time-shifted series (X_{t+h}, t=0, ±1, ±2,….) for each integer h. Strict stationarity of a time series {X_t, t=0, ±1, ±2,….) implies that the series {X_1, ….X_n} and the time shifted series {X_{1+h},…X_{n+h}} having the same joint distributions for all integers h and n > 0. Usually second order stationarity is adequate for modeling water quality time series (Fuller and Tsokas, 1971). For a second order stationary process the mean is constant and the auto-covariance function depends only on the time lag, which is consistent with a normal process. In this study the original time series variables were found to be non-stationary at 5% level of significance. The formal Dicky and Fuller unit root non-stationarity test (Fuller, 1978) was used to test the data for non-stationarity (Appendix: SAS® Code 1). First differencing of the time series was found adequate to correct the data for non-stationarity. Data was made consistent with a normal process through first differencing (constant mean and the auto-covariance dependent only on time lags).

Unobserved variation

Unobserved variations in a time series such as seasonal variation, trend, and other long and/or short term cyclical and/or non-cyclical variations due to any man-made intervention can influence regression analysis. In this study the unobserved variance components (seasonality, cycles, trend) in the original TP time series during the period from January 1995 through December 2004 were decomposed and their significance were tested at 5% level of significance using standard Chi-square tests (Appendix: SAS® Code 2). Since variations in TP due to all three unobserved components were significant ($p<0.05$), seasonal variations in TP (Summer, Winter), and possible intervention due to influential observations (X1) were computed as shown below and included in the analysis as explanatory variables. A value equal to 1 was recorded for the variable if the variable satisfied the following definition and recorded as equal to zero otherwise.

```
 X1      = 'intervention'
Summer  =  'summer months'
Winter   = 'winter months'
X1       = TP > 0.305
Summer  =  ( 5 < mm < 11 ) * ( year > 1990 );
```

Winter = (year > 1990) - Summer;

Missing values

Time series regression modeling requires data be without missing values. Missing values for any observation in any of the decisions variables can lead to missing values in the objective function. In this study after observing missing values in several of the variables missing values in the data were imputed through Marcov Chain Monte Carlo (MCMC) simulation using multiple chains (Schafer, 1997; Schafer, 1999) before fitting the non-linear mixed model (Appendix: SAS® Code 3). MCMC simulation involves sampling from probability distributions based on constructing a Markov chain that has the desired distribution as its equilibrium distribution. After multiple steps the state of the Marcov chain becomes the sample from the desired distribution. Since initial values for decision variables were used no problem with missing values were encountered during least squares minimization of TP.

Correlations and multicollinearity

All regression models including time series regression models require that the independent variables are not correlated. When a regressor is nearly a linear combination of other regressors in the model, the affected estimates are unstable and will have high standard errors. This problem is called variance inflation or multicollinearity. Data were tested for multicollenearity among independent variables. Variance inflation factors for the independent variables were obtained through fitting a linear multiple regression model for the dependent (TP) and the independent variables (Appendix: SAS® Code 4). A VIF value larger than 5 was taken to indicate the presence of multicollinearity in the variable. Simple Pearson correlation coefficients among the independent variables were also obtained along with significant probabilities at 5% level of significance.

Displaying the observed data

Histograms were constructed of original and the first differenced time series with a normal curve superposed. A bar chart of means TP levels observed at each site with lines superposed with values of overall mean TP and the compliance level for TP was created. Simple mean plots showing annual variations in TP at each site were plotted along with lines showing the compliance level for TP and the mean of observed overall TP in Truckee River superposed. Box plots of the distribution of each observed dependent and independent variable by site after counting the missing values separately for each variable at each site were developed (Appendix: SAS® Code 5). Scatter plots were constructed to display the observed relationships between the dependent and each observed independent variable and for detecting influential and/or outlying observations.

Software used in analysis

SAS® software was used to perform all the analyses and to generate all the plots and tables presented in this paper except Figure 1 and Figure 2. Figure 1 is a photograph taken of Truckee River in Reno. Figure 2 was obtained from http://truckeeriverinfo.org/gallery.

Cross sectional regression

A time series cross sectional multiple regression model between the dependent variable (TP) and the independent variables were fitted to the data to identify the significance of contribution of the individual sites towards overall TP in the Truckee River. Estimates of parameters with

probabilities were obtained as output from the model (Appendix: SAS® Code 6). Diagnostic statistics were tested for model adequacy.

Fitting non-linear mixed model

The relationship between the dependent and the independent variables were best fit into a non-linear mixed (NLMIXED) model. The NLMIXED model is a multiple regression equation that incorporates non-linearity in the dependent variables and/or one or more regression parameters (Harville, 1988; Searle, 1988). As the name implies the mixed model includes and solves for both the fixed (overall mean) and the random (site specific) effects in the relationship (Searle and others, 1992). The exact linear or non-linear relationships between dependent variable (TP) and the independent variables and the regression parameters were identified during model development. An alpha level of 0.01 (1%) was used for parameter estimation and all hypothesis testing during model fitting. Significant probabilities ($Pr>|t|<0.01$) of parameter estimates (lower the better), and hypothesis testing were used as criteria for selecting the best relationship between the dependent and independent variables, and the regression parameters.

Appropriate model diagnostic statistics (AIC, AICC, BIC and Log likelihood ratio, Akaike (1974); Buse (1973)) were used as criteria for checking for the adequacy of the fitted NLMIXED model for the data (Appendix: SAS® Code 7). Normal distribution was assumed for both fixed and random effects in the model. Estimates of parameters obtained from a best fitted linear mixed model to the same data in a previous study (Ragavan, 2008) were used as initial values for estimation of parameters of independent variables in the NLMIXED model in this study. Missing values generated in the data due to first differencing were automatically removed from the analysis during model fitting. Best parameter estimates, probabilities and the diagnostic statistics as well as model residuals from the best fitted NLMIXED model were obtained as output.

Least Squares Minimization

The NLMIXED model thus fitted was used as the objective function which was minimized with respect to TP through NLLSQ minimization. Solution to the objective function was obtained subject to non-linear and boundary constraints through NLLSQ minimization technique (Appendix: SAS® Code 8). Quasi-Newton optimization algorithm with the method of Lagrange Multiplier was used to obtain the NLLSQ minimization solution to the objective function. Gradient and the Jacobian of non-linear constraints were computed through Finite Differences. Average values of decision (independent) variables were used as initial values. Appropriate boundary and non-linear constraints for all decision variables were specified as algebraic program statements. Estimates of parameters obtained from the fitted NLMIXED model were used as initial values for estimation of parameters during NLLSQ minimization. This way minimization solution to the objective function was obtained very fast within few seconds. Lagrange multipliers used in the minimization process, gradients and solutions to the objective and Lagrange functions, and, parameter estimates after NLLSQ minimization were obtained as output from the model. Solutions to the objective and Lagrange functions for individual sites were also obtained. These values were compared to the overall observed mean TP and the compliance level for TP in Truckee River. Parameter estimates of the objective function from NLLSQ minimization were considered as final in the model.

Forecasting and sensitivity analysis

The model thus built was used to obtain monthly forecasts for TP for the period from January 2005 through December 2007. Forecasted TP values were plotted along with observed TP values. Diagnostic statistics from the final selected model were analyzed for model adequacy. Residual from the model were tested for all assumptions of residuals.

RESULTS AND DISCUSSION

Data integrity

The original TP time series in Truckee River, Nevada was non-linear, and not normal (Figure 3) and had significant autocorrelation at the tested time lags (6, 12, 18, 24) at 5% significance level ($p<0.0001$) (Table 1). After applying first differencing, the TP time series became consistent with the normal distribution (Figure 4) (constant mean and auto covariance function dependent only on time lags). The first differenced TP series was also devoid of any autocorrelations.

There were significant positive Pearson correlations between DOC and STP (0.729, $p<0.0001$), DOC and alkalinity (0.593, $p<0.0001$), and STP and alkalinity (0.749, $p<0.0001$). There were significant negative correlation between DO and Temp (-0.788, $p<0.0001$). Pearson correlations among all other independent variable pairs were not significant at 5% level. No multicollenearity were detected among any of the independent variables. Variable STP showed the largest variance inflation factor (VIF) equal to 3.47 which is less than the VIF required (= 5) for multicollinearity to occur. All other variables had VIF values less than 3.0 indicating no multicollinearity among the independent variables in the data.

Simple statistics

The mean (standard deviation[SD]), minimum, maximum and the median of the observed TP series over all sites were 0.117 mg/L (0.096), and 0.001 mg/L, 0.512 mg/L and 0.127 mg/L. The mean TP over all sites was above the compliance level (Figure 5 [SD are shown within brackets]). Mean TP at site SC (0.252[0.093] mg/L) was the largest followed by at NTD (0.208[0.092] mg/L). The mean values were much above the compliance level (0.075mg/L) at these two sites. Site MB showed the lowest mean TP value (0.027[0.088] mg/L). Mean annual variations in TP showed almost the same trend at all sites (Figure 6). Annual mean TP, were at or below the compliance level during the period between 1997 and 2000, increased largely above compliance level between the period from 2000 until 2002, and decreased thereafter until 2004 (Figure 6). Thirty five out of the 576 original observations (5.6%) had TP values above 2-SD from the mean (> 0.305 mg/L).

Simple statistics of independent variables (mean, SD, minimum, and maximum) taken over all sites are shown in Table 2. Variable SF showed larger SD than the mean indicating the presence of influential observations. Mean SF was high (643.8, [SD= 900.9] cubic feet per second [cfs]). However, SF in Truckee River was less than the overall mean, 72% of the time during the study period. Mean STP was high too (0.1 [SD=0.1] mg/L). Mean DO was equal to 10 mg/L (SD =1.8mg/L) which is much above the compliance level for DO (5 mg/L) in Truckee River. DO in Truckee River exceeded the compliance level 99% of the time and exceeded the overall mean DO, 54% of the time during the study period.

Location and variance information by sites

The location and variance information of dependent and independent variables among sites are shown as Box plots (Figures 7a through 7f)). Mean SF values did not vary significantly among sites. All sites showed very small variable response to SF with 50 percent of the observations having values within 500 cfs. Site SC showed the least variable response to SF with 50 percent of the observations within 50 to 60 cfs (Figure 7a).

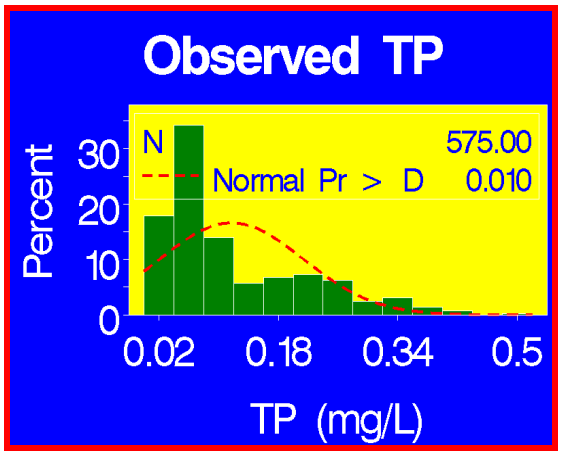


Figure 3: Histogram of the observed TP

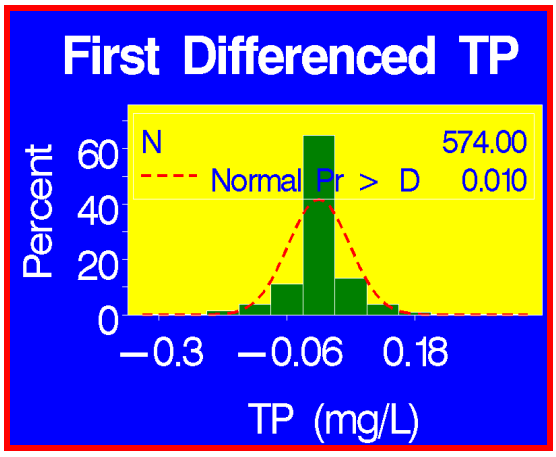


Figure 4: Histogram of the first differenced TP

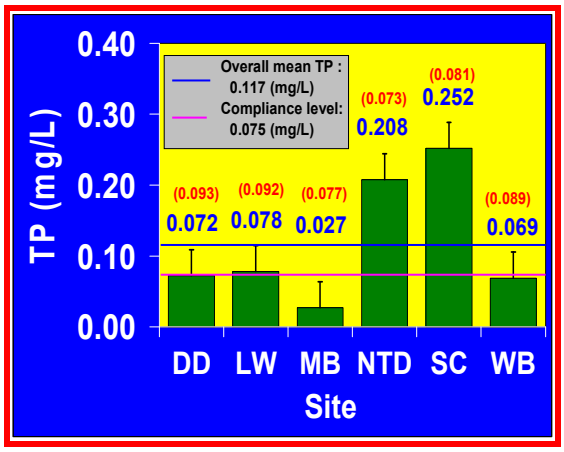


Figure 5: Observed mean TP by sites

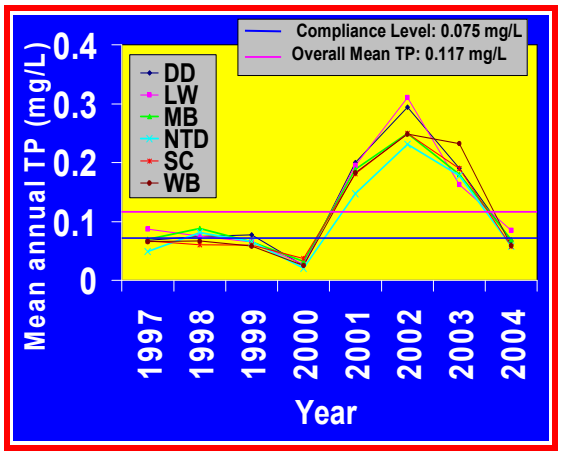


Figure 6: Annual variations in observed TP

There were no significant differences among sites in mean DO values. Mean DO values were above 7.5 mg/L at all sites with 50% of the observations showing DO values within 3 mg/L (Figure 7b). Site SC showed the lowest mean DO.

Table 1: Autocorrelation check for white noise for observed TP series

Time Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	1272.42	6	<.0001	-0.476	-0.43	0.818	-0.423	-0.419	0.871
12	2408.99	12	<.0001	-0.419	-0.41	0.782	-0.408	-0.39	0.815
18	3442.67	18	<.0001	-0.395	-0.38	0.735	-0.388	-0.371	0.785
24	4404.08	24	<.0001	-0.376	-0.368	0.702	-0.369	-0.357	0.757

There were significant differences in mean DOC and mean STP among sites. Sites NTD and SC showed significantly larger mean DOC and mean STP values than the other sites. Site NTD showed the largest variation in STP and DOC (Figures 7c and 7d respectively). The mean and the variability of STP (Figure 7d) was the same as that of TP (Figure 7g) at all sites.

Table 2: Simple statistics of independent variables over all sites

Variable	N	Mean	Std Dev	Minimum	Maximum
SF	560	643.8	900.9	5.0	5790.0
DO	574	10.0	1.8	4.5	14.4
DOC	575	3.1	2.1	0.1	22.4
STP	575	0.1	0.1	0.0	0.4
Alkalinity	575	101.6	67.6	29.0	374.0
pH	574	8.0	0.3	7.0	9.0
Temp	574	10.9	6.0	0.1	24.7

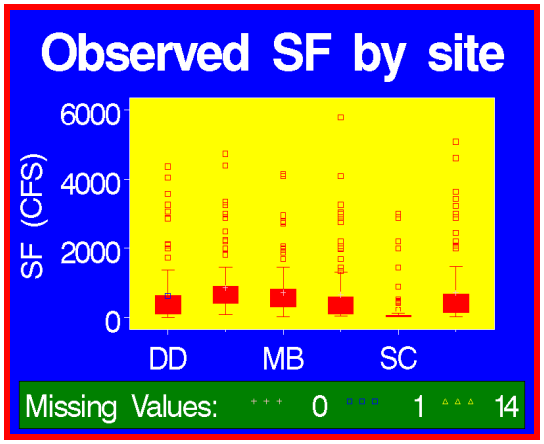


Figure 7a: Box plot of observed SF by site

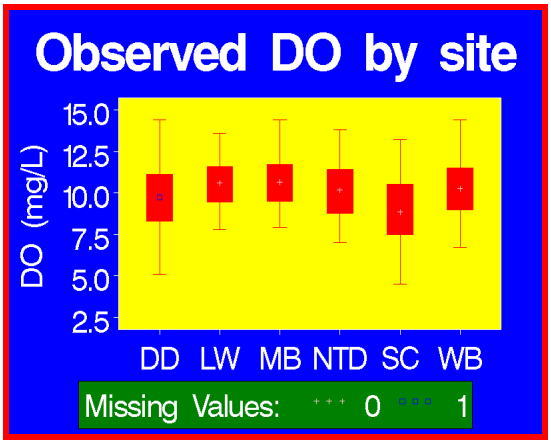


Figure 7b: Box plot of observed DO by site

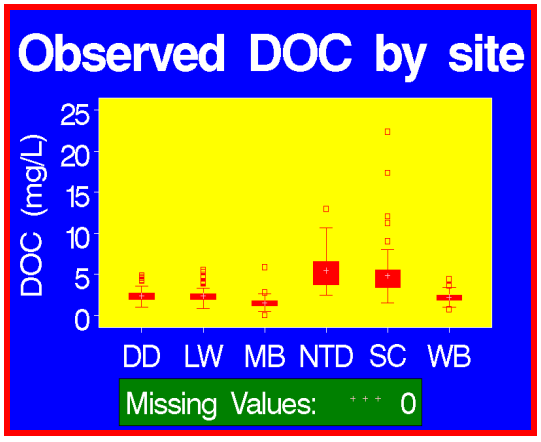


Figure 7c: Box plot of observed DOC by site

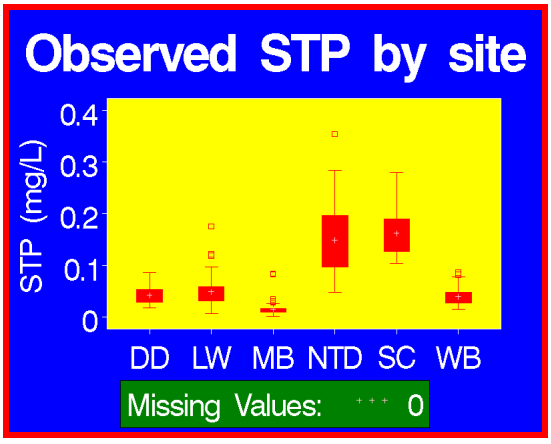


Figure 7d: Box plot of observed STP by site

Differences in mean pH or mean Temp among sites were not significant (Figures 7e and 7f). All sites showed large variability in water temperature (Temp) (Figure 7f). Site SC showed the smallest pH variability but the largest mean pH compared to other sites. The location and variance information of TP by site and by year is shown in Figure 7g, and, Figure 7h respectively. Site NTD showed the largest variation in TP followed by SC. Variations in TP were much smaller in others sites. Site MB showed the smallest variation in TP. Annual means were almost constant. Variations in mean TP increased after 2000.

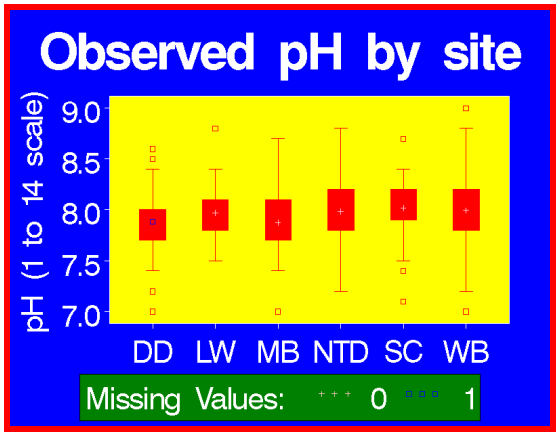


Figure 7e: Box plot of observed pH by site

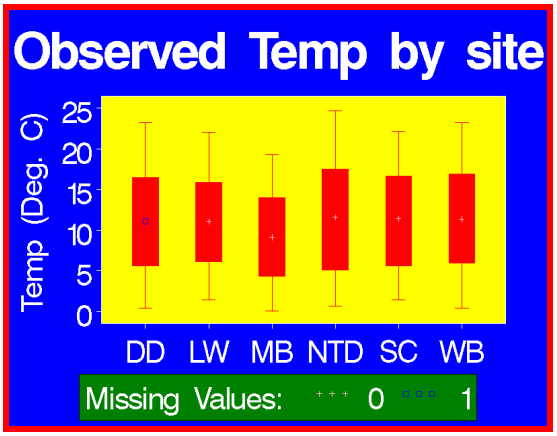


Figure 7f: Box plot of observed Temp by site

Cross sectional regression analysis

Contribution of individual sites to overall TP in Truckee River was significant at 5% level of significance for all sites ($p<0.0001$) (Table 3). All sites contributed positively towards overall TP in the Truckee River. Site SC contributed the largest (0.2517, $p<0.0001$) followed by site NTD (0.2017, $p<0.0001$). Site MB although contributed the smallest, was a significant contributor ($p<0.0001$) towards overall TP. The fitted cross sectional regression model was highly significant ($p<0.0001$) at 5% level of significance. The sites with smaller variations (DD, LW, MB, and WB) in TP could contribute significantly towards overall TP in the river.

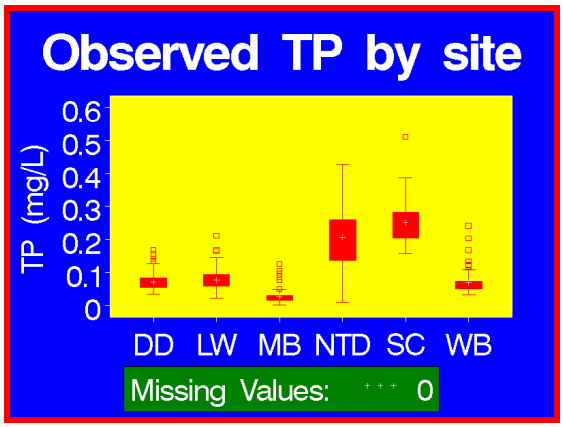


Figure 7g: Box plot of observed TP by site

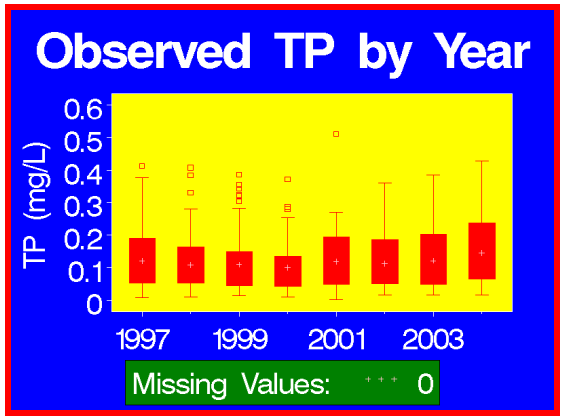


Figure 7h: Box plot of observed TP by year

Observed relationships among TP and independent variables

The observed relationship of DOC, STP and alkalinity to observed TP were linear and positive (Figures 8b, 8c and 8g respectively). The relationship of observed DO, Temp and SF to observed TP were non-linear and negative (Figures 8a, 8e and 8f respectively). The relationship of observed water pH to observed TP could not be predicted from the scatter plot although appeared to be positive (Figure 8d).

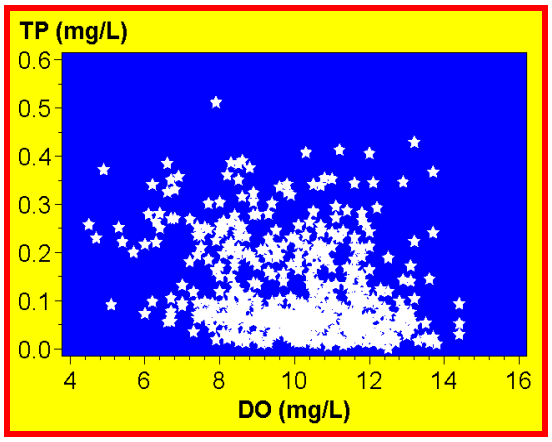


Figure 8a: Observed TP versus DO

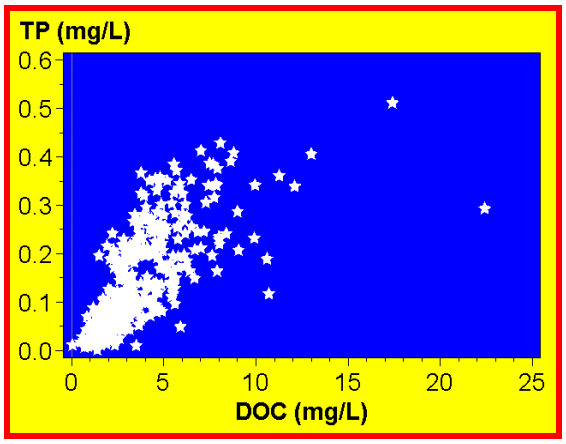


Figure 8b: Observed TP versus DOC

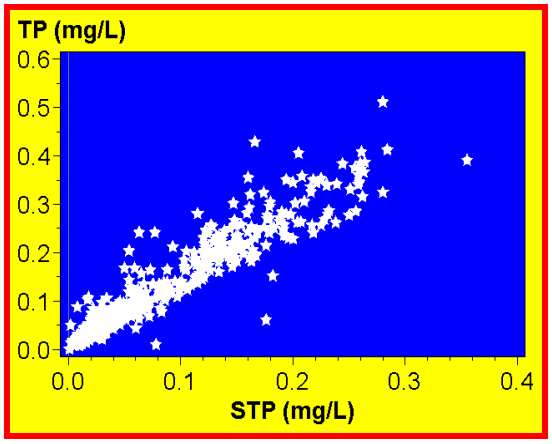


Figure 8c: Observed TP versus STP

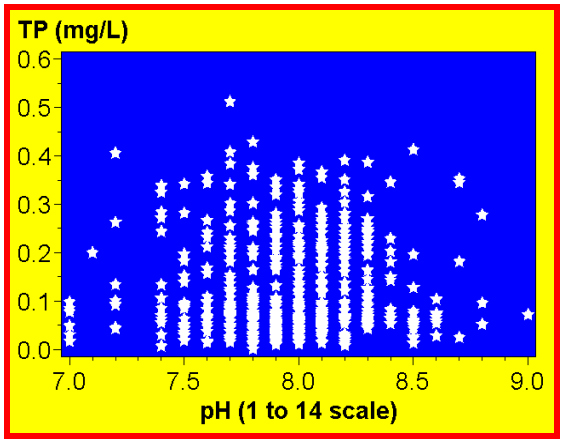
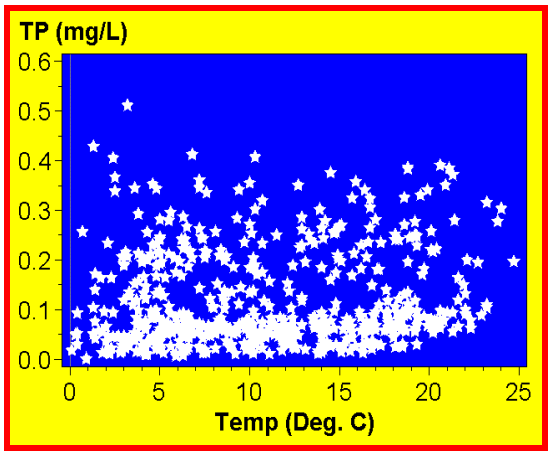


Figure 8d: Observed TP versus pH



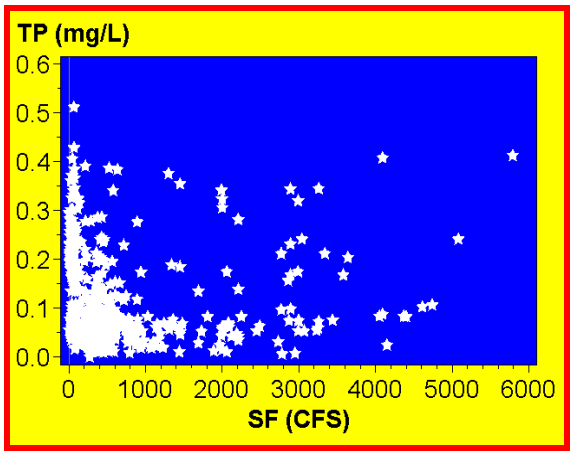


Figure 8e: Observed TP versus Temp

Figure 8f: Observed TP versus SF

Unobserved variations in original TP time series

The unobserved components (trend, seasonality, cycles) of observed TP were significant at 5% level of significance ($p<0.05$). The trend over time of TP in Truckee River was slowly increasing up until 1998 with values below the overall mean TP, increased suddenly between January 1999 and January 2000 with values above overall mean TP, and increased thereafter with values above the overall mean TP until the end of the study period (December of 2004) (Figure 9). Seasonal variations in TP were periodic with a period equal to 12 months (Figure 10). A significant ($p<0.05$) long term cyclical variation (period ~12 months) existed in the TP series which could be due to man-made intervention (Figure 11).

The above results indicate that TP in Truckee River is significantly subject to seasonal variations and man-made activities. These variations must be included when fitting any models to predict TP in Truckee River.

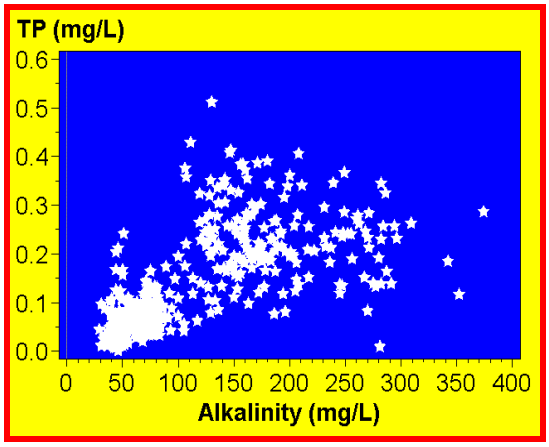


Figure 8g: Observed TP versus Alkalinity

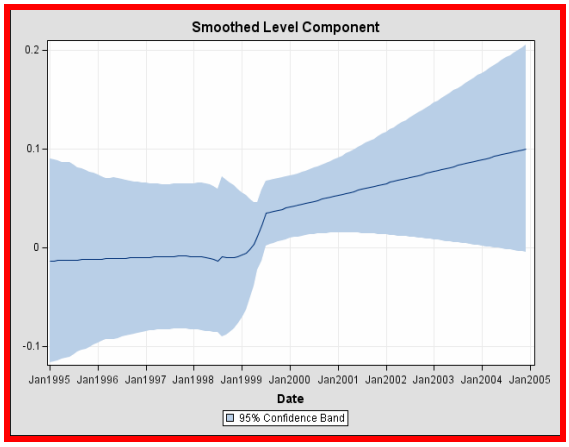


Figure 9: Trend component of observed TP

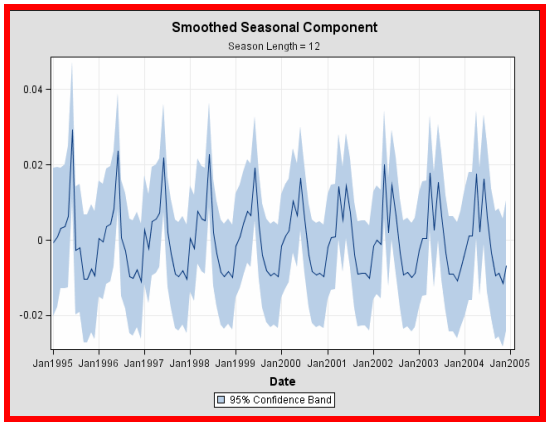


Figure 10: Seasonal component of observed TP

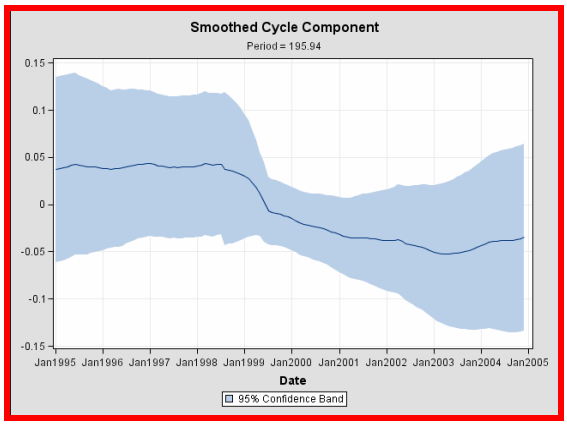


Figure 11: Cyclical component of observed TP

Non-linear mixed modeling

A negative exponential relationship was identified between observed DO and TP which was highly significant at 1% level of significance ($p<0.0001$). A negative logarithmic relationship was identified between observed SF and TP which was significant at 1% level of significance ($p=0.0046$). An inverse negative relationship was identified between observed Temp and TP, which was almost significant at 1% level of significance ($p=0.0155$) (Table 4).

The parameter estimates for DOC ($p=0.0148$) and the non linear temperature variable (1/{1-Temp}) ($p=0.0155$) were significant at the 5% level, and almost significant at 1% level, hence were not excluded from the final NLMIXED model. Parameter estimates for seasonal variations (Summer) and the man-made intervention (X1) were periodic (period=12 months), positive, and highly significant ($p<0.0001$) at 1% level of significance.

Table 3: Significance of contribution of individual sites to overall TP

Site	Estimate	Standard Error	DF	t Value	Pr > |t|
MB	0.02696	0.00514	569	5.25	<.0001
WB	0.06892	0.00514	569	13.41	<.0001
SC	0.2517	0.00517	569	48.72	<.0001
LW	0.07804	0.00514	569	15.19	<.0001
DD	0.07227	0.00514	569	14.06	<.0001
NTD	0.2075	0.00514	569	40.38	<.0001

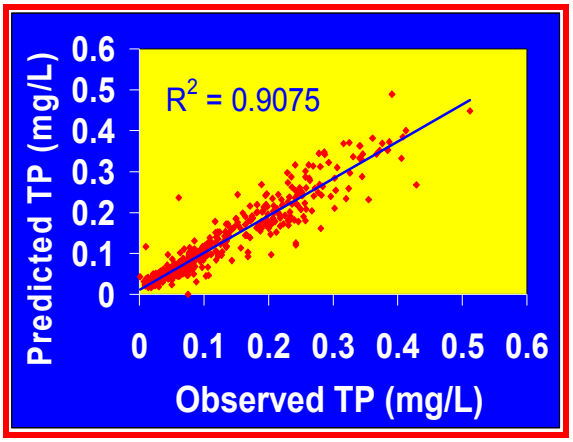


Figure 12: Overall observed TP against TP predicted by the NLMIXED model

Table 4: Parameter estimates and probabilities from the NLMIXED model

Variable	Parameter	Estimate	Standard Error	DF	Pr > |t|	Alpha
Intercept	beta1	-0.3973	0.000306	559	<.0001	0.01
DOC	beta2	0.005812	0.002377	559	0.0148	0.01
EXP(DO)	beta3	-7.495	0	559	<.0001	0.01
LOG(SF)	beta4	-0.00883	0.003100	559	0.0046	0.01
pH	beta5	0.03741	0.002664	559	<.0001	0.01
STP	beta6	1.2065	0.000043	559	<.0001	0.01
EXP(Summer-12)	beta7	0.10000	0	559	<.0001	0.01
EXP(X1-12)	beta8	0.03674	0	559	<.0001	0.01
1/(1-Temp)	beta9	-0.01616	0.006658	559	0.0155	0.01
Alkalinity	beta10	0.000390	0.000078	559	<.0001	0.01

Water pH, STP and alkalinity showed positive, linear, highly significant parameter estimates ($p<0.0001$) at 1% level of significance (Table 4).
All the independent variables studied were included in the objective function as decision variables. Significant non-linearity was identified in the parameter of DO ($p<0.0001$) and SF ($p<0.0001$). An exponential parameter for DO and a logarithmic parameter for SF fitted the model best. According to the above results variations in TP in Truckee River is influenced significantly by non-linearity in its relationship with the independent variables.

Predicted mean (SD value is shown within brackets), minimum, maximum, and, median TP over all sites from the NLMIXED model were 0.113 (0.089) mg/L, 0.017 mg/L, 0.489 mg/L, and 0.076 mg/L respectively. Predicted mean and the maximum TP values agreed well with the observed mean and maximum TP values respectively. Predicted minimum TP (slightly larger) and the predicted median TP (slightly smaller) were slightly deviant from observed minimum and median TP values. Predicted mean TP values at individual sites along with observed mean TP values are shown in Table 4 (SD are shown within brackets).

The best fitted NLMIXED model explained 96.7 % of the total variation and predicted observed TP very closely (R^2=0.908; n=576 before first differencing) (Figure 12). The mean TP values predicted by the NLMIXED model at individual sites also closely agreed with observed mean TP values (Table 5). However, the predicted TP values were slightly underestimated at SC, and were slightly overestimated at NTD and LW by the NLMIXED model (Table 5). Mean TP at sites SC and NTD were significantly larger than that at other sites. A box plot of the model predicted TP by individual sites is shown in Figure 13. Site NTD showed the largest variation in predicted TP, while MB showed the smallest mean predicted TP and the smallest variation in predicted TP. The mean and variation information for predicted TP by year are shown in Figure 14. Mean annual predicted TP is almost constant. Variations in predicted TP are slightly larger than that of observed TP.

Non linear least squares minimization

The objective function used in the NLLSQ minimization is shown in Equation 1. The boundary values and non-linear constraints used in the minimization process are listed following the objective function. Solution to NLLSQ minimization for the objective function was 0.0694 mg/L which is below the current compliance value for TP set at LW. This value is also below the mean observed TP and the mean predicted TP over all sites. Solutions for individual sites were below the observed mean TP values at the respective sites. Solution to the respective Lagrange function was negative (-0.0684). Gradients for the objective and the Lagrange functions were close to zero for all parameters. Parameter estimates after NLLSQ minimization were same as that obtained from the NLMIXED model for all variables. Hence the parameter estimates and from the NLMIXED model, were accepted as final values. The solutions to objective and Lagrange functions for individual sites are shown in Table 6. NLLSQ minimization solution to objective function at sites SC and NTD were above the overall compliance level for TP in Truckee River. These two sites require better management practices towards reducing overall TP loading into the Truckee River.

The Objective function

$$TP = \mathbf{beta1} + \mathbf{beta2} * DOC + (EXP(-\mathbf{beta3} * DO)) + ((LOG(\mathbf{beta4}^2) * LOG(SF))$$

+ **beta5** * pH + **beta6** * STP + **beta7** * (EXP(Summer-12)) + **beta8** * (EXP(X1-12)) - **beat9** * (1/(1-Temp)) + **beta10** * Alkalinity (2)

Boundary values:

-3.2E-10 ≤ beta3 ≤ **3.2E-10.**
0 < beta7 ≤ **1.**
0 < beta8 ≤ **1.**

Where beta3, beta7 and beta8 are coefficients for DO, seasonality variable (Summer), and, the intervention variable (X1).

Non-Linear Constraints:

-5 ≤ LOG(SF) ≤ **5**
6.14421E-06 ≤ EXP(Summer-12) ≤ **1.67017E-05.**
6.14421E-06 ≤ EXP(X1-12) ≤ **1.67017E-05.**
1 ≤ EXP(DO) ≤ **1.586E+15.**
1 ≤ pH ≤ **14.**
0 ≤ DOC ≤ **100.**
0 ≤ STP ≤ **5.**
0 ≤ Alkalinity ≤ **500.**
(1-Temp) ≠ **0**

Table 5: Observed and predicted mean (SD) TP at sites along with overall TP

Site	Mean Observed	Mean Predicted
DD	0.072 (0.093)	0.072 (0.028)
LW	0.078 (0.092)	0.081 (0.017)
MB	0.027 (0.077)	0.034 (0.006)
NTD	0.208 (0.073)	0.220 (0.046)
SC	0.252 (0.081)	0.223 (0.015)
WB	0.069 (0.089)	0.067 (0.009)
OVERALL	0.117 (0.096)	0.113 (0.089)

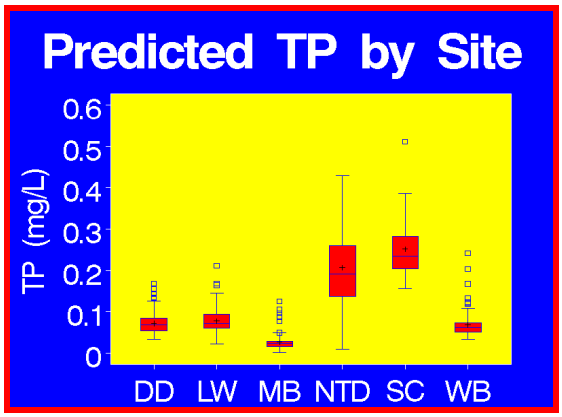


Figure 13: Box plot of predicted TP by site

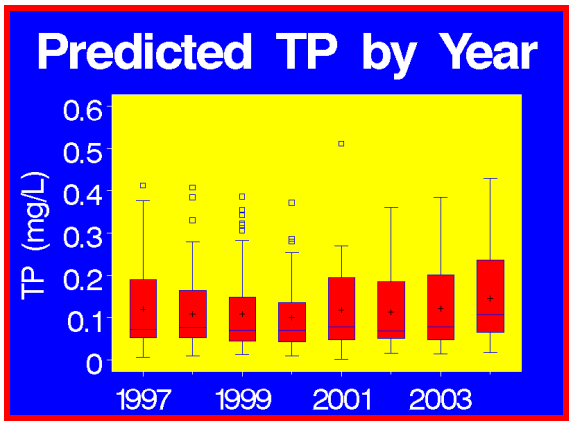


Figure 14: Box plot of predicted TP by year

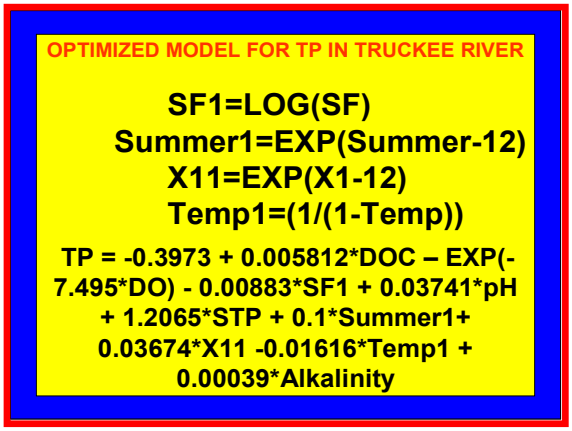


Figure 15: Objective function for TP

Model validity

Diagnostic statistics obtained from the best fitted NLMIXED model were sufficiently low (Table 7). The Residuals from the best NLMIXED model did not show any significant pattern (Figure 16). The residuals also did not show any particular relationship to the observed and/or to the predicted TP (Figure 17). The residuals were normal.

Forecasting using fitted model

Forecasted TP values from the final NLMIXED model closely agreed with observed TP for the forecast period (January 2005 to December 2007, n=36) (Figure 18). The observed (0.1386 [SD=0.045] mg/L) and forecasted (0.1373 [SD=0.027], mg/L) mean TP values during the forecast period agreed well. Forecasted TP decreased from January 2005 to May 2005, and increased from June 2005 until December 2005. Forecasted TP also decreased during the period between January 2006 and December of 2006, and increased thereafter until December 2007.

Table 6: Solution to objective and Lagrange functions at individual sites

Site	Objective Function	Lagrange Function
MB	0.0327	-0.0618
WB	0.0460	-0.0663
SC	0.1347	-0.0574
LW	0.0487	-0.0668
DD	0.0463	-0.0663
NTD	0.1288	-0.0591
Overall	0.0694	-0.0685

Table 7: Diagnostic statistics of the best fitted non-linear mixed model

Statistic	Value
-2 Log Likelihood	-1052
AIC (smaller is better)	-1026
AICC (smaller is better)	-1026
BIC (smaller is better)	-970.2

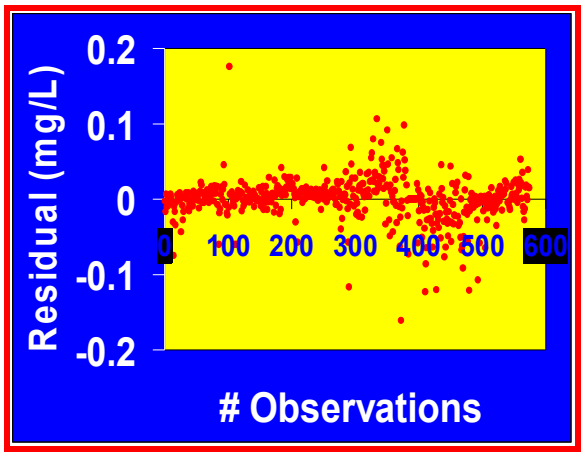


Figure 16: Pattern of model residuals

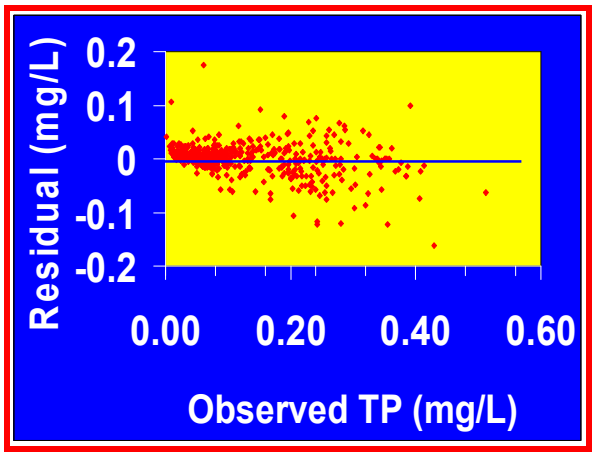


Figure 17: Residuals against observed TP

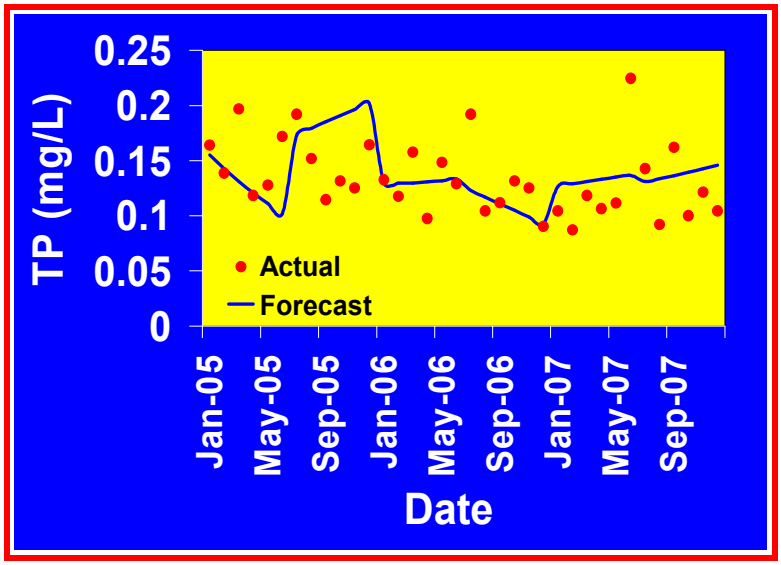


Figure 18: TP forecasted by the NLMIXED model along with observed TP (January 2005 through December 2007)

CONCLUSIONS

Trend of observed TP in Truckee River is significantly on the increase. Currently SC is the major contributor of TP to Truckee River followed by NTD. TP loading from these two sites require careful monitoring to reduce build up of TP and subsequent Eutrophication in Truckee River. The non-linear objective function built for TP in this study predicted TP closely (R^2=0.908) and forecasted the original TP in Truckee River closely accurately explaining 96.7 percent of the total variation in TP. Parameter estimates of all the independent variables studied were significant at 1% significance level ($p<0.01$) except the non-linear variation of water Temperature which was significant at 5% and was almost significant at 1% level ($p=0.0155$).

Results indicate that TP in Truckee River can be predicted accurately as a function of the seven independent variables studied. Since non-linearity in the relationship of the independent variables to TP can be a significant contributor toward prediction, nonlinear mixed modeling is

an appropriate way of modeling the relationship of independent variables to TP, which can explain both the random and the fixed effects in the relationships. .

As predicted by the model, the relationships of SF, DO and water temperature to TP are non-linear and negative. TP in Truckee River are also affected by significant seasonal variations, and significant cyclical man-made intervention events. These variations must be taken into account when fitting a model to predict and forecast TP values in Truckee River. Non-linear LSQ minimization provided just one solution, reproduced the fitted NLMIXED model exactly and gave reasonably accurate solutions for TP at all sites. Non-linear optimization is an efficient approach to predict possible and accurate range of TP values in Truckee River close to their target sites. The NLLSQ minimization solution of the objective function (0.069 mg/L) for TP is below the current compliance level for TP (0.075 mg/L) at LW which indicates a need for revising the existing criteria for phosphorus loading into the river for lowering Eutrophication.

REFERENCES

Akaike, H., 1974, A new look at the statistical model identification, IEEE trans.: Autom. Control AC-19, p. 716-723.

Box, G.E.P., and Jenkins, G.M., 1976, Time series analysis forecasting and control, (2nd ed.): Holden-Day, San Francisco, Ca.

Buse, A., 1973, Goodness of fit in generalized least squares estimation, American Statistician, v. 27, p. 106-108.

DaSilva, J.G.C., 1975, The analysis of cross-sectional time series data, Ph.D. dissertation, Department of Statistics, North Carolina State University.

Dodds, W.K., Smith, V.H., and Lohman, K., 2002, Nitrogen and phosphorus relationships to benthic algal biomass in temperature streams, Can. J. Fish. Aquat. Sci., v. 59, p. 865-874.

Dodds, W.K., and Gudder, D.A., 1992, The Ecology of Cladophora, Journal of Psychology, v. 28, n. 4, p. 415-427.

Fuller, W., 1978, Introduction to time series, New York: John Wiley & Sons, Inc.

Harville, D.A., 1988, Mixed-model methodology: Theoretical justifications and future directions, Proceedings of the Statistical Computing Section, American Statistical Association, New Orleans, p. 41-49.

Lewis, W.M. Jr., McCutchan, Jr., J.H., December 2005, Environmental thresholds for nutrients in streams and rivers of the Colorado Mountains and Foothills Report.

NDEP, 1994, Truckee River final total maximum daily loads and waste load allocations: Nevada Division of Environmental Protection, Carson City, Nevada.

NDEP, 2007, Nevada's nutrient assessment protocols for Wadeable streams: Nevada Division of Environmental Protection, Carson City, Nevada.

NDEP, 1993, Truckee river strategy: Nevada Division of Environmental Protection, Carson City, Nevada.

Parks, R.W., 1967, Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated, Journal of the American Statistical Association, v. 62, p. 500-509.

Peternel, K., and Laurel, S., May 15-May 19, 2005, Truckee River Restoration Modeling, World Water and Environmental Resources Congress. Anchorage, Alaska, USA.

Ragavan, A., 2008, Data Mining Application of Non-Linear Mixed Modeling in Water Quality

Analysis, Proceedings of the Data Mining and Predictive Modeling Section, SAS® Global Forum, San Antonio, TX, Paper 140-2008.

Salas, J.D., and Obeysekera, J.T.B., 1988, ARIMA models Identification of Hydrologic Time Series, Water Resources Research, v. 18, no. 4, p. 1011-1021.

Schafer, J.L., 1999, Multiple Imputation: A Primer, Statistical Methods in Medical Research, v. 8, p. 3-15.

Schafer, J.L., 1997, Analysis of Incomplete Multivariate Data, New York: Chapman and Hall.

Searle, S. R., 1988, Mixed Models and Unbalanced Data: Wherefrom, Whereat, and Whereto?, Communications in Statistics - Theory and Methods, v. 17, n. 4, p. 935-968.

Searle, S.R., Casella, G., and McCulloch, C.E., 1992, Variance Components, New York: John Wiley & Sons, Inc.

Tetra Tech Inc., 2005, Technical Approach to Develop Nutrient Numeric Endpoints for California, U.S. EPA Region, IX.

Truckee Meadows Water Reclamation Facility: www.tmwrf.com

Truckee River Geographic Response Plan, 2005: http://ndep.nv.gov/bca/emergency/truckee_river_plan05.pdf

USEPA, 1991, Guidance for water quality-based decisions: The TMDL process EPA 440/4-91-001, U.S.EPA, Office of Water, Washington, DC.

USEPA, July 2000, Nutrient Criteria Technical Guidance Manual: Rivers and Streams, U.S.EPA-822-B-00-002.

USEPA, March 2007, N-Steps: http://n-steps.tetratechffx.com/NTSCHome.com

APPENDIX

SAS® CODE 1

```
PROC ARIMA DATA=Monthly;
IDENTIFY VAR= TP
STATIONARITY=(ADF=(1,2,4,6,12));
RUN;

DATA Monthly; SET Monthly;
      TP =DIF(TP);
RUN;
```

SAS® CODE 2

```
PROC UCM DATA=Monthly PRINTALL;

   ID Date INTERVAL=Month;
   MODEL TP;
   IRREGULAR plot=smooth;
   LEVEL variance=0 noest
      plot=smooth;
   SLOPE variance=0 noest
      plot=smooth;
   CYCLE rho=1 noest=rho
      plot=smooth;
   SEASON length=12
      plot=smooth;
RUN;
```

SAS® CODE 3

```
PROC MI DATA=Monthly SEED=21355417
NOINT NIMPUTE=6 MU0=50 10 180
   OUT=outmi;
MCMC CHAIN=multiple DISPLAYINIT
   INITIAL=em(ITPRINT);
 VAR TP Alkalinity DO2 DOC STP
     SF pH Temp;

RUN;
```

SAS® CODE 4

```
DATA Monthly; SET Monthly;
sflog=Log(ABS(SF));doexp=EXP(DO2);
summer1=exp(summer-12); x11= exp
(summer-12); Temp1=(1/(1-Temp));
RUN;
PROC REG DATA=Monthly;
MODEL tp= sflog doexp DOC STP pH
Alkalinity Temp1 X11 Summer1 /VIF;
RUN;
PROC CORR DATA=comp;
VAR SF DO2 DOC STP Alkalinity pH
Temp;
RUN;
```

SAS® CODE 5

```
PROC MEANS DATA=Monthly NOPRINT;
      VAR TP;
      BY site;
OUTPUT OUT=Cancel NMISS=ncancel;

 DATA Comp;
      MERGE Monthly Cancel;
      by site;
RUN;
    symbol1 v=plus    c=black;
    symbol2 v=square   c= red;
    symbol3 v=triangle
      c=yellow;
TITLE 'Distribution of Original
TP Among Sites';

PROC BOXPLOT DATA=Comp;

PLOT TP *site = ncancel /
   boxstyle = schematicid
   cboxes=blue cboxfill = red
   cframe=vligb nohlabel
   symbollegend = legend1
   notches;
   legend1 label=('Missing
      Values:')
 cborder = black cframe=ligr;
 label TP ='TP (mg/L)';

RUN;
```

SAS® CODE 6

```
PROC TSCSREG Data=Monthly
      OUTEST=out1 COVOUT CORROUT;
ID site date;
MODEL TP=Alkalinity DOC DO2 SF
    pH STP Summer X1 Temp
      /NOINT RANTWO DaSilva;

TEST MB =0; TEST WB=0; TEST SC=0;
  TEST LW=0; TEST DD=0; TEST LW=0;

RUN;
```

SAS® CODE 7

```
PROC NLMIXED DATA=Monthly QPOINTS=10 ALPHA=0.01 TECH=QUANEW;
UPDATE=DDFP;

PARMS
   beta1=0.000246 beta2=0.006199 beta3=-0.000000548 beta4= -0.000001
   beta5=-0.0000002455 beta6=-0.7867 beta7=0.1 beta8=0.03665 beta9=-
   0.00066 beta10=-0.002 g11=-0.001428 to 0.02 by 0.001 g12=-0.001 to
   0.01 by 0.001;

 eta = beta1+ beta2*DOC+ exp(-beta3*DO))+ {(log(beta4^2))*(log(abs(SF)))
    + beta5*pH + beta6*STP + beta7*(EXP(Summer-12)) + beta8*(EXP(X1-12))
    + beta9*(1/(1-Temp)) + beta10* (Alkalinity) + g12*b1;
       num = eta; mu= num;
 MODEL TP ~ NORMAL(mu,g12);
 RANDOM b1 ~ NORMAL(0,g11) SUBJECT=SITE;
 PREDICT mu OUT=cdf;

RUN;
```

SAS® CODE 8

```
PROC NLP PALL TECH=quanew CLPARM=BOTH BEST=10 FD=Forward OUTMOD=model;
  LSQ TP;

   PARMS
      beta1= -0.3973, beta2= 0.005812, beta3 = -0.000000032, beta4=
      -0.000001, beta5= 0.03741, beta6= 1.2065, beta7= 0.1, beta8=
      0.03674, beta9=-0.01616, beta10= 0.00039, sf =374.0, Temp=10.94,
      Summer=1, x1=1, do2=10.04, pH=8.0, doc=3.15, stp=0.076,
      Alkalinity=101.62;

   BOUNDS
      -3.2E-10 <= beta3 <= 3.2E-10, 0 < beta7 <= 1, 0 < beta8 <= 1;

   NLINCON
      nlc1 = Log(SF); nlc2= (1-Temp);  nlc3=Exp(Summer-12);
      nlc4 = Exp(X1-12); nlc5=Exp(DO); nlc6=pH; nlc7=DOC;
      nlc8=STP; nlc9=Alkalinity; nlc10=1/nlc2;

     -5 <= nlc1 <= 5, nlc2 ≠ 0, 6.14421E-06 <= nlc3 <= 1.67017E-05,
      6.14421E-06 <= nlc4 <= 1.67017E-05, 1 <= nlc5 <= 1.586E+15,
      1 <= nlc6 <= 14, 6.14421E-06, 0<= nlc7 <= 100, 0 <= nlc8 <= 5,
      0 <= nlc9 <= 500;

   TP = ((beta1 + beta2*nlc7 + exp(-beta3*nlc5) +(log(beta4^2)*nlc1)
          + beta5*nlc6+ beta6*nlc8 + beta7*nlc3 + beta8*nlc4
          + beta9*nlc10 + beta10*nlc9));
 RUN;
```

Neural Networks vs Genetically Optimized Neural Networks in Time Series Prediction

Ion Railean[1,2] Sorin Moga[1] Monica Borda[2] Cristina Stolojescu[1,3]

[1] Institut TELECOM; TELECOM Bretagne, UMR CNRS 3192
Lab-STICC; Université européenne de Bretagne, France
(e-mail: firstname.lastname@telecom-bretagne.eu)
[2] Technical University of Cluj-Napoca, Romania
Faculty of Electronics, Telecommunications and Information Technology
(e-mail: firstname.lastname@com.utcluj.ro)
[3] Politehnica University of Timisoara, Romania
Faculty of Electronics and Telecommunications
(e-mail: firstname.lastname@etc.upt.ro)

Abstract. This paper deals with methods for finding the suitable weights in an Artificial Neural Network (ANN) using Genetic Algorithms (GA). We study the weakness and strengthness of the proposed approach in case of a statistical data forecasting. We describe a different approach when using the input data during optimization phase. Besides GA, we applied stationary wavelet transform (SWT) as a signal preprocessing, and time-delay neural networks (TDNN) approach for the system's inputs. Our results show that this optimization is suitable only for certain purposes in case of a statistical data prediction.
Keywords: Genetic Algorithms, Artificial Neural Networks, forecasting.

1 Introduction

The optimization of Artificial Neural Networks using Genetic Algorithms applied in forecasting have been proposed in many papers [1], [5], [6]. In [1] is presented the way of determining the optimal size of the hidden layer and the number of connections between layers. In [2] an approach using genetic computing is given, used for establishment of the optimum number of layers and the number of neurons on layer, for a given problem. A proposal of an intelligent algorithm to select the optimal architecture for ANN model in hot rolling process based on GA is shown in [3]. Venkatesan [4] proves the importance of the accuracy of algorithm-based ANN model for the turning process in manufacturing industry. The simultaneous optimization of the network architecture and the training of weights is presented in [7]. Most of the papers present the use of optimized ANNs in forecasting only in industrial processes, which are described by well-predefined formulas and the selection of parameters is required, and do not depend on statistical and human behavior.

In this paper we try to understand the influence of the ANN optimization using GA in a domain implying statistical data: WiMAX network traffic. We make a comparison between prediction accuracy of the optimized and un-optimized ANNs. Our optimization consists in setting the weights of the neural networks. In comparison to other researchers, we propose a new approach in selection of the training data.

Another aspect, is that we use wavelet transform as a signal preprocessing, and the ANN optimization is done for each of the signal's decomposition level.

The rest of the paper is organized as follows. The sections 2, 3, and 4 describe the basic aspects of the GA, ANN, and SWT. The section 5 shows the simplified forecasting framework used in our analysis. The experiments, results, and the comparison between regular ANN training and optimized ANN are given in section 6. While section 7 contains the main conclusions of the current research.

2 Genetic Algorithms

A GA is a search technique for optimization and machine learning applications. It is based on natural selection, the process that drives biological evolution. It consists of a set of individual elements (the population). At each step, the GA selects individuals randomly from the current population to be parents and uses them to produce the children for the next generation. Over successive generations, the population "evolves" toward an optimal solution. There are several steps in a GA:

- Encoding technique: gene, chromosome
- Initialization procedure: creation
- Evaluation function: environment
- Selection of parents: reproduction
- Genetic operators: mutation, recombination
- Parameter settings: practice and art

The population members are strings or chromosomes. The GA selects a subset (usually pairs) of solutions from a population, called parents, and combines them to produce new solutions called children or offsprings. The rules of combination to yield children are based on the genetic notion of crossover, which consists of interchanging solution values of particular variables. There are also occasional operations such as random value changes, which are called mutations. The children produced by the mating of parents, and that pass a survivability test, are then available to be chosen as parents for the next generation.

3 Artificial Neural Networks

The ANN is a mathematical model that simulates the structure and functions of the real biological neural networks. It is composed by interconnected simple elements, called artificial neurons. An ANN is characterized by three things:

- Its architecture: the pattern of nodes and connections between them
- Its learning algorithm, or training method: the method for determining the weights of the connections
- Its activation function: the function that produces an output based on the input values received by a node

The two most important types of ANNs are feed-forward (FFANN) and recurrent networks (RANN). A FFANN has its neurons organized in a layered structure. Each layer consists of units which receive their input from the units situated on a layer directly below and send their output to the units from a layer directly above. RANN are characterized by the fact that they contain feedback connections and they take into consideration the dynamical properties of the network.

In this paper we used feed-forward ANN, and we discuss the setting of weights of the connections. One way is to set them explicitly, using a priori knowledge. Another way is to "train" the ANN by feeding it teaching patterns and letting it change its weights according to some learning rule. While our approach, consists in applying GA to find the optimal weights between the input and the hidden layer.

4 The wavelet analysis

Multi-resolution analysis (MRA) is a signal processing technique that takes into account the signal's representation at multiple time resolutions. Using wavelet MRA, the collected measurements can be smoothed until the overall long-term trend is identified. Fluctuations around the obtained trend are further analyzed at multiple time scales. The level of decomposition depends on the length of the data set (the number of values). At each temporal resolution two categories of coefficients are obtained: approximation and detail coefficients. We used the à trous methodology in MRA implementation, also known as Shensa's algorithm [9], which corresponds to the computation of the Stationary Wavelet Transform (SWT).

The à trous wavelet transform decomposes a signal X_t as follows:

$$X_t = a_{p,t} + \sum_{j=1}^{p} d_{j,t} \tag{1}$$

where $a_{p,t}$ represents the smooth version of the original signal (the approximation at the p^{th} level of decomposition), while $d_1 ... d_p$ represent the details of X_t at scale 2^{-j}. This equation can be seen as a multiple linear regression model also, where the original signal is expressed in terms of its coefficients. Among different mother wavelets (Daubechies, Symlet, Meyer, etc. [8]), we used Daubechies 2 wavelet.

5 Forecasting framework

The simplified forecasting framework of our analysis is presented in Figure 1.It implies a series of steps as presented below:

1. Use SWT to decompose the data for input and for test
2. Apply an Artificial Neural Network for each level of decomposition obtained from the input and build the forecasting model. Choose between having or not the GA Optimization for the ANN
3. Use the decomposed test data and the obtained model in order to predict each decomposition level of the future forecasted signal
4. Use the Inverse SWT in order to obtain the final predicted signal

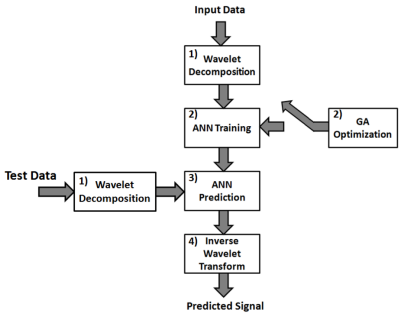


Fig. 1. The main block diagrams of the forecasting framework. Each block represents one of the steps taken in the construction of our model

6 Experiments and results

6.1 Data analysis

Our WiMAX traffic data used in analysis was obtained by monitoring the traffic from 67 Base Stations (BS) during eight weeks, from March 17th till May 11th, 2008. Each BS has its own data set which is composed of numerical values representing the total number of packets from the uplink channel. Each value is recorded every 15 minutes. It results that for a given BS we have the following number of samples: 96 samples/day, 672 samples/week, and a total number of 5376 samples.

6.2 ANN approach

The goal in our experiments was to make one day ahead forecasting. Taking into account this information we used for ANN's architecture only one neuron for the output, which consists of an array of 96, 32, or 16 samples. 32 and 16 samples were obtained by making a downsampling of the signal with 3 and 6. These downsamplings were done because of the existence of observed periodicities in the WiMAX traffic [13]. Regarding the number of layers, we used one hidden layer network. In [2] is pointed out the fact that one hidden layer network is able to approximate most of the nonlinear functions demanded by practice.

For the input layer, we used the approach of TDNN described in [10], [11], and [7]. The time-delay of the input information was set to 4, 8, 12, and 24 hours shifting. The data used for the input layer was the wavelet transform obtained from the weeks 1-6 during training process, and weeks 2-7 during test. Also, we did not apply all the data at ANN's inputs, we used only the days corresponding to the same period of the week as the forecasted day. The number of neurons for the hidden layer was 2. The training algorithm was the combination between adaptive learning rate with momentum.

6.3 Genetically Optimized ANNs

For the optimized neural networks we used an approach permitting us to train a given ANN using two data sets at the same time. In the first part we used the information from weeks 1-5 while having as a target the given day from week 6, and in the second part we used the information from weeks 2-6, while having as a target the information from week 7. During testing process, we applied at the optimized ANN's inputs the wavelet transform from corresponding to the days from weeks 3-7. An example of this training is given in Figure 2 (because of the space, we present in the figure a simplified optimization using only 2 days at the input). The final predicted signal, obtained after applying inverse stationary wavelet transform on all forecasted sequences, was compared to the original signal from 8th week.

The designing of training the ANN using the Genetic Algorithms is as follows:

- each individual contains a set of weights for all the links between layers
- each gene represents a single weight
- we had a population size of 100 individuals, meaning 100 different possibilities at each generation for the network
- the number of generations is 100: less generations resulted in not finding an acceptable solution for our problem, while more generations resulted in a longer time processing. However, above this value, we didn't manage to observe better performance of the final results
- the fitness function is calculated as follows:

$$F = \frac{1}{N}\sum_{i=1}^{N}(x_i^{f_1} - x_i^{o_1}) + \frac{1}{N}\sum_{i=1}^{N}(x_i^{f_2} - x_i^{o_2}),$$

where N is the number of samples, $x_i^{o_1}$and $x_i^{o_2}$ represent certain level decomposition of the original signal used for inputs, while $x_i^{f_1}$ and $x_i^{f_2}$ are the output targeting signals for the two data sets.

6.4 Evaluation criteria

In order to evaluate the prediction performance between ANNs and genetically optimized ANNs, we used the following well-known evaluation criteria: Symmetrical Mean Absolute Percentage Error (SMAPE) and R-Square (RSQ):

- SMAPE: calculates the symmetric absolute error in percent between the actual X and the forecast F across all observations t of the test set of size n for each time series s:

$$SMAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|X_t - F_t|}{(X_t + F_t)/2} \tag{2}$$

- RSQ: the coefficient of determination R^2, in statistics, is the proportion of variability in a data set that is accounted for by a statistical model. In this definition, the term *variability* is defined as the sum of squares:

$$R^2 = \frac{SS_R}{SS_T} = \frac{\sum_t (X_t - \overline{X_t})^2}{\sum_t (F_t - \overline{F_t})^2} \tag{3}$$

in which X_t, F_t are the original data values and predicted values respectively, while $\overline{X_t}$ and $\overline{F_t}$ are the means of the observed data and modeled (predicted) values, respectively. SS_T is the total sum of squares, SS_R is the regression sum of squares. In the ideal case the value of RSQ is 1, while SMAPE must be 0.

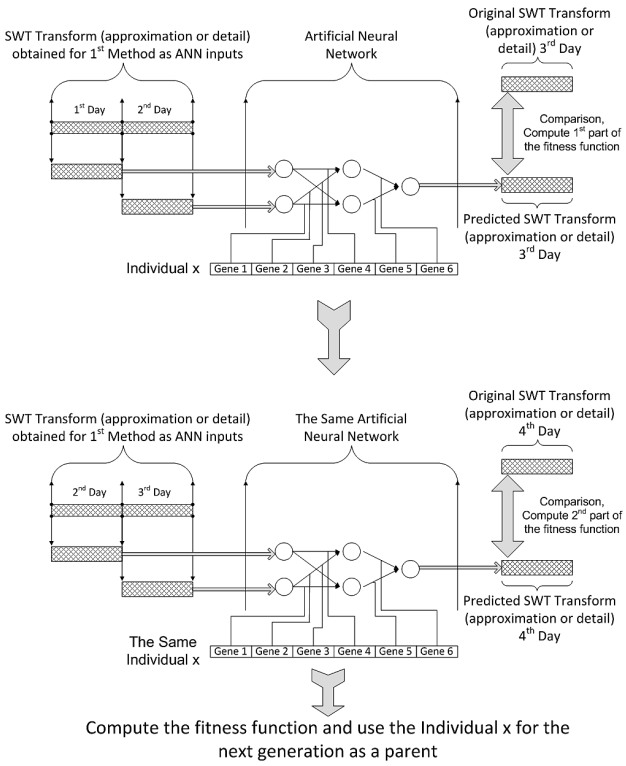


Fig. 2. Example of ANN Optimization. The top diagram represents the computation of the first part of the fitness function, while the bottom diagram exemplifies the computation of the second part of the fitness function

6.5 Results

In both ANN and genetically optimized ANNs we made 8232 simulations from 11256 possible (after extracting the erroneous data). We made a combination between all the possibilities in choosing the number of BS (from a total of 67), number of samples per day (16, 32, or 96), time-delay interval for inputs (4, 8, 12, or 24 hours shifting), and the day of the week (from Monday till Sunday). The results for RSQ and SMAPE values are presented in the tables 1 and 2. They present

Time Delay	Measured Value	96 samples	32 samples	16 samples
4 hours	RSQ	1.212	1.249	1.317
	SMAPE	0.905	0.927	0.802
8 hours	RSQ	**1.159**	**1.120**	**1.205**
	SMAPE	**0.886**	**0.858**	**0.756**
12 hours	RSQ	1.330	1.219	1.287
	SMAPE	0.904	0.924	0.814
24 hours	RSQ	1.178	1.201	1.263
	SMAPE	0.910	0.946	0.780

Table 1. RSQ and SMAPE values (one day prediction) using ANN. The best values are represented in bold

the medium value of the results from all 67 BS and days of the week during a given configuration of samples per day and the number of shifted hours.

By comparing the results, we can see that the SMAPE values in case of GA optimization are not better then the ones of the usual ANN training. However, the RSQ value is closer to the ideal 1. This is true for all used configurations of TDNN.

Time Delay	Measured Value	96 samples	32 samples	16 samples
4 hours	RSQ	1.127	1.200	1.176
	SMAPE	1.001	**0.913**	0.867
8 hours	RSQ	**1.108**	**1.129**	**1.099**
	SMAPE	**0.983**	0.924	**0.820**
12 hours	RSQ	1.211	1.155	1.168
	SMAPE	1.008	0.951	0.851
24 hours	RSQ	1.087	1.189	1.215
	SMAPE	1.067	0.950	0.872

Table 2. RSQ and SMAPE values (one day prediction) using Genetically Optimized ANN. The best values are represented in bold

7 Conclusions

In this paper we presented a comparison between neural networks and optimized neural networks used in statistical data forecasting. We proposed a combination between wavelet transform, time delay neural networks, and a different approach of using the input data when applying GA for our ANN training.

Our results show that in case of the optimization the SMAPE value increased with about 2-7% in comparison to the regular ANN training, while the RSQ value decreased with about 2-10%. It means that the optimized ANN is able to express better the variability of the statistical data. This is because it takes into consideration a longer time interval in optimization, as we managed to use two data sets at the same time during training process. However, the value of SMAPE is not better. The

reason for these differences might be as follows: according to the RSQ formula, we make a ratio between the variabilities according to the medium values of predicted and real signal; while SMAPE presents the differences between predicted and real samples of the signal. This means that in case of optimized networks we have a more shifted medium value from the medium of the original signal in comparison to the regular ANNs, because it keeps the behavior of the data from twice longer time then in case of the other approach. While regular ANN, expresses the behavior closer to the data we want to predict, and it does not necessarily make a generalization of the earlier time intervals.

One of our future task is to test our approach on other non-statistical data sets. We will try also to integrate the obtained model of calculating more precisely the data variability in other methods of prediction. Also, we will use it especially in data classification, because we have obtained an approach that uses multiple data sets during network training, while in case of regular network training we would need a more complex architecture for the ANN in order to obtain the desired results.

Acknowledgements

We would like to thank Alcatel Lucent Timisoara who sent the data used in our research. We gratefully appreciate their effort and assistance.

References

1. R. Garcia Ojeda et al., *Genetic Algorithms in the Optimal Choice of Neural Networks for Signal Processing*, IEEE, 0-7803-2972-4, 1996
2. I. Ileana et al., *The Optimization of Feed Forward Neural Networks Structure using Genetic Algorithms*, ICTAMI, Thessaloniki, Greece, 2004
3. J. S. Son et al., *A study on genetic algorithm to select architecture of a optimal neural network in the hot rolling process*, Journal of Materials Processing Technology 153-154, pp. 643-648, 2004
4. D. Venkatesan et al., *A genetic algorithm-based artificial neural network model for the optimization of machining processes*, Neural Comput and Applic 18, pp. 135-140, 2009
5. J. C. Yu, Y. L. Tseng, *Evolutionary Engineering Optimization Using Recursive Regional Neural Network and Genetic Algorithm*, IEEE, 0-7695-2882-1/07, 2007
6. A. Fiszelew et. al., *Finding Optimal Neural Network Architecture Using Genetic Algorithms*, Research in Computing Science 27, pp. 15-24, 2007
7. B. T. Zhang, H. Muhlenbein, *Evolving Optimal Neural Networks Using Genetic Algorithms with Occam's Razor*, Complex systems 7, pp. 199-228, 1993
8. S. Mallat, *A Wavelet Tour of Signal Processing*, Second Edition, (1999)
9. M.J.Shensa, *Discrete Wavelet Transform. Wedding the a trous and Mallat algorithms,* IEEE Transactions and Signal Processing, 40, pp. 2464-2482,(1992)
10. T. Taskaya-Temizel et al., *Configuration of Neural Networks for the Analysis of Seasonal Time Series*, ICAPR 3, Lecture Notes in Computer Science, vol. 1, pp. 297-304, 2005
11. D. S. Clouse et al., *Time-Delay Neural Networks: Representation and Induction of Finite-State Machines*, IEEE Transaction on Neural Networks, vol. 8, no. 5, September 1997
12. G. P. Zhang, M. Qi, *Neural network forecasting for seasonal and trend time series*, European Journal of Operational Research 160, pp. 501-514, 2005
13. C. Stolojescu et al., *Forecasting WiMAX BS Traffic by Statistical Processing in the Wavelet Domain*, in Proceedings of ISSCC, Iasi, Romania, pp. 177-183, 2009

SMTDA 2010: Stochastic Modeling Techniques and Data Analysis
International Conference, *Chania, Crete, Greece, 8 - 11 June 2010*

Grouping Ordinal Variables by Using Fuzzy Cluster Analysis

Hana Rezankova, Dusan Husek, Michaela Rysankova

Department. of Statistics and Probability, University of Economics, Prague
Prague, Czech Republic
Email: hana.rezankova@vse.cz, michaela.rysankova@vse.cz
Institute of Computer Science, Academy of Sciences of the Czech Republic
Prague, Czech Republic
Email: dusan.husek@cs.cas.cz

Abstract: Over recent years data mining has been establishing itself as one of the major disciplines in computer science with growing industrial impact. Data sets usually carry information about objects in the form of features sets encoded as vectors. These are now a day as a rule high-dimensional, so there is even necessary to reduce its dimension in advance. Here we study problem of searching groups of similar variables which are of ordinal type. As an example we analyzed the data from the survey on "Active lifestyle of university students". Variables expressing a satisfaction concerning different points of view of the students' life were included into analysis. Here we suggest application of non-metric multidimensional scaling and categorical principal component analysis followed by the obtained results interpretation using fuzzy cluster analysis. The soft version of CSPA (cluster-based similarity partitioning algorithm) is applied for ensembles of fuzzy clustering results obtained on the basis of different techniques.
Keywords: Dimensionality reduction, Ordinal variables, Multidimensional scaling, Categorical principal component analysis, Fuzzy cluster analysis, Silhouette plot

1 Introduction

Modern automated methods for measurement, collection, and analysis of data in all fields of science, industry, and economy are providing more and more data with drastically increasing complexity of its structure. These data carry objects features encoded as high-dimensional vectors, so there is as a rule necessity to reduce its dimension before an analysis. It is worth to mention some examples of such data: a textual document characterized its vocabulary; a web page characterized by graphical user interface patterns; or a respondent characterized by its opinions. In the last case, objects are described by variables of ordinal type.
In this study we suggest the procedure for dimensionality reduction which consists of application of two methods followed by the obtained results interpretation using fuzzy cluster analysis. We apply non-metric multidimensional scaling on the basis of Kendall's tau-b for creation the similarity matrix and categorical principal component analysis. In fuzzy cluster analysis we determine the optimal number of

clusters by using average silhouette widths. The soft version of CSPA (cluster-based similarity partitioning algorithm) is applied for ensembles of fuzzy clustering results obtained on the basis of different techniques. The final assignment of variables to the groups is graphically presented by a silhouette plot.

2 Similarity Measures for Ordinal Variables

In the process of searching groups of similar variables, coefficients of dependency are usually applied as similarity measures, see Rezankova (2009). Dependency of the ordinal variables is denoted as a rank correlation and their intensity is expressed by correlation coefficients. The best known among them is Spearman's correlation coefficient. Let us have the n x p data matrix $\mathbf{X}$ with the elements x_{ij} where n is a number of objects and p is a number of variables. If investigated ordinal variables $\mathbf{X}_g$ and $\mathbf{X}_h$ express the unambiguous rank, the following formula can be used for *Spearman's correlation coefficient*:

$$r_S = 1 - \frac{6 \cdot \sum_{i=1}^{n} (x_{ig} - x_{ih})^2}{n(n^2 - 1)}.$$

If this assumption is not satisfied, the process described e.g. in Rezankova (2010) must be applied.

Other measures investigate pairs of objects. If, in a pair of objects, the values of both investigated variables are greater (less) for one of these objects, this pair is denoted as concordant. If for one variable the value is greater and for the second one it is less, then the pair is denoted as discordant. In other cases (the same values for both objects exist for at least one variable), the pairs are tied. For the sake of simplification, we will use the following symbols:

Γ – a number of concordant pairs,

Δ – a number of discordant pairs,

Ψ_g – a number of pairs with the same values of variable $\mathbf{X}_g$ but distinct values of variable $\mathbf{X}_h$,

Ψ_h – a number of pairs with the same values of variable $\mathbf{X}_h$ but distinct values of variable $\mathbf{X}_g$.

On these numbers of pairs, *Kendall's tau-b* (Kendall's coefficient of the rank correlation) is based, for example. It is expressed as

$$\tau_b = \frac{\Gamma - \Delta}{\sqrt{(\Gamma + \Delta + \Psi_g)(\Gamma + \Delta + \Psi_h)}}.$$

It is a symmetric measure, as well as Spearman's correlation coefficient.

Some other measures and their features are described in Rezankova (2009) and Rezankova (2010).

3 Methods for Searching Groups of Similar Variables

There are several types of techniques which are assigned to dimensionality reduction methods. Some of them are based of the projection of a high-dimensional space into a low-dimensional space. Usually, an object characterized by the vector of values of variables is plotted as a point in two-dimensional space where two of the found dimensions are used. Similarly, values of components or new dimensions can be calculated for variables which can then be plotted in two-dimensional space by means of a dot graph. If groups of variables exist, it can be seen in this graph. Multidimensional scaling and categorical principal component analysis are examples of such techniques.

Multidimensional scaling, see Torgerson (1952), Cox and Cox (2001), is based on the proximity matrix. Nonmetric multidimensional scaling, see Kruskal (1964) is used for the further analysis. It both finds a non-parametric monotonic relationship between the dissimilarities in the proximity matrix and the Euclidean distance between variables, and the location of each variable in the low-dimensional space. The user must pre-specify number of dimensions.

Categorical principal component analysis, see De Leeuw et al. (1976), quantifies categorical variables using optimal scaling, resulting in optimal principal components for the transformed variables. The variables can be given mixed optimal scaling levels and no distributional assumptions about the variables are made. This type of analysis can easily deal with nonlinear relationships between the variables to be analyzed.

4 Algorithms for Fuzzy Clustering and Visualization Results

Fuzzy clustering is studied very intensively in last decades. A lot of papers in journals and proceedings from conferences and also some monographs have been published, e.g. Abonyi and Feil (2007) and Hoppner et al. (2000). There are some various techniques for fuzzy (soft) cluster analysis. One of them is the fuzzy *k*-means algorithm, see e.g. Kruse et al. (2007). It is a generalization of the classical (hard) *k-means* (also HCM – *hard c-means*).

Let $\mathbf{x}_i$ be a vector of feature values, which characterizes the *i*th object. Then the distance between the *i*th and *j*th objects can be calculated as Euclidean distance between vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ for example (in the following text we will consider an object and a representing vector as synonyms), i.e.

$$d_{\mathrm{E}}(\mathbf{x}_i, \mathbf{x}_j) = d_{ij} = \sqrt{\sum_{l=1}^{m} (x_{il} - x_{jl})^2} = \left\| \mathbf{x}_i - \mathbf{x}_j \right\|$$

here m is a number of variables.
We suppose that the data set consisting of n objects, i.e. $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, should be partitioned into k clusters $C_1, \ldots, C_k$. In some algorithms, the representative of each cluster is determined. This can be either one object from a cluster (so called *medoid*) or a new vector characterizing the center of a cluster (*centroid*).
In the latter case, the *centroid* is usually created by average values of individual variables. Further on, the centroid of the hth cluster will be denoted as $\overline{\mathbf{x}}_h$. Then the distance between the ith object and the corresponding centroid can be expressed as

$$d_{\mathrm{E}}(\mathbf{x}_i, \overline{\mathbf{x}}_h) = d_{ih} = \left\| \mathbf{x}_i - \overline{\mathbf{x}}_h \right\|.$$

The *fuzzy k-means* (frequently FCM – *fuzzy c-means*) algorithm minimizes the objective function

$$J_{\mathrm{FCM}} = \sum_{h=1}^{k} \sum_{i=1}^{n} u_{ih}^{q} d_{ih}^2$$

where k is a number of clusters, the elements $u_{ih} \in [0, 1]$ are membership degrese, d_{ih} is the distance between the j-th object and the center of the h-th cluster, and the parameter q ($q > 1$) is called the fuzzifier or weighting exponent (usually $q = 2$ is chosen). Further, the following conditions have to be satisfied:

$\sum_{h=1}^{k} u_{ih} = 1$ for $i = 1, \ldots, n$ and $\sum_{i=1}^{n} u_{ih} > 0$ for for $h = 1, \ldots, k$.

In some cases, the proximity matrix is only known instead of original vectors of feature values. In this matrix, each pair of objects is numerically described by a real-valued relation. Cluster analysis based on the proximity matrix is sometimes called *relation clustering*, see Runkler (2007).
One algorithm is *relation fuzzy k-means* (RFCM – *relation fuzzy c-means*). For $q = 2$ it is called FANNY, see Kaufman and Rousseeuw (2005), and it is implemented in the statistical software system S-PLUS. In this algorithm, the objective function

$$J_{\mathrm{FANNY}} = \sum_{h=1}^{k} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} u_{ih}^2 u_{jh}^2 d_{ij}}{2 \sum_{j=1}^{n} u_{ih}^2}$$

is minimized. In this function, d_{ij} is the known distance between the ith and jth objects, and u_{ih} and u_{jh} are unknown.
Behind of some special graphical techniques, see Wiswedel et al. (2007), traditional graphs for disjunctive clustering can be also used for a representation of memberships. One of them is a *silhouette plot*, see Kaufman and Rousseeuw (2005), which is implemented in the S-PLUS system. The width and the direction of the rectangles for the ith object is determined by the value

$$\psi_i = \frac{\eta_i - \mu_i}{\max\{\eta_i, \mu_i\}} \text{ where } \eta_i = \sum_{j \in C_g} d_{ij} \Big/ (n_g - 1),\ \mu_i = \min_{h \neq g} \left(\sum_{j \in C_h} d_{ij} \Big/ n_h \right),$$

and n_g is a number of objects in the gth cluster. It is supposed that the object is assigned only to one cluster (C_g) according to the highest value of a membership degree. It means that $u_{ig} = \max_h(u_{ih})$.
For the cluster number determination, we can use the average silhouette width $\overline{\psi}$. The higher value represents better partitioning objects to clusters.

5 Ensembles of Fuzzy Clustering

Sometimes, the user has results of clustering (assignments of individual objects to the certain number of clusters) obtained by different way and he has not any access the original features of the objects. For example in marketing research customers are segmented in multiple ways based on different criteria (need-based, demographics, etc.). The user can be interested in obtaining a single, unified segmentation.
Combining clustering is more difficult than combining the results of multiple classifiers. Before combining the clustering one has to identify which clusters from different clusterings correspond to each other. Moreover, the number of clusters in individual solutions might vary, see Punera and Ghosh (2007).
For results of hard clustering, graph-theoretic approaches have been proposed in the literature. They are based on the hypergraph representation of clustering, see Table 1. In this table $C_h^{(q)}$ denotes the hth cluster in the qth clustering and $u_{ih}^{(q)}$ represents the membership degree of the ith object to the hth cluster in the qth clustering. In the case of hard clustering $u_{ih}^{(q)} \in \{0, 1\}$.

Table 1. Hypergraph representation of clustering

cluster	$C_1^{(1)}$	$C_2^{(1)}$	…	$C_{k^{(1)}}^{(1)}$	…	$C_1^{(r)}$	$C_2^{(r)}$	…	$C_{k^{(r)}}^{(r)}$

object									
$\mathbf{x}_1$	$u_{11}^{(1)}$	$u_{12}^{(1)}$	…	$u_{1k^{(1)}}^{(1)}$	…	$u_{11}^{(r)}$	$u_{12}^{(r)}$	…	$u_{1k^{(r)}}^{(r)}$
$\mathbf{x}_2$	$u_{21}^{(1)}$	$u_{12}^{(1)}$	…	$u_{1k^{(1)}}^{(1)}$	…	$u_{21}^{(r)}$	$u_{12}^{(r)}$	…	$u_{1k^{(r)}}^{(r)}$
…	…	…	…	…	…	…	…	…	…
$\mathbf{x}_n$	$u_{n1}^{(1)}$	$u_{n2}^{(1)}$	…	$u_{1k^{(1)}}^{(1)}$	…	$u_{n1}^{(r)}$	$u_{n2}^{(r)}$	…	$u_{1k^{(r)}}^{(r)}$

Cluster-based similarity partitioning algorithm (CSPA) is an example of the graph-theoretic approach. In the CSPA technique a similarity matrix is computed as $\mathbf{W} = (1/r)\mathbf{U}\mathbf{U}^{\mathrm{T}}$ where r is a number of clusterings. Then a clustering algorithm based on a proximity matrix can be used.
In *soft version of CSPA* (sCSPA), one can use either the $\mathbf{U}\mathbf{U}^{\mathrm{T}}$ matrix or the similarity matrix created on the basis of Euclidean distance. In the latter, the distance between objects $\mathbf{x}_i$ and $\mathbf{x}_j$ is calculated as

$$d_{\mathrm{E}}(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{\sum_{q=1}^{r} \sum_{h=1}^{k^{(q)}} \left(u_{ih}^{(q)} - u_{jh}^{(q)}\right)^2} .$$

On the basis of the proximity matrix the objects can be clustered into k clusters.

6 Applications to Real Data File

In this section, the analysis of one real data file will be described. The file is the result of the research "Active lifestyle of university students" which was realized at the Faculty of Electrical Engineering of the Czech Technical University in Prague by dr. Z. Valjent from the Institute of Physical Education and Sports of this university in 2008, see Valjent (2009) and Valjent and Flemr (2009). This file contains answers from 1 453 respondents.
For the purpose of investigation of relationships between variables, 15 variables expressing a satisfaction concerning different points of view of the students' life were analyzed. Respondents evaluated their satisfaction on the scale from 1 (no satisfaction) to 7 (very satisfied).
First, the similarity matrix based on Kendall's tau-*b* was created in the SPSS system. This matrix was transformed to the dissimilarity matrix by subtraction of the values from 1 in Microsoft Excel. The transformed matrix was analyzed by non-metric multidimensional scaling (MDS) in the STATISTICA system. Further, principal component analysis (PCA) for the categorical data (the CATPCA procedure) in the SPSS system was applied. In both cases, three dimensions were considered (on the basis plot interpretation advisability). Both dimension values

obtained by non-metric MDS and component loadings obtained by the CATPCA procedure were analyzed by the FANNY method in the S-PLUS system.
The suitable number of clusters was searching on the basis of average silhouette widths. In both cases, four clusters were identified as optimal, see Table 2.
The membership degrees matrices for four clusters for both cases were joined into one matrix as in Tables 1 is shown. However, instead of objects, the variables in the rows were characterized. The data from this matrix were further clustering with use of the FANNY algorithm. The resulting clustering is displayed in Fig. 1.
We obtained the following view on the relationships between variables. The best cluster is formed by variables V17 and V18 (satisfaction with partner in life and sex life). The fourth cluster is formed by the pair of variables V19 and V20 (study generally and studying results) and the variable V21 (financial situation), which differs a little. The second cluster is represented by the close variables V15 and V16 (family life and friends), variables V14 and V26 (success and acknowledgments from others and total quality of life), and variable V22 (housing). In the first cluster there are variables V25 (fitness), V24 (health state) and V13 (body weight), and also the variables V12 (visage) and V23 (leisure time).

Table 2. Average silhouette widths characterizing results obtained by algorithm FANNY (for dimension values obtained by non-metric MDS and for component loadings obtained by the CATPCA procedure)

Methods	Number of clusters				
	2	3	4	5	6
MDS	0.22	0.22	**0.39**	0.36	0.31
PCA	0.30	0.41	**0.56**	0.49	0.48

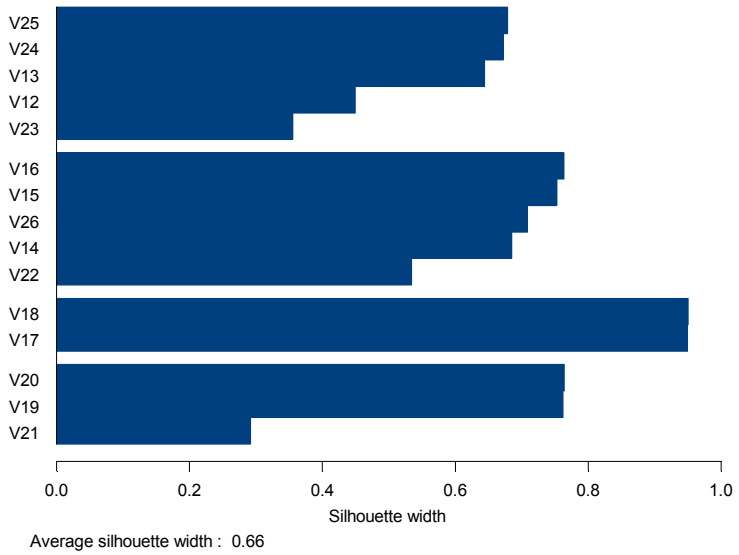


Fig. 1. Silhouette plot for four clusters of variables (FANNY algorithm for the combination of clustering results)

7 Conclusions

For dimensionality reduction is useful to discover groups of similar variables. This can be achieved using different clustering methods. In the complex task presented here we solved the problem of interpretation of dimension values obtained by non-metric MDS and component loadings obtained by the CATPCA procedure. For this, we used fuzzy cluster analysis, including cluster number identification, following by ensembles of obtained fuzzy clustering and graphical presentation of the results by the silhouette plot.

When we used fuzzy cluster analysis, rows of the input matrix were considered as the variables and the columns as the dimensions (components). Using both, the fuzzy cluster analysis and the silhouettes plot one can identify as more similar as less similar variables in the clusters. In the former it is possible by membership coefficients and in the latter by silhouette widths.

Using CATPCA we obtained the higher values of membership degrees. The average silhouette widths used for cluster number determination are also higher in the case of CATPCA. This means that by categorical principal components analysis we obtained better clusters than by multidimensional scaling based on Kendall's tau-b.

Acknowledgement. This work was supported by projects AV0Z10300504, GACR P202/10/0262, 205/09/1079, MSM6138439910, and IGA VSE F4/3/2010.

References

1. Abonyi, J. and Feil, B., *Cluster Analysis for Data Mining and System Identification*. Birk-häuser Verlag AG, Berlin (2007).
2. Cox, T. F. and Cox, M. A. A.: *Multidimensional Scaling*, *Second Edition*. Chapman & Hall/CRC, New York (2001).
3. De Leeuw, J., Young, F. W., and Takane, Y., Additive structure in qualitative data: an alternating least squares method with optimal scaling features. *Psychometrika*, **31,** 33-42 (1976).
4. Hoppner, F., Klawon, F., Kruse, R., and Runkler, T., *Fuzzy Cluster Analysis. Methods for Classification, Data Analysis and Image Recognition*. John Wiley & Sons, New York (2000).
5. Kaufman, L. and Rousseeuw, P., *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons, Hoboken (2005).
6. Kruse, R., Döring C., and Lesot, M.-J., Fundamentals of Fuzzy Clustering, in *Advances in Fuzzy Clustering and Its Applications* (J. V. Oliveira and W. Pedrycz, eds.), John Wiley & Sons, Chichester, 3-30 (2007).
7. Kruskal, J. B., Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, **29**: 115–129 (1964).
8. Punera, K. and Ghosh, J., Soft Cluster Ensembles, in *Advances in Fuzzy Clustering and Its Applications* (J. V. Oliveira and W. Pedrycz, eds.), John Wiley & Sons, Chichester, 69-91 (2007).
9. Rezanková, H., Cluster analysis and categorical data. *Statistika*, **89**(3), 216-232 (2009).
10. Rezanková, H., *Analysis of Data from Questionnaire Surveys, Second Edition* (in Czech). Professional Publishing, Prague (2010).
11. Runkler, T. A., Relational Fuzzy Clustering, in *Advances in Fuzzy Clustering and Its Applications* (J. V. Oliveira and W. Pedrycz, eds.), John Wiley & Sons, Chichester, 31-51. (2007).
12. Torgerson, W. S., Multidimensional scaling: I. Theory and method. *Psychometrika*, **17**, 401-419 (1952).
13. Valjent, Z., *Active Lifestyle of University Students* (in Czech). Doctoral thesis, FTVS UK, before defense.

14. Valjent, Z. and Flemr, L., Quality of life of students of the technical university (in Czech). *Acta* (2009), in press.
15. Wiswedel, B., Patterson, D.E., and Berthold, M.R., Interactive Exploration of Fuzzy Clusters, in *Advances in Fuzzy Clustering and Its Applications* (J. V. Oliveira and W. Pedrycz, eds.), John Wiley & Sons, Chichester, 123-136 (2007).

Arbitrage opportunities between NYSE and XETRA?: A comparison of simulation and high frequency data

Jörg Rieger, Kirsten Rüchardt and Bodo Vogt

Faculty of Economics and Management, Chair of Empirical Economics
Otto-von-Guericke-University Magdeburg, Magdeburg, Germany
Email:Kirsten.Ruechardt@ovgu.de

Abstract: This paper investigates the no-arbitrage condition of financial markets by comparing two stock markets: the New York Stock Exchange (NYSE) and the German Exchange Electronic Trading System (XETRA). We analyze German stocks that are traded simultaneously at both exchanges using high frequency data for XETRA, the NYSE, and the foreign exchange rates. Converting Euro-prices into Dollar-prices and vice versa reveals possibilities to discuss the efficiency of these two stock markets and arbitrage opportunities. One measure of efficiency is stock price clustering and we obtain the result that XETRA is more efficient if the exchange rate is taken into account. The observed difference in the clustering effect would not be observable, if the no-arbitrage condition held. We propose a trading strategy that exploits these differences. Furthermore, we compare our empirical findings with the results we obtain by simulating financial markets using a Random Walk as a model for the price movement.
Keywords: financial markets; simulation; no-arbitrage condition; stochastic processes

1 Introduction

When comparing different stock markets the following questions arise immediately: which stock market is more liquid or more efficient and are there arbitrage opportunities? According to the Efficient Market Hypothesis, financial markets are "informational efficient" and there are no arbitrage possibilities (Grossman 1976). In this paper, we analyze the intraday trades of selected German stocks (Daimler and Deutsche Bank) that are traded simultaneously at the New York Stock Exchange (NYSE) and at the German Exchange Electronic Trading System (XETRA). The conversion of the XETRA Euro stock prices into US-Dollar stock prices by the foreign exchange rates and vice versa enables us to discuss the question which stock market is more efficient: XETRA or the NYSE? To investigate this question we use the phenomenon of stock price clustering as a possible indication about the degree of the efficiency of a stock market. Stock price clustering describes the tendency of prices to deviate from a uniform distribution, tending instead to cluster at certain prices and avoiding others. This anomaly can be observed for different stock markets with different market structures and has been widely discussed in the literature (see for example Osborne, 1962, Niederhoffer, 1965, 1966, Ball et al., 1985, Harris, 1991, Christie et al., 1994, Kahn et al., 1999, Vogt et al., 2001, Huang and Stoll, 2001, Sopranzetti and Datar, 2002, Sonnemans, 2006).

The phenomenon of price clustering contradicts any strict definition of the Efficient Market Hypothesis and can be used for measuring the efficiency of stock markets (Ikenberry and Weston, 2008). Stock markets with a higher degree of stock price clustering are considered as less efficient stock markets. Our data yield the result of different extents of stock price clustering for stocks that are traded simultaneously at XETRA and the NYSE. We use different approaches to decide whether XETRA or the NYSE is the more efficient stock market which results in different answers. The first approach directly compares the stock prices on the two exchanges while for the second approach the foreign exchange rate is taken into account. To be more precise, the latter approach indicates that XETRA is the more efficient stock market when comparing converted XETRA-prices and actually observed NYSE-prices. But according to both approaches we observe a difference in the clustering structure between XETRA and the NYSE. The observed difference indicates a violation of the Efficient Market Hypothesis and therefore inefficiency between both analyzed stock markets. Furthermore, it puts some question on the no-arbitrage condition of financial markets. The no-arbitrage condition of financial markets implies that the Dollar-prices at the NYSE should be obtained by converting the Euro-prices at XETRA and vice versa (for companies that are traded simultaneously at both stock markets). We propose a trading strategy that exploits the differences in the observed clustering structure between converted and actually observed stock market prices (quasi-arbitrage opportunities). As these results apply to empirical intraday data of selected German stocks, we want to check whether we obtain the same results by simulating the stock markets. For this purpose, we use the Random Walk as a model for the price movement. The simulated data is in line with the Efficient Market Hypothesis and the no-arbitrage condition as well. Although the assumptions of applying a Random Walk as a model describing our empirical data are fulfilled, we observe substantial differences in the clustering structure. These results reinforce our empirical findings.

2 Empirical Approach

2.1 Data Description

We use high frequency data (all intraday trades) of German stocks (Daimler and Deutsche Bank) that are traded simultaneously at XETRA and the NYSE in November and December 2004 (15th of November through 29th of December). The data is obtained from the Trade and Quote (TAQ) database of the NYSE and the XETRA stock market. In addition, we use high frequency data of the foreign exchange rate Euro vs. US-Dollar. This data was recorded by using a computer program. In 2004, stock prices at XETRA were listed in Euros with a tick size (smallest trading unit or minimum price variation) of 1 Euro-cent while at the NYSE prices were listed in Dollar with a tick size of 1 Dollar-cent. To be more precise, we analyzed all intraday trades of 30 trading days and a time period between 3:30 pm and 5:30 pm for XETRA and 9:30 am to 11:30 am for the NYSE.

2.2 Empirical Results

One approach to answer the question whether XETRA or the NYSE is more efficient is analyzing the last digits of the stock prices of Daimler and Deutsche Bank at both stock markets. We obtain the following frequency distributions.

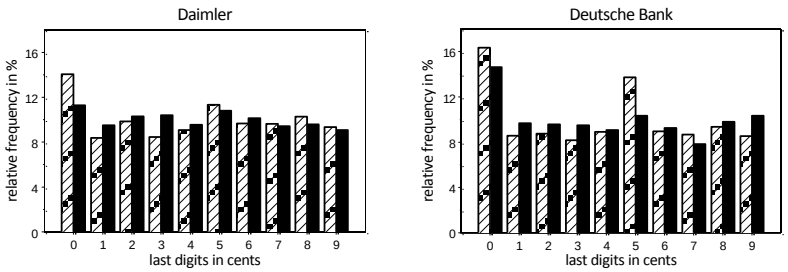


Fig. 1. Frequency distribution of the last digits of the stock prices of Daimler and Deutsche Bank at XETRA (dashed bar) and at the NYSE (solid black bar).

As can be seen from Figure 1, clustering exists and seems to be more pronounced at XETRA. This first impression of different extents of price clustering can more formally be tested by applying a measure *D* of clustering that is used in Ikenberry and Weston (2008). Under standard regularity conditions, the statistic *D* is Chi-squared distributed where large values of *D* imply a significant deviation from the expected distribution (uniform distribution). The test statistic *D* can be calculated for both XETRA (D_{XETRA}) and the NYSE (D_{NYSE}). However, this test does not address whether XETRA is more or less clustered compared to the NYSE. For this purpose, in a second step Ikenberry and Weston (2008) suggest comparing D_{XETRA} and D_{NYSE} by examining the ratio $\tilde{D}$ between both (it is assumed that the numerator has to be greater (or equal) than the denominator, otherwise the inverse has to be calculated). The statistic $\tilde{D}$ is *F*-distributed and enables us to test whether the degree of stock price clustering is the same for XETRA and the NYSE. Large values of $\tilde{D}$ imply that price clustering at XETRA is greater compared to the NYSE. Table 1 presents the numerical values of D_{XETRA} and D_{NYSE} for Daimler and Deutsche Bank and the corresponding ratio $\tilde{D}$. The latter indicates that XETRA is more clustered compared to the NYSE (*F*-test, 1% level of significance) and therefore the NYSE is the more efficient stock market according to this analysis. In addition, the numerical values of D_{XETRA} and D_{NYSE} imply that the last digits of the stock prices of Daimler and Deutsche Bank are not uniformly distributed (Chi-squared goodness of fit test, 1% level of significance).

	D_{XETRA}	D_{NYSE}	$\tilde{D}$
Daimler	1243*	38.82*	32.02*
Deutsche Bank	4314*	119.18*	36.2*

Table 1: Chi-squared test statistics and *F* test statistics,
* denotes significance at the 1% level.

Considering the no-arbitrage condition of financial markets, stock prices at the NYSE should be obtained by multiplying the stock prices observed at XETRA with the corresponding exchange rate (that is valid for the observed time point) and maybe rounding these converted prices to the next possible Dollar-price (and vice versa).To be more precise, we do not expect any difference in the clustering structure between actually observed Dollar-prices and Dollar-prices that result from converting the Euro prices (and vice versa). The resulting frequency distributions of the last digits of those converted stock prices and of the last digits of the actually observed stock prices are presented in Figure 2 for the German company Deutsche Bank[1]. Analyzing the last digits of the transactions data of Deutsche Bank, we observe substantial differences between the clustering of converted and actually observed stock prices. In addition, the converted Euro-prices and the converted Dollar-prices seem to be uniformly distributed. We used a Chi-squared 'goodness of fit' test to check whether the observed distribution of the last digits differs from the expected distribution (that results from converting the stock prices). The numerical values of the test statistic are presented in Table 2, indicating statistical significance (at the 1% level) that the actually observed distribution of the last digits differs from the distribution we would expect after converting stock prices.

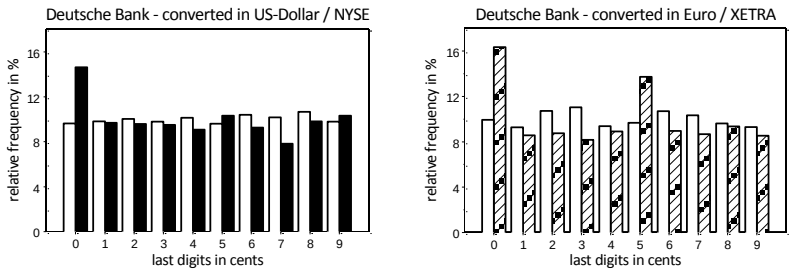


Fig. 2. Frequency distributions of the last digits of the stock prices of converted Deutsche Bank-prices (solid white bars) and of actually observed stock prices of Deutsche Bank at the NYSE (solid black bar) and at XETRA (dashed bar).

[1] Analyzing the stock prices of Daimler we obtain similar results. The corresponding plots can be made available by the corresponding author on request.

	Converted in US-Dollar vs.actually observed US-Dollar *D*	Converted in Euro vs. actually observed Euro *D*
Daimler	67.12	191.96
Deutsche Bank	149.42	4891.10

Table 2: Chi-squared test statistics, * denotes significance at the 1% level

Iit is obvious that the actually observed Dollar-prices at the NYSE reveal a clustering pattern, while there is only a low degree of stock price clustering for the last digits of into US-Dollar converted Euro-prices (Figure 2). Furthermore, the latter seem to be uniformly distributed. But the degree of stock price clustering of Dollar-prices that result from converting the Euro-prices at XETRA corresponds to the degree of stock price clustering we observe at XETRA for these are the same prices, only in different currencies. This leads to the result that the NYSE reveals an additional stock price clustering and therefore we can conclude that the NYSE has a higher degree of stock price clustering compared to XETRA and also that XETRA is a more efficient stock market. That means, comparing converted XETRA-prices and actually observed NYSE-prices yields that XETRA is more efficient.

Summarizing our empirical findings concerning the phenomenon of stock price clustering, we observe different extents of price clustering for the same stocks traded simultaneously on two stock markets. This implies inefficiency between both analyzed stock markets XETRA and the NYSE (for the stock prices of Daimler and Deutsche Bank). But the different approaches how to compare the degree of stock price clustering (with and without using the exchange rate) yield that we cannot strictly respond to our question whether XETRA or the NYSE is the more efficient stock market. If the exchange rate is taken into account we obtain the result that XETRA is more efficient when comparing converted Euro-prices and actually observed Dollar-prices. Nevertheless, we do not expect this observed inefficiency between both stock markets if the no-arbitrage condition of financial markets held.

In the following, we want to provide a trading strategy how to benefit from this observed inefficiency and we calculate a proxy of possible profits. For the purpose of investigating possible arbitrage opportunities (or quasi-arbitrage opportunities to be more precise) it is necessary to know for example the bid price for a specific stock at XETRA, the ask price at NYSE and the corresponding exchange rate at a point in time. As these prices are in most cases not available, we use transaction prices as proxies for this procedure. Analyzing those trades provides a strong indication about the existence or non-existence of (quasi-)arbitrage opportunities. That means we are noting a stock price at XETRA at one point in time and we are converting this Euro-price into a Dollar-price by using the exchange rate that is valid at the time (and vice versa). In a next step we compare the difference between this converted price and the next possible transaction that occurs at the NYSE (and vice versa). If the no-arbitrage condition is fulfilled this difference is zero. Our data and analysis provide empirical evidence that the differences are not zero in most

cases. Table 3 presents the proportions of zero differences and non-zero differences between converted and actually observed stock prices for Daimler and Deutsche Bank. The proportion of non-zero differences exceeds 80% (the results are significant at the 1% level, binomial test).

	Converted in US-Dollar vs. actually observed US-Dollar		Converted in Euro vs. actually observed Euro	
	Difference<>0	Difference=0	Difference<>0	Difference=0
Daimler	83.77%	16.23%	81.15%	18.85%
Deutsche Bank	92.11%	7.89%	91.10%	8.90%

Table 3: Proportions of zero differences between Dollar-prices that result from converting the Euro-prices and actually observed Dollar-prices (and between Euro-prices that result from converting the Dollar-prices and actually observed Euro-prices, respectively).

As the proportion of non-zero differences clearly indicates possible (quasi-) arbitrage opportunities, we want to provide a trading strategy that takes advantage of the observed inefficiency between the two analyzed stock markets and of the observed non-zero differences.
As a signal to buy or sell shares (in this context, a sale also can be a short sale) we consider the most recent observable difference (between converted XETRA-prices and actually observed NYSE-prices or converted NYSE-prices and actually observed XETRA-prices, respectively). If this difference exceeds 0.05 US-Dollar or Euro, we buy the stock on the observed market and sell the same on the other market. We propose a short sale of shares on the main market and a buy (or buy to cover) on the other market, if the observed difference is less than -0.05 US-Dollar or Euro. The following Table 4 presents the average profits of this strategy for the period of investigation and one traded share.

	Converted in US-Dollar vs. actually observed US-Dollar		Converted in Euro vs. actually observed Euro	
	Mean profit per one share per trade in US-Dollar	Number of trades	Mean profit per one share per trade in Euro	Number of trades
Daimler	0.01295	227	0.01917	73
Deutsche Bank	0.03432	1172	0.03352	565

Table 4: Mean profit and number of trades for the suggested trading strategy.

It can be argued that the proposed strategy yields not enough profit to achieve a considerable net profit when taking transaction costs into account. But considering the fee structure of the U.S. broker TradeStation Securities, Inc., a complete transaction can for example be traded for less than 0.00699 US-Dollar per share by

using a so called flat fee[2]. For this case, the calculated average profit per trade seems to be quite lucrative. We can conclude that the suggested trading strategy yields a positive profit when considering transaction costs.

3 Simulation Approach

As the results in Section 2 apply to empirical transactions data of selected German stocks we want to verify, whether we obtain similar results by simulating financial markets. The observed extent of stock price clustering in the empirical data already puts some question on the Efficient Market Hypothesis and the no-arbitrage condition as well. In this Section, we want to check more formally, whether a well known stochastic process emphasizes or contradicts our empirical findings (Fama, 1969). For this purpose, we use the Random Walk as a model for the price movement (which is in line with the Efficient Market Hypothesis) for both stock markets XETRA and the NYSE (Cootner, 1962). According to the Random Walk Hypothesis differences of successive stock prices are normally distributed with expectation zero and variance σ^2[3]. The following Figure 3 presents the frequency distributions of the price differences of the stock prices of Deutsche Bank[4] (for XETRA and the NYSE), and Figure 4 presents the frequency distribution of the price differences of the simulated stock prices[5]. According to these plots, the assumption of a Random Walk as a model for the price movement seems to be suitable and in the next Subsection, we want to have a detailed look at the last digits of simulated stock prices.

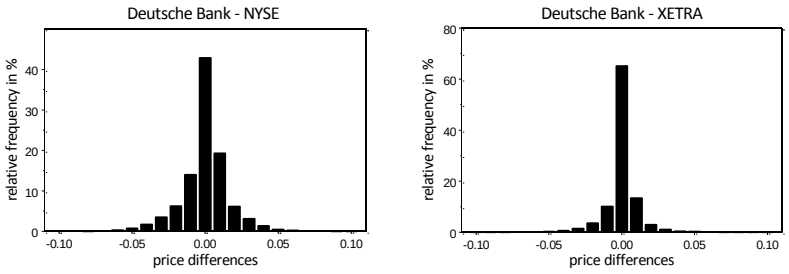


Fig. 3. Frequency distributions of the price differences of the stock prices of Deutsche Bank.

[2] TradeStation Securities, Inc Flat fee 6.99$ per trade max 5000 shares per trade, minimum 30 Trades per month on account. The example is calculated with 2000 shares per trade.
[3] Our data reveal an expected value of zero for the price differences. For the purpose of simulation, the variance has to be estimated using the mean squares error.
[4] The stock prices of Daimler reveal a similar pattern and all Figures can be shown on request by the corresponding author.
[5] The frequency distributions of price differences of converted stock prices and of converted simulated stock prices reveal the same pattern.

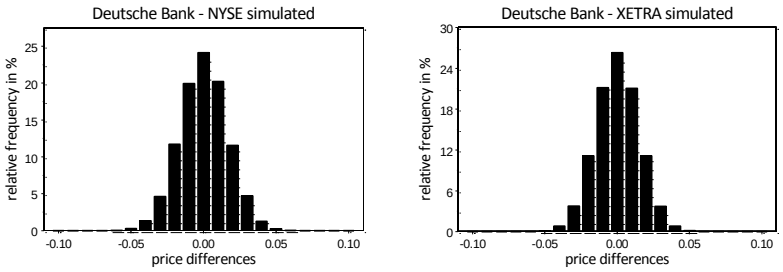


Fig. 4. Frequency distributions of the price differences of the simulated stock prices of Deutsche Bank.

3.1 Stock Price Clustering of Simulated Data

As in Subsection 2.2, the last digits of the simulated stock market prices of Deutsche Bank imply the frequency distributions presented in Figures 5. The stock prices of Daimler reveal the same results. The usage of converted simulated stock prices or converted stock prices shows the same results concerning the stock price clustering phenomenon.

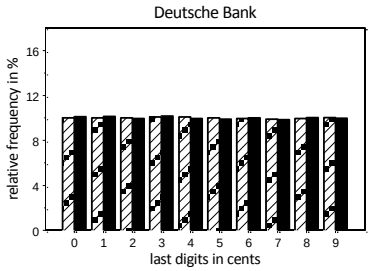


Fig. 5. Frequency distributions of the last digits of simulated stock prices at XETRA (dashed bar) and the NYSE (solid black bar).

Figure 5 implies a uniform distribution of the last digits of simulated stock prices (the same argument holds for the last digits of converted stock prices and converted simulated stock prices) and the degree of stock price clustering does not differ between the two exchanges (Chi-squared goodness of fit test, 1% level of significance). This is a result that contradicts our empirical findings. Although the price differences seem to be normally distributed, the stock price clustering reveals substantial differences between empirical and simulated data. We would expect the results of the simulated data for our empirical data if the Efficient Market

Hypothesis, and the no-arbitrage condition of financial markets as well, held. Therefore, our analysis reinforce the efficiency differences between XETRA and the NYSE and that trading German stocks simultaneously at both stock markets reveals (quasi-)arbitrage opportunities from the Behavioral Finance point of view.

4 Conclusion

This paper investigates the question which stock market is more efficient: XETRA or the NYSE? We examine this question by analyzing high frequency data of selected German stocks (Daimler and Deutsche Bank) that are traded simultaneously at both stock markets. If we take the exchange rate into account we show that German stocks are traded more efficiently at the German stock exchange XETRA. This result also hints at arbitrage possibilities. Furthermore, we analyzed quasi-arbitrage opportunities by suggesting a trading strategy. We have shown that simultaneous trading on both stock markets leads to lucrative profits after subtracting transaction costs. Even if this can only serve as a proxy, it provides a clear indication of arbitrage. As these results apply to empirical intraday data, we want to check whether we obtain similar results by simulating the stock markets. The simulation results reveal that the stock price clustering and the differences between the two exchanges cannot be explained by two Random Walks or two processes which are the same on both stock exchanges.

References

1. Ball, C.A., Torous, W.N., Tschoegl, A.E., The degree of price resolution: The case of the gold market. *The Journal of Futures Markets*, **5**, 29-43 (1985).
2. Christie, W.G., Harris, J.H., Schultz, P.H., Why did NASDAQ market makers stop avoiding odd-eighth quotes? *The Journal of Finance*, **49**, 1841-1860 (1994).
3. Cootner, Paul H., Stock Prices: Random vs. Systematic Changes. *Industrial Management Review*, **3**, 24-45 (1962).
4. Fama, Eugene F., Random Walks in Stock Market Prices. *Financial Analysts Journal*, **21**, 5, 55-59 (1965).
5. Fama, Eugene F., Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, **25**, 2, 383-417 (1969).
6. Grossman, Sanford J., On the Efficiency of Competitive Stock Markets Where Traders Have Diverse Information. *The Journal of Finance*, **31**, 573-585 (1976).
7. Harris, L., Stock price clustering and discreteness. *Review of Financial Studies*, **4**, 389-415 (1991).
8. Huang, R.D., Stoll, H.R., Tick size, bid-ask spreads, and market structure. *The Journal of Financial and Quantitative Analysis*, **36**, 503-522 (2001).

9. Ikenberry, D.L., Weston, J.P., Clustering in US stock prices after decimalization. *European Financial Management*, **14**, 30-54 (2008).

10. Kahn, C., Pennachi, G., Sopranzetti, B., Bank deposit rate clustering: Theory and empirical evidence. *Journal of Finance*, **54**, 2185-2214 (1999).

11. Niederhoffer, V., Clustering of stock prices. *Operations Research*, **13**, 258-265 (1965).

12. Niederhoffer, V., A new look at clustering of stock prices. *Journal of Business*, **39**, 309-313 (1966).

13. Osborne, M.F.M., Periodic structure in the Brownian motion of stock prices. *Operations Research*, **10**, 345-379 (1962).

14. Sonnemans, J., Price clustering and natural resistance points in the Dutch stock market: A natural experiment. *European Economic Review*, **50**, 1937-1950 (2006).

15. Sopranzetti, B.J., Datar, V., Price Clustering in foreign exchange spot markets. *Journal of Financial Markets*, **5**, 411-417 (2002).

16. Vogt, B., Uphaus, A., Albers, W., Numerical decision processing causing stock price clustering? *Homo Oeconomicus*, **18**, 1-12 (2001).

SMTDA 2010: Stochastic Modeling Techniques and Data Analysis
International Conference, *Chania, Crete, Greece, 8 - 11 June 2010*

Generalizability Analysis: An Example Using Unbalanced Data

Teresa Rivas and Rosa Bersabé

Faculty of Psychology, University of Málaga. Spain
Email: moya@uma.es
Email: bersabe@uma.es

Abstract: The Generalizability Theory (GT) was developed by Cronbach et al. (1972). Only a small number of applications - in educational testing, marketing, etc. - have appeared (e.g., Brennan et al., 1995; Bruckner, et al., 2006; Finn, 2004). GT Analysis was restricted to balanced and non-missing data. Brennan (2001a) has provided extensions of GT models which are better adapted to real situation measurements in psychological and educational practice (unbalanced, missing data, etc). Software is also now available (Brennan, 2001b).
This paper shows an application of a GT model for unbalanced data, with a design $p^{\bullet} \times i^{\circ}$, being p (subjects) crossed with the fixed facet (scales) and i (items) nested in each fixed facet. The data are answers given by 778 females to the questionnaire, Eating Attitudes Test (EAT-26; Garner, Olmsted, Bohr & Garfinkel, 1982). Generalizability Coefficients obtained in Generalizability (G) and Decision (D) Studies are described and interpreted. From these results a composite score across scales is also given.
Keywords: Generalizability, Unbalanced design, Composite Score

1. Introduction

Within the frame of Test Theory, the Generalizability Theory (GT) provides models which analyse the error from which an observed score of a subject can be generalised to its universe score.
For each of the designs which can be planned to collect data, there is a GT model. Each GT model provides *Generalizability coefficients* for absolute or relative (to a reference group) decisions/measures. Both types of coefficients can assume values between 0 and 1. A value close to 1 shows that the essential source of variation to account for the observed scores is the variance of universe scores of the measured object (generally subjects). A value close to 0 indicates that other important sources of variation, due to particular conditions of measure (items, occasions, situations, etc.), are also present (Martinez-Arias, 1995, p. 219). Brennan (2001a) and Shavelson & Webb (1991) present the characteristics of the different GT models for balanced data, and completely crossed or nested designs. Balanced data are analysed using ANOVA statistic models.
Brennan (2001a) also presents some extensions of GT models. In particular, he shows several unbalanced G study designs and the statistical procedure *Analogous-ANOVA* for the estimation of G study variance components. He also points out that general procedures applicable to any unbalanced D study design are unknown. However, he shows estimators of error variances and coefficients for several frequently encountered unbalanced D study designs (Brennan, 2001a, pp: 215-

216). This author solves the problem of unbalanced mixed effects designs. In practice balanced designs, or unbalanced designs with respect to nesting, can be fitted by free computer programs provided by Brennan (2001b). Recently, Cardinet, Johnson & Pini (2010) have also provided the EduG program.
From real data and an unbalanced mixed effects design, this paper presents the results and interpretation of G and D study coefficients obtained with the mGENOVA program.

2. Method

Instrument

The *Eating Attitudes Test* (EAT-26; Garner et al., 1982) is a 26-item self-report questionnaire. Items are presented in a 6-point forced-choice Likert scale ranging from 1 (never) to 6 (always). The total score is obtained by recoding scores as follows: scores from 1 to 3 are recoded 0, 4 is recoded 1, 5 as 2, and 6 is recoded 3. The only exception is item 25 whose answers score as follows: 1 as 3, 2 as 2, 3 as 1, and 4 to 6 as 0. The EAT-26 total score ranges from 0 to 78.
Garner et al., (1982) determined three factors or scales, that were labelled Dieting (D), Bulimia and Food Preoccupation (B) and Oral Control (OC). Item numbers in each scale being 13, 6 and 7, respectively.

Data

778 women (aged 12-21) answered all the items of the questionnaire EAT-26. They were randomly sampled from different High Schools in Malaga Province (Spain).

Unbalanced $p^{\bullet} \times i^{\circ}$ Design. Composite model

A mixed model and an unbalanced design have been considered. The term $(p^{\bullet})$ indicates that p (participant) is crossed with the fixed facet (scales) and the term (i°) denotes that i is nested in a fixed facet (following the notation in Brennan, 2001b, p.38). Facet i (items) is nested in the levels of fixed facet j (scales), being p (participants: 778), i (items: 13,6,7), j (scales: 3) . The three levels in the scales are D, B and OC.
Scores in the Composite model are given as:

$$X_C = w_D\left(0D_{i0} + 1D_{i1} + 2D_{i2} + 3D_{i3}\right) + w_B\left(0B_{i0} + 1B_{i1} + 2B_{i2} + 3B_{i3}\right) +$$
$$+ w_{OC}\left(0OC_{i0} + 1OC_{i1} + 2OC_{i2} + 3OC_{i3}\right)$$

being

D_{ik} item *i* of Dieting scale scored in category k $\quad i:1,...,13 \quad k:0,..,3$

B_{ik} item *i* of Bulimia scale scored in category k $\quad i:1,...,6 \quad k:0,..,3$

OC_{ik} item *i* of Oral Control scale scored in category k $\quad i:1,...,7 \quad k:0,..,3$

w_D, w_B, w_{OC} weights for each scale are defined as the ratio of the number of scale items to the total number of EAT-26 items. In the first D Study, these weights are $w_D = 13/26$, $w_B = 6/26$, $w_{OC} = 7/26$.

X_c gives a participant's score in EAT-26.

A participant scoring the maximum value (3) in all the items obtains a maximum score of 3 in EAT-26 from this Composite model. A participant scoring the minimum value (0) in all the items obtains a minimum score of 0 in EAT-26.

Results

The fitting of GT model is carried out using the mGENOVA program (Brennan, 2001b). Results of Generalizability Study (G Study) and four Decision Studies (D Study) are shown in Table 1.

G Study

(a) Variance components for each scale (Columns 4 - 6).

D Study (1) is obtained from levels of scales used in G Study.

Other D Studies (2 - 4) are obtained by successively increasing the number of items in the different scales:

(b) Number of scale items in each Study (Column 2) and weights for the composite model

(c) Error variance for relative and absolute measures in different scales (Columns 7- 8)

(d) Generalizability coefficients for relative and absolute measures in different scales (Columns 9 - 10)

(e) Contributions to variances $\hat{\sigma}^2{}_p\%, \hat{\sigma}^2_\delta\%$, $\hat{\sigma}^2_\Delta\%$ in the composite model (Columns 11- 13).

Taking the D Study (1) as reference (with the real number of items in the scales), in the successive D studies (2) – (4), the greater the number of items, the higher the value of the G coefficient. In each study, there is hardly any difference between the G coefficients for absolute and relative measures. Figure 1 shows a summary of Generalizability coefficients - for relative and absolute measures - obtained in the different scales used in the four D studies.

In the fourth D Study, if the number of items in the Bulimia and Oral Control scales is increased to 13, the G coefficients will be $\hat{\rho}^2_D = 0.849$; $\hat{\rho}^2_B = 0.793$; $\hat{\rho}^2_{OC} = 0.712$ for relative measures, and $\hat{\Phi}_D = 0.833$; $\hat{\Phi}_B = 0.782$; $\hat{\Phi}_{OC} = 0.707$ for absolute measures.

The composite model gives G coefficients close to 0.9, for both relative and absolute measures. These coefficients increase very slowly across the different D studies (Col 9-10, Table 1, Figure 1). However, the contributions to $\hat{\sigma}^2_p$, $\hat{\sigma}^2_\delta$, $\hat{\sigma}^2_\Delta$ vary meaningfully in each scale across the different D studies. Taking each previous study D as reference, the contributions to variances (a) decrease in each scale in which the number of items is fixed, and (b) increase in the scales in which the number of items is increased.

Table 1. Design $p^{\bullet} \times i^{\circ} : p$ crossed with the fixed facet j, and i nested in j; G and D Studies

	n_i	w_i	$\hat{\sigma}^2(p)$	$\hat{\sigma}^2(i)$	$\hat{\sigma}^2(pi)$	$\hat{\sigma}^2_{\delta}$	$\hat{\sigma}^2_{\Delta}$	$\hat{\rho}^2$	$\hat{\Phi}$	Contributions to $\hat{\sigma}^2{}_p\%$	$\hat{\sigma}^2_{\delta}\%$	$\hat{\sigma}^2_{\Delta}\%$
G Study												
Dieting			0.276	0.080	0.638							
Bulimia			0.106	0.025	0.361							
Oral Control			0.117	0.016	0.615							
D Study (1)												
Separate												
Dieting	13		0.276			0.049	0.055	0.849	0.833			
Bulimia	6		0.106			0.060	0.064	0.638	0.623			
Oral Control	7		0.117			0.088	0.090	0.572	0.565			
Composite			0.150			0.022	0.024	0.873	0.863			
Dieting	13	13/26								65.42	56.19	58.11
Bulimia	6	6/26								17.75	14.68	14.41
Oral Control	7	7/26								16.84	29.13	27.48
D Study (2)												
Separate												
Dieting	13		0.276			0.049	0.055	0.849	0.833			
Bulimia	**10**		0.106			0.036	0.039	**0.746**	**0.734**			
Oral Control	**10**		0.117			0.061	0.063	**0.656**	**0.650**			
Composite			0.133			0.017	0.018	**0.889**	**0.881**			
Dieting	13	13/26								53.65	45.95	47.87
Bulimia	**10**	10/26								25.13	20.01	19.79
Oral Control	**10**	10/26								21.22	34.04	32.34

Contd.

Table 1(Continuation). Design $p^{\bullet} \times i^{\circ} p$ crossed with the fixed facet j, and i nested in j; G and D Studies

	n_i	w_i	$\hat{\sigma}^2_p$	$\hat{\sigma}^2_i$	$\hat{\sigma}^2_{pi}$	$\hat{\sigma}^2_\delta$	$\hat{\sigma}^2_\Delta$	$\hat{\rho}^2$	$\hat{\Phi}$	Contributions to $\hat{\sigma}^2_p$%	Contributions to $\hat{\sigma}^2_\delta$%	Contributions to $\hat{\sigma}^2_\Delta$%
D Study (3)												
Separate												
Dieting	13		0.276			0.049	0.055	0.849	0.833			
Bulimia	**13**		0.106			0.028	0.030	**0.793**	**0.782**			
Oral Control	10		0.117			0.061	0.063	0.656	0.650			
Composite			0.129			0.015	0.016	**0.897**	**0.890**			
Dieting	13	13/26								49.73	43.35	45.19
Bulimia	**13**	13/26								30.70	24.54	24.28
Oral Control	10	10/26								19.57	32.11	30.53
D Study (4)												
Separate												
Dieting	13		0.276			0.049	0.055	0.849	0.833			
Bulimia	13		0.106			0.028	0.030	0.793	0.782			
Oral Control	**13**		0.117			0.047	0.049	**0.712**	**0.707**			
Composite			0.124			0.014	0.015	**0.900**	**0.893**			
Dieting	13	13/26								46.19	39.54	41.40
Bulimia	13	13/26								28,75	22.38	22.24
Oral Control	**13**	13/26								25.06	38.07	36.36

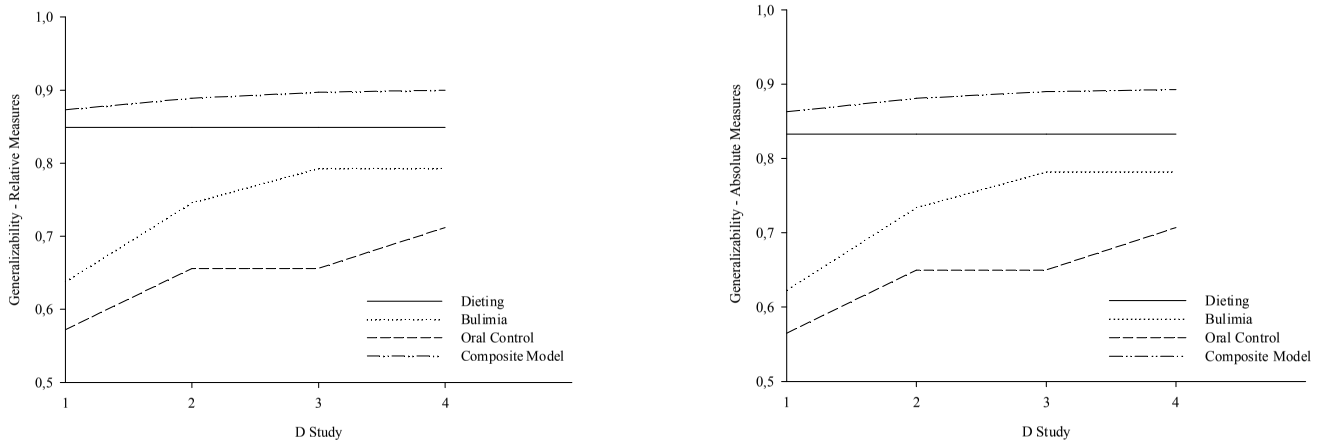


Figure 1. Generalizability for Relative and Absolute Measures in the D Studies

In the fourth D Study, the composite model for the three scales is given as $X_c = (13/26)D + (13/26)B + (13/26)OC$.

When the X_c total score is considered: a high degree of generalizability $\hat{\rho}^2 = 0.900$ is obtained for relative measures, and $\hat{\Phi} = 0.893$ for absolute measures.

3. Discussion

The GT model with this design has been used to show an application with unbalanced data, which can be very useful to validate questionnaires in Clinical Psychology. This design is also of special interest in the validation of Criterion Referenced Tests, e.g. in educational measurement, (Rivas, González & Delgado, in press).

Traditionally, few applications using balanced or unbalanced designs have appeared in psychometric literature.

In GT, the balanced designs application to assign scores to different objects - in psychological and educational measurement - has theoretical limitations such as (a) samples must be randomly drawn from the population, and (b) the assumptions of linearity, homoscedasticity, and normality in ANOVA models. In addition, on occasion in practice, the balanced designs can be very restrictive. This restriction diminishes using an unbalanced design (e.g. the unbalanced design in regard to nesting, as in the above example).

From the ANOVA statistical model it is relatively easy to define variance component estimations for any balanced G or D study. However, it is not easy to generalise how to obtain variance component estimations for any unbalanced design. Brennan (2001a) presents in detail some estimators for unbalanced G and D studies. This author also discusses different statistical questions arising from variance component estimation for unbalanced designs.

In practice, for some unbalanced designs it is possible that there may be no statistical procedure for the estimation of variance components. If there were, there could be more than one solution, thus complicating the choice of the solution. Therefore, the design and appropriate statistical analysis procedure require careful planning. Also, estimators of variance components and coefficients of G or D studies require careful interpretation.

Acknowledgments: This research was supported by grants from the Ministerio de Ciencia y Tecnología. (Project Ref. BSO2001-1945) and Consejería de Educación y Ciencia de la Junta de Andalucía (Research Group CTS-278)

References

1. Brennan, R.L. (2001a). *Generalizability Theory*. Springer: New York.
2. Brennan, R.L. (2001b). *Manual for mGENOVA*. Iowa Testing Programs Occasional Papers No. 50.
3. Brennan, R.L., Gao, X. & Colton, D.A. (1995). Generalizability analyses of work keys listening and reading. *Educational and Psychological Measurement 55*, 157–176.
4. Bruckner, C. T.,Yoder, P.J. & McWilliam, R.A. (2006). Generalizability and Decision Studies: An Example Using Conversational Language Samples. *Journal of Early Intervention 28*, 139-153
5. Cardinet, J., Johnson, S. & Pini, G. (2010). *Applying Generalizability Theory using EduG*. Routledge: New York
6. Cronbach, L.J., Gleser, G.C., Nanda, H. & Rajaratnam, N. (1972). *The Dependability of Behavioural Measurements: Theory of Generalizability for Scores and Profiles*. Wiley: New York.
7. Finn, A. (2004). A reassessment of the dimensionality of retail performance: a multivariate generalizability theory perspective. *Journal of Retailing and Consumer Services, 11,* 235-245
8. Garner, D.M., Olmsted, M.P., Bohr, Y., & Garfinkel, P.E. (1982). The Eating Attitudes Test: psychometric features and clinical correlates. *Psychological Medicine, 12*, 871-878.
9. Martínez-Arias, M.R. (2005). *Psicometría. Teoría de los Tests Psicológicos y Educativos*. Síntesis: Madrid
10. Rivas, T., González, M.J. & Delgado, M. (in press). Descripción y Propiedades Psicométricas del Test de Evaluación del Rendimiento Académico (TERA). *Revista Interamericana de Psicología*
11. Searle, S., Casella, G., & McCulloch, C.E. (1992). *Variance Components*. Wiley: New York.

SMTDA 2010: Stochastic Modeling Techniques and Data Analysis International Conference, *Chania, Crete, Greece, 8 - 11 June 2010*

Procedure to Calculate Efficiency Assessment of Three-Category Classification

Teresa Rivas and Félix Caballero

Faculty of Psychology, University of Malaga. Spain

Email: moya@uma.es

Email: ffcaballero@uma.es

Abstract: An efficiency index is defined to show the utility of a classifier when two cut-off points and three categories are considered.

Given two cut-off points, a scoring classifier $\vec{x}$ is determined by the point (x_1, x_2, x_3, x_4, x_5, x_6) which represents the possible six types of errors. A graphical representation of $\vec{x}$ in 3D is given as (F*a*_R, F*b*_R, F*c*_R), F*i*_R being the proportion of cases whose real state is *i* (*i* = *a, b, c*) although they have been classified into categories other than *i*.

Given any classifier $\vec{x}$ in the unit cube, with $\vec{x}$ lying on the plane $Fa\_R + Fb\_R + Fc\_R = 2$, the Index based on the Tetrahedron Volume (ITV) is defined as

$$ITV = 1 - \frac{Fa\_R + Fb\_R + Fc\_R}{2} \quad \text{being} \quad 0 \le Fa\_R + Fb\_R + Fc\_R \le 2$$

The procedure to obtain ITV is shown and its properties are described. An example is also given.

Keywords: Measurement, Multiple Cut-offs, Three-Category Classification**,** Efficiency Assessment

1. Introduction

Traditional ROC Analysis (Receiver Operating Curve) assumes that diagnostic tests distinguish between two mutually exclusive classes, generally, with or without disorder. In many situations of measurement, a more precise classification could be obtained by considering three diagnostic categories (e.g. asymptomatic, symptomatic and with disorder). Considering 3 diagnostic categories (instead of 2) produces notable differences in a 3 x 3 classification table in relation to a 2x2 table. This is because, in a 3x3 table, the errors cannot be defined in the same way as negative and positive falses given in 2x2 tables (See Table 1).

ROC indices that summarise the results of classification are necessary when there are three diagnostic categories, as also in the case of two categories.

Some authors define different goodness of fit measures to assess the precision of a trichotomic classification (e.g. Ferri, Hernández Orallo & Salido, 2003a, 2003b). For such generalizations of ROC analysis to the multidimensional case, different authors have defined precision and utility measures of classifications. Hand & Till (2001) present a generalization of AUC (*Area Under the ROC Curve*) measure for classifiers who assign a different score or probability to each prediction. They define an *M* measure from the AUC index associated with each pair of categories. Ferri, Hernández-Orallo & Salido (2003) present an extension of AUC in the form of VUS (Volume Under the ROC Surface). They also give a set of measures or indices of efficiency (performance) for a classifier in three categories (e.g. the extension AUC for a point, AUC-1PT3, and different variants of the Hand & Till measure). Based on Ferri, Hernández-Orallo & Salido (2003), Caballero & Rivas (2009) consider the problem of convex polygon – formed in a two-class ROC analysis - and extend it to the problem of determining the convex surface which is obtained when a three-category analysis is considered. From these concepts, they define the Index based on the Tetrahedron Volume (ITV) :

Given any classifier $\vec{\boldsymbol{x}} = \left(\boldsymbol{Fa\_R}, \boldsymbol{Fb\_R}, \boldsymbol{Fc\_R}\right)$ from a unit cube,

$$ITV = 1 - \frac{Fa\_R + Fb\_R + Fc\_R}{2} \quad \text{being} \quad 0 \le Fa\_R + Fb\_R + Fc\_R \le 2$$

The possible values of ITV are found in the interval [0, 1] and high values indicate a better efficiency of classifier

From cut-off scores $\left(X = x\right)$ obtained with real data, Rivas & Caballero (2009) compare ITV with indices defined by other authors, such as *AUC-1PT3* and extensions (*HT1, HT3*) of Hand-Till's *M measure*.

From a 3x3 classification Table, this paper gives the procedure to calculate the ITV index. A 3x3 Table contains classifications established by a categorical dependent (actual or real categories) and a predictor variable (e.g. test) on which two cut-off scores (classifiers) have been established. These cut-offs give the predicted categories (classification intervals).

2. Procedure

In a trichotomic classification, let the three categories of study be *A*, *B*,*C*.

In accordance with a criterion (e.g. diagnosis), the sample which comprises the study is classified in one of three real states, labelled A_R, B_R , C_R.

In accordance with a predictor variable or instrument of measure (e.g. test), let [t_0, t_3] be the range of scores obtained on the test by subjects of the sample.

From some procedures (e.g., ROC curves, Multinomial Logistic Regression model, etc.), two cut-off scores $t_1, t_2 \in [t_1, t_3]$ are obtained on the total score of subjects. These cut-off scores establish the following classification intervals on total scores [t_0, t_1), [t_1, t_2), [t_2, t_3]. Sample subjects fall into one of these intervals or predicted categories, which can be labelled A_p, B_p, C_p, respectively.

These two cut-off scores also establish a classifier $\vec{x} = (x_1, x_2, x_3, x_4, x_5, x_6)$. The efficiency of this error classifier is assessed.

Table 1 shows the number of subjects classified according to the real state (diagnostic criterion) and the classification established by the instrument of measure used. Table 2 gives the proportions associated with the frequencies given in Table 1.

Table 1. Absolute frequencies associated with three categories

		Real Category			
		A_R	B_R	C_R	Total
Predicted Category	A_p	$n_{A_R A_p}$	$n_{B_R A_p}$	$n_{C_R A_p}$	n_{A_p}
	B_p	$n_{A_R B_p}$	$n_{B_R B_p}$	$n_{C_R B_p}$	n_{B_p}
	C_p	$n_{A_R C_p}$	$n_{B_R C_p}$	$n_{C_R C_p}$	n_{C_p}
	Total	n_{A_R}	n_{B_R}	n_{C_R}	

Table 2. Proportions associated with three categories

		Real Category			
		A_R	B_R	C_R	Total
Predicted Category	A_p	$\frac{n_{A_R A_p}}{n_{A_R}}$	x_1	x_2	n_{A_p}
	B_p	x_3	$\frac{n_{B_R B_p}}{n_{B_R}}$	x_4	n_{B_p}

C_p	x_5	x_6	$\frac{n_{C_R C_p}}{n_{C_R}}$	n_{C_p}
Total	1	1	1	

To facilitate notation, in Table 2, x_i, *i = 1, 2, ..., 6,* is noted as the proportion of subjects predicted wrongly in each real category.

$$x_1 = \frac{n_{B_R A_p}}{n_{B_R}} \qquad x_2 = \frac{n_{C_R A_p}}{n_{C_R}} \qquad x_3 = \frac{n_{A_R B_p}}{n_{A_R}}$$

$$x_4 = \frac{n_{C_R B_p}}{n_{C_R}} \qquad x_5 = \frac{n_{A_R C_p}}{n_{A_R}} \qquad x_6 = \frac{n_{B_R C_p}}{n_{B_R}}$$

A graphical representation of classifier $\vec{x} = (x_1, x_2, x_3, x_4, x_5, x_6)$ could be made in a 6D space, but this has graphical limitations. A graphical representation in 3D, where the classifier $\vec{x}$ is determined by the point, is proposed.

$$(Fa\_R, Fb\_R, Fc\_R) = (x_3 + x_5, x_1 + x_6, x_2 + x_4),$$

Fi_R being the proportion of subjects whose real state is *i* but, according to the established cut off points, they have been classified into a different category to *i* (*i* = *a*, *b*, *c*).

Ignoring all decision rules, the trivial classifiers for three categories are those that classify all subjects into category *a*, *b* or *c*. These three trivial classifiers, in 6D notation, are represented by (1, 1, 0, 0, 0, 0), (0, 0, 1, 1, 0, 0), (0, 0, 0, 0, 1, 1) and in 3D notation by (0, 1, 1), (1, 0, 1), (1, 1, 0), respectively. The general equation of plane, in which the trivial classifiers lie, is $Fa\_R + Fb\_R + Fc\_R = 2$.

Given that any of these classifiers would be less efficient than a trivial classifier, those lying above this plane can, therefore, be discarded (Caballero & Rivas, 2009).

Let $\vec{x} = (Fa\_R, Fb\_R, Fc\_R)$ be a classifier given by a point that satisfies the condition $0 \le Fa\_R + Fb\_R + Fc\_R \le 2$. The tetrahedron whose vertices are this point and the three trivial classifiers are considered. If the tetrahedron volume is multiplied by 3, an index which assumes values between 0 and 1 is defined. This Index based on Tetrahedron Volume, labelled ITV, is defined as

$$ITV = 1 - \frac{Fa\_R + Fb\_R + Fc\_R}{2}$$

ITV is defined exclusively for $\vec{x} = (Fa\_R, Fb\_R, Fc\_R)$, such that $0 \le Fa\_R + Fb\_R + Fc\_R \le 2$.

The tetrahedron whose vertices are the point (0, 0, 0) – or ideal classifier - and the points that represent the trivial classifiers, in 3D notation, is the tetrahedron of

greater volume (Figure 1). Thus, this tetrahedron is that which presents a greater ITV. In particular, ITV is equal to 1 for the ideal classifier, and ITV is equal to 0 for trivial classifiers.

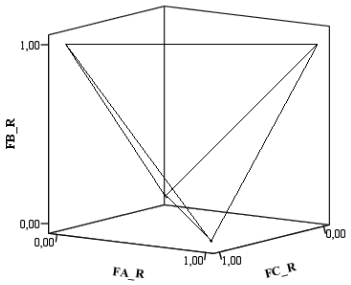


Figure 1. Tetrahedron defined by ideal and trivial classifiers

ITV is not defined for classifiers lying above the plane of trivial classifiers. The classifiers lying above this plane (in 3D notation, (Fa_R, Fb_R, Fc_R) satisfying $Fa\_R + Fb\_R + Fc\_R \geq 2$) are less efficient than the classifiers (Fa_R, Fb_R, Fc_R) lying on the same plane as trivial classifiers. (They are also less efficient than the classifiers lying below this plane). This is because the sum of proportions of different error types is greater in the first than in the second case. The classifiers in the first case can be discarded, because they show less utility or efficiency than that of a trivial classifier. This trivial classifier is a random classifier that does not involve any decision rule or its associated cost.

3. Application

The Eating Attitudes Test (EAT–26; Garner, Olmstead, Bohr & Garfinkel, 1982) is a self-report questionnaire with 26 items. It is the abbreviated version of the EAT–40 (Garner & Garfinkel, 1979). EAT-26 has frequently been used in epidemiological studies to detect possible cases with an eating disorder (ED) in subject samples drawn from the community. Each item is scored on a Likert type scale from 1 (never) to 6 (always). EAT–26 total score ranges 0 – 78.

The Spanish version of the *Questionnaire for Eating Disorder Diagnoses* (Q-EDD; Mintz, O'Halloran, Mulholland, & Schneider, 1997) was developed by Rivas, Bersabé & Castro (2001). The Q-EDD is a self-report questionnaire with 50 items which operationalizes the DSM-IV diagnostic criteria for ED (American Psychiatric Association, 1994). Q – EDD establishes the classification of subjects into several diagnostic categories of ED.

This paper uses scores in the EAT-26 questionnaire and categories (asymptomatic, symptomatic and with ED) according to Q-EDD diagnostic criteria assessed in a subject sample (778 females).

Two cut-off scores or classifiers obtained in the EAT-26 scores have been used to predict the three categories of classification given by the Q-EDD.

Different pairs of cut-off scores have been obtained by two procedures:

(a) Multinomial Logistic Regression Model: Bersabé & Rivas (2010), obtain cut-off scores 20 and 56. These provide score intervals: 0 – 20 (Asymptomatic), 21 – 56 (Symptomatic) and 57 – 78 (ED). The classifier obtained is labelled C1 (Table 3).

Table 3. Classification for C1

			Real Category (Q-EDD)			
			ED	Symptomatic	Asymptomatic	Total
Predicted Category (EAT-26 score)	ED	[57,78]	**0**	3	0	3
	Symptomatic	[21,56]	18	**43**	22	83
	Asymptomatic	[0, 20]	18	171	**503**	692
		Total	36	217	525	778

These cut-off scores are associated with classifier $\left(\frac{3}{217},0,\frac{1}{2},\frac{22}{525},\frac{1}{2},\frac{171}{217}\right)$, whose ITV value is

$$ITV = 1 - \frac{Fa\_R + Fb\_R + Fc\_R}{2} = 0{,}078$$

(b) Two-category ROC analysis for pairs of groups: Rivas, Bersabé, Jiménez & Berrocal (in press) obtain a cut-off score of 18 on EAT-26 to differentiate between categories Asymptomatic – Symptomatic. For this, the sample proportion of symptomatic subjects has been considered. A cut-off score of 43 to separate the categories Symptomatic – ED has also been obtained. In this case, the sample proportion of subjects with ED has also been considered. These cut-offs provide the score intervals on EAT-26 scores: 0 – 18 (Asymptomatic), 19 – 43 (Symptomatic) and 44 – 78 (ED). The classifier is labelled C2 (shown in Table 4).

Table 4. Classification for C2

	Real Category (Q-EDD)			
	ED	Symptomatic	Asymptomtic	Total

Predicted Category (EAT-26 score)						
	ED	[44,78]	**9**	7	2	18
	Symptomatic	[19,43]	11	**49**	26	86
	Asymptomatic	[0, 18]	16	161	**497**	674
		Total	36	217	525	778

These cut-off scores, obtained by both two-category ROC Analyses, define the classifier $\left(\frac{7}{217},\frac{2}{525},\frac{11}{36},\frac{26}{525},\frac{16}{36},\frac{161}{217}\right)$, whose ITV value is

$$ITV = 1 - \frac{Fa\_R + Fb\_R + Fc\_R}{2} = 0{,}211$$

From the cut-off scores obtained in (a) and (b) above, ITV shows that the classification given in (a) is less efficient than that given in (b).

4. Discussion

In a Three-Category classification, measures are necessary to determine the utility of a particular classifier. To this end, ITV quantifies the efficiency of the classification obtained from a particular classifier. It is defined from the tetrahedron whose vertices are the points (3D) associated with the possible better classifier and the three random classifiers which have ignored any decision procedure. As it is unnecessary to consider classifiers less efficient than a trivial classifier, this ITV has been defined exclusively for classifiers located on the same plane or below the plane where the points associated with the trivial classifiers lie.

The range of possible ITV values is 0-1, a greater value being an index of a greater efficiency or performance of a particular classifier. The better classifier (called here the 'ideal classifier') is that which classifies all the subjects into the correct category. It is associated with an ITV value of 1. Each one of the trivial classifiers, and the classifiers located on the plane defined by them, has an ITV value of 0. The object of this index is to discriminate between different classifiers associated with a particular diagnostic test, and to establish which of the possible classifiers shows greater efficiency. In the above application, the classifier obtained by procedure (b) is more efficient than that obtained by procedure (a)

In conclusion, the errors $(x_1, x_2, x_3, x_4, x_5, x_6)$ in a 3x3 Table could be seen in a similar way to positive and negative false proportions in a 2x2 Table . While F*a*_R, F*b*_R and F*c*_R could be seen as measures of error associated with each category A, B, C, in a 3x3 Table, they cannot be seen as extensions of positive and negative false proportions associated with each category in a 2x2 Table.

Extensions of ITV can be made using different weights – in function of the error cost - for each type of error in a 3x3 Table. In addition, the measure of global

precision (AUC) in a two-category ROC analysis could be extended for three-category ROC analysis.

Acknowledgments: This research was supported by grants from the Ministerio de Ciencia y Tecnología. (Project Ref. BSO2001-1945) and Consejería de Educación y Ciencia de la Junta de Andalucía (Research Group CTS-278)

References

1. American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: Author.
2. Bersabé, R., & Rivas, T. (2010). A general equation to obtain multiple cut-off scores on a test from multinomial logistic regression. *The Spanish Journal of Psychology, 13*, 1, 487-495
3. Caballero, F. F. & Rivas, T. (2009, September). Indices de Eficacia en un clasificación tricotómica (I). [Efficiency indices for trichotomic classification (I)]. Paper presented at the XII Spanish Conference of Biometry, University of Cádiz, Spain.
4. Ferri, C., Hernández-Orallo, J., & Salido, M.A. (2003). Volume under the ROC Surface for Multi-class Problems. In N. Lavrac, D. Gamberger, L. Todorovski, & H. Blockeel (eds.), *Proceedings of the 14th European Conference on Machine Learning* (pp. 108–120). Cavtat-Dubrovnik, Croatia: Springer.
5. Ferri, C. Hernández-Orallo, J. & Salido, M.A. (2003) *Volume Under the ROC Surface for Multi-class Problems. Exact Computation and Evaluation of Approximation.* (Technical Report, DSIC, 1-40). Valencia: Universidad Politécnica, Retrieved May 20, 2009, from http://users.dsic.upv.es/grupos/elp/cferri/VUS.pdf
6. Garner, D.M., & Garfinkel, P.E. (1979). The Eating Attitudes Test: An index of the symptoms of anorexia nervosa. *Psychological Medicine*, *9*, 273-279.
7. Garner, D. M., Olmstead, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The Eating Attitudes Test: Psychometric features and clinical correlates. *Psychological Medicine*, *12*, 871-878.
8. Hand, D.J. & Till, R.J. (2001). A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems. *Machine Learning, 45*, 171–186
9. Mintz L.B., O'Halloran, M.S., Mulholland, A.M., & Schneider, P.A. (1997). Questionnaire for Eating Disorders Diagnoses: Reliability and validity of operationalizing DSM-IV criteria into a self-report format. *Journal of Consulting Psychology*, *44*, 2, 132.
10. Rivas, T., Bersabé, R., Jiménez, M., & Berrocal, C. (in press). The Eating Attitudes Test (EAT – 26): reliability and validity in Spanish Female Samples. *The Spanish Journal of Psychology*.

11. Rivas, T. & Caballero, F. F. (2009, September). *Indices de Eficacia en un clasificación tricotómica (II)* [Efficiency indices for trichotomic classification (II)]. Paper presented at the XII Spanish Conference of Biometry, University of Cádiz, Spain.

A new Local EM Estimation Method for Latent Factorial Generalized Linear Models

Mohamed Saidane[1], Christian Lavergne[1,2], and Xavier Bry[1]

[1] Université Montpellier II, I3M UMR-CNRS 5149
Place Eugène Bataillon CC 051 - 34095, Montpellier, France,
(e-mail: saidane@math.univ-montp2.fr)
(e-mail: bry@math.univ-montp2.fr)
[2] Université Paul-Valéry, Montpellier III
route de Mende - 34095, Montpellier, France,
(e-mail: Christian.Lavergne@math.univ-montp2.fr)

Abstract. Factor models have been fully developed and dealt with in the case where observations are assumed to be normally distributed. Here, we consider the less restrictive framework in which the distribution of the observations is assumed to belong to the exponential family. Thus, we introduce a new class of factor models allowing to analyze and predict discrete data (binomial, Poisson...), but also non-normal continuous data (gamma, for instance). These Generalized Linear Factor Models (GLFM) are built up combining standard Factor Models with Generalized Linear Models (GLM). A new parameter estimation method is presented for the GLFM. It is based on Fisher's Score algorithm for non-standard GLM, combined with an Expectation-Maximization (EM) type iterative algorithm for latent factors. Extensive Monte Carlo simulations show promising results.
Keywords: Factor Models, Generalized Linear Models, EM Algorithm, Scores Algorithm, Simulations..

1 Introduction

Latent variable models are widely used in social sciences for studying the interrelationships among observed variables. More specifically, latent variable models are used for reducing the dimensionality of multivariate data, for assigning scores to sample members on the latent dimensions identified by the model, and for constructing measurement scales (e.g., in psychometrics). [6,7] proposed a generalized linear latent variable model framework for any type of observed data (metric or categorical) in the exponential family. They extended the work of [5] and [11] for mixed binary and metric variables (the latter with covariate effects as well) and [2] for categorical variables. A similar framework was also discussed by [12] that includes multilevel models (random-effects models) as a special case.

In this paper we develop a general approach to factor analysis that involves observed variables that are assumed to be distributed in the exponential family. It accommodates a great variety of data, including rating, ordering,

choice, frequency, and timing data and entails a number of special cases of factor analysis not considered previously.

The framework is that of factor models (FM): a set of q observed random variables (RV) $\{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_q\}$ is assumed to be produced by fewer ($k < q$) unobserved (latent) ones, $\{\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_k\}$, called factors. The factors are supposed to account for the dependencies among the response variables in the sense that if the factors are held fixed, then the observed variables are independent. This is known as the assumption of conditional or local independence. So far, most developments on FM's were limited by the assumption that $\{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_q\}$ are normally distributed, and used this specific distribution to carry out their estimation, through the EM algorithm.

Here, we want to extend FM's to any type of distribution belonging to the exponential family: binomial, gamma, Poisson, etc. Therefore, we must also deal with the framework of generalized linear models (GLM), which only take observed variables as predictors, and are estimated using these observed values. So far, FM's and GLM's have been developed independently.

In this work, we propose a class of models - Generalized Linear Factor Models (GLFM) - in which, conditional to the factors $\{\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_k\}$, each $\mathbf{y}_i$ is modeled with a GLM taking these factors as predictors. For identification purposes, the factors are taken uncorrelated and normally distributed with 0 mean and unit variance. The independence assumption for the latent variables can be relaxed. Moreover, [1] showed that the choice of the latent variable distribution has a negligible effect on the interpretation of the results. He suggested using the normal distribution because it has rotational advantages when it comes to more than one latent variable.

In this paper we intend to estimate the model by maximizing the likelihood function. In the estimation we consider the unobserved factor scores, $\{\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_k\}$, as missing data. Then, we use the EM algorithm to learn parameters from the incomplete data. The problem here is that the the EM algorithm - using explicit expression of the expected completed log-likelihood of parameters conditional to observations - does not directly extend to non-normal distributions. To circle this difficulty, we consider the GLM's estimation algorithm that iteratively linearizes the model and performs Generalized Least Squares on it, and we propose to apply the EM procedure "locally" to this linearized GLM.

2 General structure of the GLFM

2.1 Model of the dependent variable Y conditional to factors

We consider n observation units $\{1, ..., t, ...n\}$. Let $\mathbf{y}_t = (y_{it})_{i=1,q}$ and $\mathbf{f}_t = (f_{jt})_{j=1,k}$ respectively be the vector of observed variables $\{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_q\}$ and latent factors $\{\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_k\}$ for unit t.

Conditional to factors $\mathbf{f}_t$, $(y_{it})_{i=1,q}$ are independently distributed according to a model having an exponential structure [8]:

$$\ell_i(y_{it}|\delta_{it},\phi) = \exp\left\{\frac{(y_{it}\delta_{it} - b_i(\delta_{it}))}{a_{it}(\phi)} + c_i(y_{it},\phi)\right\}$$

Let us recall classical results for this structure:

$$\mu_{it} = \mathbb{E}(y_{it}) = b_i'(\delta_{it}); \quad Var(y_{it}) = a_{it}(\phi)b_i''(\delta_{it}) = a_{it}(\phi)b_i''\left[b_i'^{-1}(\mu_{it})\right]$$

Let $\nu_i = b_i''\left[b_i'^{-1}(\mu_{it})\right]$. Independence of $(y_{it})_{i=1,q}$ conditional to $\mathbf{f}_t$ implies that they have conditional variance matrix:

$$Var(\mathbf{y}_i) = diag\left\{a_{it}(\phi)v_i(\mu_{it})\right\}_{t=1,\ldots,n}$$

2.2 Linear predictors

Stacking vectors $\mathbf{f}_t$, we get the (n,k) factor matrix $\mathcal{F} = [\mathbf{f}_1,...,\mathbf{f}_t,...,\mathbf{f}_n]'$. We assume that, underlying variables $\{\mathbf{y}_1,\mathbf{y}_2,...,\mathbf{y}_q\}$, are predictors $\eta = \{\eta_1,\eta_2,...,\eta_q\}$ that are linear combinations of the factors. Let $\boldsymbol{\theta} = (\theta_1,...,\theta_q)'$ be the vector of fixed effects. Most generally, these effects may depend on covariates, but in order to simplify our developments, we here take them constant.

For all i, let $\widetilde{\boldsymbol{\theta}}_i = \theta_i\mathbf{1}_n$. Then, the linear predictor of $\mathbf{y}_i$ conditional to $\mathcal{F}$ may be written as a vector in $\mathbb{R}^n$: $\eta_i = \widetilde{\boldsymbol{\theta}}_i + \mathcal{F}a_i$, where a_i is a k-coefficient vector. Let $A = (a_1,...,a_q)'$ be the (q,k) coefficient matrix. In matrix form, we have:

$$\eta = \boldsymbol{\theta}\mathbf{1}_n' + A\mathcal{F}'$$

Column t corresponds to unit t:

$$\eta_t = \boldsymbol{\theta} + A\mathbf{f}_t$$

The distribution assumption of factors is such that:

$$\forall t,\ \mathbf{f}_t \sim \mathcal{N}(\mathbf{0},\mathbf{I}_k)$$

2.3 Link function

The linear predictor and the expectation of the dependent variable $\mathbf{y}_i$ are linked through a *link function* g_i:

$$\forall i,\ t:\ \eta_{it} = g_i(\mu_{it})$$

Amongst all link functions, that which allows to equate the linear predictor and the canonical parameter is called canonical link function. We have:

$$\mu_{it} = b_i'(\delta_{it}) \ \Rightarrow\ \eta_{it} = g_i\left(b_i'(\delta_{it})\right)$$

So, the canonical link function is: $g_i = b_i'^{-1}$

3 Estimation of the GLFM

As, conditional to the factors, the GLFM boils down to a GLM, we first recall the overall structure of the GLM estimation algorithm, which also allows to introduce our notations. Then, we give back their latent random variable status to the factors, and adapt the estimation procedure to this situation by including an EM step in its current iteration.

3.1 Estimation of a GLM

Consider the GLM of some variable y, with $\mu = \mathbb{E}(y)$. Let $X = (x_1, ..., x_t, ..., x_n)'$ be the (n, k) observed predictor matrix. Let g be the link function, and η the linear predictor:

$$\eta = X\beta, \quad \beta \in \mathbb{R}^k$$

For each unit t, we have:

$$\eta_t = g(\mu_t) \;\Rightarrow\; x_t'\beta = g\left(b'(\delta_t)\right)$$

The problem is to estimate β. The log-likelihood of the model is:

$$\mathcal{L}(\delta; y) = \sum_{t=1}^{n} \mathcal{L}_t(\delta_t; y_t) = \sum_{t=1}^{n} \left[\frac{y_t\delta_t - b(\delta_t)}{a_t(\phi)} + c(y_t, \phi)\right]$$

Derivation with respect to β yields:

$$\begin{aligned}\frac{\partial \mathcal{L}_t}{\partial \beta_j} &= \frac{\partial \eta_t}{\partial \beta_j}\frac{\partial \mu_t}{\partial \eta_t}\frac{\partial \delta_t}{\partial \mu_t}\frac{\partial \mathcal{L}_t}{\partial \delta_t} = x_{tj}\frac{1}{g'(\mu_t)}\frac{1}{b''(\delta_t)}\frac{y_t - \mu_t}{a_t(\phi)} \\ \Rightarrow\; \frac{\partial \mathcal{L}}{\partial \beta_j} &= \sum_{t=1}^{n} x_{tj}\frac{1}{g'(\mu_t)^2 var(y_t)}g'(\mu_t)(y_t - \mu_t)\end{aligned}$$

Let:

$$W_\beta = diag\left[g'(\mu_t)^2 V(y_t)\right]_{t=1,n} = diag\left[g'(\mu_t)^2 a_t(\phi)v(\mu_t)\right]_{t=1,n}$$

and

$$\frac{\partial \eta}{\partial \mu} = diag\left(\frac{\partial \eta_t}{\partial \mu_t}\right)_{t=1,n} = diag\left(g'(\mu_t)\right)_{t=1,n}$$

Then, likelihood equations can be written:

$$X'W_\beta^{-1}\frac{\partial\eta}{\partial\mu}(y-\mu)=0 \tag{1}$$

This equation system not being linear in β, it is solved using an iterative process, known as Fisher's *scores algorithm*. If $m^{[e]}$ denotes the value of element m after iteration e:

$$\begin{aligned}\beta^{[e+1]} &= \beta^{[e]} - \left(\mathbb{E}\left[\left\{\frac{\partial^2\mathcal{L}}{\partial\beta\partial\beta'}\right\}\right]^{[e]}\right)^{-1}\frac{\partial\mathcal{L}^{[e]}}{\partial\beta}\\ &= \beta^{[e]} + \left(X'W_{\beta^{[e]}}^{-1}X\right)^{-1}X'W_{\beta^{[e]}}^{-1}\frac{\partial\eta^{[e]}}{\partial\mu}\left(y-\mu^{[e]}\right)\\ &= \left(X'W_{\beta^{[e]}}^{-1}X\right)^{-1}X'W_{\beta^{[e]}}^{-1}z^{[e]}\end{aligned} \tag{2}$$

where:

$$z_\beta=\eta+\frac{\partial\eta}{\partial\mu}(y-\mu)=X\beta+\frac{\partial\eta}{\partial\mu}(y-\mu) \tag{3}$$

Then, (1) becomes:

$$X'W_\beta^{-1}(z_\beta-X\beta)=0 \tag{4}$$

Equations (4) with given z_β may be interpreted as GLS normal equations in the linear model:

$$\begin{gathered}\mathcal{M}: \quad z_\beta=X\beta+\zeta\,, \quad \text{where}: \ \mathbb{E}(\zeta)=0\ ;\ V(\zeta)=W_\beta\\ \text{(indeed: } V(\zeta_t)=V(z_{\beta,t})=g'(\mu_t)^2Var(y_t)\text{)}\end{gathered}$$

So, current iteration e of the estimation algorithm consists in solving $X'W_{\beta^{[e]}}^{-1}(z_{\beta^{[e]}}-X\beta)=0$ with respect to β, and updating β in W_β and z_β with the solution.

We shall refer to $\mathcal{M}^{[e]}: z_{\beta^{[e]}}=X\beta+\zeta^{[e]};\ \ \mathbb{E}(\zeta^{[e]})=0;\ \ V(\zeta^{[e]})=W_{\beta^{[e]}}$ as the (current) *linearized model*. One important point is that GLS estimation of this model is nothing but a Quasi-Likelihood Estimation (QLE). This estimation by maximum of QL mimics MLE on each step, under a normality and independence assumption of the $z_{\beta,t}$'s with a fixed covariance structure.

Note: as the 1st order development of g at point μ yields:

$$g(y)\approx g(\mu)+g'(\mu)(y-\mu)=z$$

we may perform OLSR of $g(y)$ on X, in order to get an initial value $\beta^{[0]}$.

3.2 Estimation of a GLFM

In the case of a classical FM [10], estimation is handily carried out using the EM algorithm [4], which then requires that all variables be normally distributed, and maximizes the expectation of the completed log-likelihood conditional to observations, integrated with respect to the factors. According to the previous section, this normality assumption may be formally used with the linearized GLM within current step e, since GLS mimics normal MLE [3].

In the case of a GLFM, the estimation principle we propose is then informally straightforward. We consider the model alternately as:

- a GLM model conditional to $\mathcal{F} = (\mathbf{f}_1, ..., \mathbf{f}_t, ..., \mathbf{f}_n)'$
- a FM within the current estimation step of this GLM, as this step uses a linearized version of the GLM.

To be more precise, conditional to the current values of $\boldsymbol{\theta}$, A, $\mathcal{F}$, and following (3), we introduce the pseudo-dependent working variable z, which is then known:

$$z_{i,\mathcal{F}} = \widetilde{\theta}_i + \mathcal{F}a_i + \frac{\partial \eta_{i,\mathcal{F}}}{\partial \mu_{i,\mathcal{F}}}(\mathbf{y}_i - \mu_{i,\mathcal{F}}) = \widetilde{\theta}_i + \mathcal{F}a_i + g'(\mu_{i,\mathcal{F}})(\mathbf{y}_i - \mu_{i,\mathcal{F}})$$
$$\text{let} \quad \zeta_{i,\mathcal{F}} = g'(\mu_{i,\mathcal{F}})\varepsilon_{i,\mathcal{F}} \quad \text{with} \quad \varepsilon_{i,\mathcal{F}} = \mathbf{y}_i - \mu_{i,\mathcal{F}}$$

This intermediate z variable is used in the following estimation algorithm. Let $\forall\, t\ z_t = (z_{1t}, ..., z_{qt})'$, and $Z = (z_1, ..., z_t, ..., z_n)'$:

(i) Given Z and $V(\zeta)$, the model - called *linearized marginal model* - is:

$$\forall\, t = 1, n\ z_t = \boldsymbol{\theta} + A\mathbf{f}_t + \zeta_t$$

It is viewed as a non-standard FM, and estimated through an EM step, yielding $\mathcal{F}$. Since $\mathbf{f}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_k)$, we have:

$$V(z_t) = \boldsymbol{\Sigma} = AA' + \boldsymbol{\Psi}$$

with

$$\boldsymbol{\Psi} = \mathbb{E}(\boldsymbol{\Psi}_t) = \mathbb{E}\left(diag(g'(\mu_{i,\mathbf{f}_t})^2 var(\varepsilon_{it}|\mathbf{f}_t))\right)_{i=1,q}$$

If g is the canonical link function, we have:

$$\boldsymbol{\Psi} = \mathbb{E}(\boldsymbol{\Psi}_t) = \mathbb{E}\left(diag(a_{it}(\phi)g'(\mu_{i,\mathbf{f}_t}))\right)_{i=1,q}$$

Matrix $\boldsymbol{\Sigma}$ is the variance matrix used through the EM algorithm. It is analytically calculated for all classical canonical link functions.

(ii) Given $\mathcal{F}$, the model - called linearized *conditional* model - is viewed as a GLM, and parameters $\boldsymbol{\theta}$ and A are updated through Fisher's scores algorithm. This algorithm uses the variance matrix of ζ *conditional* to $\mathcal{F}$:

$$V(z_t|\mathbf{f}_t) = V(\zeta_t) = \boldsymbol{\Psi}_t$$

(ii) Variance matrix $V(\zeta)$ and z are then updated.

4 Experimental results

We present simulations carried out on a GLFM with two common factors, based on the Poisson distribution ($g = \log$). The simulated data vector has size $q = 40$ with $k = 2$ and $n = 400$. The convergence threshold

$$N = \max_{i \in \{1,\ldots,k\}} \left\{ \sum_{t=1}^{n} \left(f_{it}^{[e+1]} - f_{it}^{[e]} \right)^2 \right\}$$

was taken equal to 10^{-5}.[1] Initial parameter values for the EM algorithm were obtained through random perturbation of the real parameter values. As EM also requires an initial value for z, we used the following approximation:

$$\forall\, i = 1, q\,; t = 1, n \quad z_{it}^{[0]} = \log\left[\alpha y_{it} + (1-\alpha)\overline{\mathbf{y}}_i\right], \quad \text{with } \alpha = 0.5$$

The rationale behind the use of $\alpha < 1$ is to circle difficulties due to zero-values in data.

Our tests showed a good behaviour of the algorithm both at parameter and factor estimation. The sample means are very close to the true ones, and the standard deviations are small. Furthermore, the convergence threshold was reached after 7 iterations.

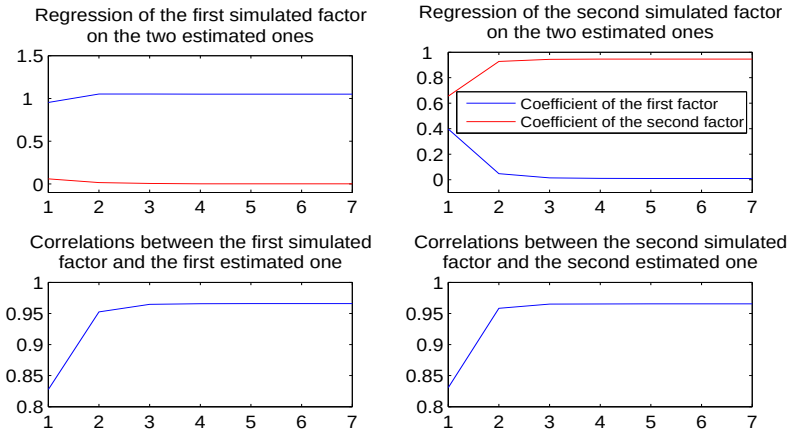


Fig. 1. Correlations between simulated factors and their estimation.

Results from the regression of the simulated factors $\mathbf{f}_t$ on the estimated factors $\widetilde{\mathbf{f}}_t$ (e.g., $f_{1t} = \beta_1 \widetilde{f}_{1t} + \gamma_1 \widetilde{f}_{2t} + \nu_t$ and $f_{2t} = \beta_2 \widetilde{f}_{1t} + \gamma_2 \widetilde{f}_{2t} + \nu_t$) given in figure 2 show that the regression coefficients β_1 and γ_2 converge to one,

[1] $[e]$ is the iteration number

while β_2 and γ_1 are close to zero. The correlations between simulated factors and their estimation was very close to 1 ($r_{f_1,\tilde{f}_1} > 95\%$, $r_{f_1,\tilde{f}_2} \approx r_{f_2,\tilde{f}_1} \approx 0$ and $r_{f_2,\tilde{f}_2} > 95\%$) which reconfirm again this result (Figure 1).

Finally, using the empirical Kullback-Leibler $\widetilde{K}_n$ divergence (see [9]) in order to measure the distance of estimators $\widetilde{\Theta}$ from the true parameters Θ_0 we have concluded a general decrease in average and spread of the distances with increasing n. Given that small values of $\widetilde{K}_n$ imply similarity between Θ_0 and $\widetilde{\Theta}_n$, the results of this experiment suggest an increasing accuracy and stability of the estimators as n increases.

We are currently performing tests on a variety of situations involving variables $\mathbf{y}_i$ modeled through *distinct* GLM's conditional on the same factors.

References

1. Bartholomew, D.J., "The Sensitivity of Latent Trait Analysis to Choice of Prior Distribution", *British Journal of Mathematical and Statistical Psychology*, 41, 101–107 (1988).
2. Bartholomew, D. J., and Knott, M., *Latent Variable Models and Factor Analysis*, London: Arnold (1999).
3. Lavergne C. and Trottier C., "From a conditional to a marginal point of view in GL2M", in *Proceedings in Good Statistical Practice*, Seeber, 205–209 (1997).
4. McLachlan, G.J., and Krishnan, T., *The EM algorithm and extensions*, Wiley Series in Probability and Statistics, NY: John Wiley & Sons, Inc (2008).
5. Moustaki, I., "A Latent Trait and a Latent Class Model for Mixed Observed Variables", *British Journal of Mathematical and Statistical Psychology*, 49, 313–334 (1996).
6. Moustaki, I., and Knott, M., "Generalized Latent Trait Models", *Psychometrika*, 65, 391–411 (2000).
7. Moustaki, I and Victoria-Feser, M.P., "Bounded-influence robust estimation in generalized linear latent variable models", *Journal of the American Statistical Association*, 104, 644–653 (2006).
8. Nelder, J.A., and Wedderburn, R.W.M., "Generalized linear models", *Journal of the Royal Statistical Society: Series A*, 135, 370–384 (1972).
9. Saidane, M. and Lavergne, C., "An EM-based Viterbi Approximation Algorithm for Mixed-State Latent Factor Models", *Communications in Statistics - Theory and Methods* 37, 2795–2814 (2008).
10. Saidane, M. and Lavergne, C., *Modelling and Forecasting Volatility Dynamics Using Quadratic GARCH-Factor Models: Empirical Evidence from International Foreign Exchange Markets*, in Stock Returns: Cyclicity, Prediction and Economic Consequences, Nova Science Publishers, (2009).
11. Sammel, M. D., Ryan, L.M., and Legler, J.M., "Latent Variable Models for Mixed Discrete and Continuous Outcomes", *Journal of the Royal Statistical Society, Ser. B*, 59, 667–678 (1997).
12. Skrondal, A., and Rabe-Hesketh, S., *Generalized Latent Variable Modelling: Multilevel, Longitudinal, and Structural Equation Models*, London: Chapman & Hall (2006).

Estimation of Skew *t*-Distribution by the Monte-Carlo Markov Chain Approach

Leonidas Sakalauskas and Ingrida Vaiciulyte

Institute of Mathematics & Informatics
Vilnius, Lithuania
Email: sakal@ktl.mii.lt
Email: ingrida_vaiciulyte@yahoo.com

Abstract. The Monte-Carlo Markov Chain (MCMC) method for estimation of the skew *t*-distribution is presented. The skew *t*-distribution is represented by a multivariate skew - normal distribution with the covariance matrix depending on the parameter, distributed according to the inverse - gamma distribution (Azzalini and Genton, 2008). Thus, the density of the skew *t*-distribution is expressed through a multivariate integral. Next, the MCMC procedure is constructed for recurrent estimation of the skew *t*–distribution by maximum likelihood, where the Monte-Carlo sample size is regulated so that to ensure the convergence and to decrease the total amount of Monte-Carlo trials needed for estimation. The confidence intervals of Monte-Carlo estimators are introduced because of the asymptotic Gaussian distribution of Monte-Carlo estimators and the termination rule is implemented testing statistical hypotheses on an insignificant change of estimates in two contiguous chains of the procedure (Sakalauskas, 2000).
Keywords: Statistical simulation, Monte-Carlo method, Maximum likelihood, Gaussian approximation, EM - algorithm

1 Introduction

In recent time, there has been a growing interest in the analysis of parametric classes of distributions that exhibit various shapes of skewness and kurtosis. To model departures of such a distribution from normality, a well-known approach consists in modifying the probability density function of a random vector in a multiplicative fashion (Azzalini and Genton, 2008). A multivariate skew *t*–distribution, which is often applied to model non–Gaussian errors, is constructed in this way, too. In general, the skew *t*-distribution is represented by a multivariate skew - normal distribution with the covariance matrix depending on the parameter, distributed according to the inverse-gamma distribution. According to this representation, the density of the skew *t*–distribution as well as the likelihood function are expressed through multivariate integrals that are convenient to estimate numerically by Monte-Carlo simulation (Cabral et all, 2008). In this paper, the maximum likelihood method for estimating the parameters of the multivariate skew *t*-distribution is developed using the adaptive Monte-Carlo Markov chain approach.

2 The Maximum Likelihood Estimation of Multivariate Skew *t*-Distribution

Denote the skew *t*–variable by $ST(\mu,\Sigma,\Theta)$. In general, a multivariate skew *t*–distribution defines a random vector *X*, that is distributed as a multivariate Gaussian vector:

$$f(x,a,t,\Sigma)=(t/\pi)^{\frac{d}{2}}\cdot\left|\Sigma\right|^{-\frac{1}{2}}\cdot e^{-t\cdot(x-a)^T\cdot\Sigma^{-1}\cdot(x-a)}, \tag{1}$$

where the vector of mean *a*, in it's turn, is distributed as a multivariate Gaussian $N(\mu,\Theta/2t)$ in the half–plane $q\cdot\omega^{-1}\cdot(a-\mu)\ge 0$, where $\omega=\mathrm{diag}(\Sigma)$, $\Sigma\ge 0,\ \Theta\ge 0$ are the full rank $d\times d$ matrices, *d* is the dimension, and the random variable *t* follows from the Gamma distribution:

$$f_1(t)=\frac{t^{\frac{b}{2}-1}}{\Gamma(b/2)}\cdot e^{-t}. \tag{2}$$

Assume, for simplicity, the parameter *b* to be fixed. By definition, the *d*–dimensional skew *t*–distributed variable *X* has the density:

$$p(x,\mu,\Theta,\Sigma)=2\cdot\int_0^\infty\int_{q\cdot\omega^{-1}\cdot(a-\mu)\ge 0} f(x,a,t,\Sigma)\cdot f(a,t,\mu,\Theta)\cdot f_1(t)\,da\,dt=$$

$$=\int_0^\infty\int_{q\cdot\omega^{-1}\cdot(a-\mu)\ge 0}\frac{2}{\pi^d\cdot\left|\Sigma\right|^{\frac{1}{2}}\cdot\left|\Theta\right|^{\frac{1}{2}}\cdot\Gamma\left(\frac{b}{2}\right)}\cdot t^{\frac{b}{2}+d-1}\times \tag{3}$$

$$\times e^{-t\cdot\left(\cdot(x-a)^T\cdot\Sigma^{-1}\cdot(x-a)+\cdot(a-\mu)^T\cdot\Theta^{-1}\cdot(a-\mu)+1\right)}da\,dt.$$

Let a matrix of observations be given $X=\left(X^1,X^2,\ldots,X^K\right)$, where X^i are independent vectors, distributed as $ST(\mu,\Sigma,\Theta)$. We will examine the estimation of parameters μ,Σ,Θ by the maximum likelihood method. Thus, the log–likelihood function is as follows:

$$L(\mu,\Sigma,\Theta)=-\sum_{i=1}^K\ln\left(p\left(X^i,\mu,\Sigma,\Theta\right)\right)=-\sum_{i=1}^K\ln\left(Ef\left(X^i,a,t,\Sigma\right)\right), \tag{4}$$

where the expectation is taken with respect to *a* and *t*, distributed as described above. The estimates $\hat{\mu},\hat{\Sigma},\hat{\Theta}$ of parameters of multivariate skew *t*–distribution (3) are found by taking and setting equal to zero the first derivatives, and then solving the equations obtained by this way subject to $\Sigma\ge 0,\ \Theta\ge 0$.

Derivatives of the likelihood function are expressed through derivatives of the density function:

$$\frac{\partial L(\mu,\Sigma,\Theta)}{\partial\mu}\equiv-\sum_{i=1}^K\frac{\partial p(X^i,\mu,\Sigma,\Theta)}{\partial\mu}\cdot\frac{1}{p(X^i,\mu,\Sigma,\Theta)},$$

$$\frac{\partial L(\mu,\Sigma,\Theta)}{\partial \Sigma} \equiv -\sum_{i=1}^{K} \frac{\partial p(X^i,\mu,\Sigma,\Theta)}{\partial \Sigma} \cdot \frac{1}{p(X^i,\mu,\Sigma,\Theta)},$$

$$\frac{\partial L(\mu,\Sigma,\Theta)}{\partial \Theta} \equiv -\sum_{i=1}^{K} \frac{\partial p(X^i,\mu,\Sigma,\Theta)}{\partial \Theta} \cdot \frac{1}{p(X^i,\mu,\Sigma,\Theta)}.$$

Differentiation of the density function of skew *t*-distribution (3) provides us:

$$\frac{\partial p(x,\mu,\Sigma,\Theta)}{\partial \mu} = \int_0^{\infty} \int_{q\cdot\omega^{-1}\cdot(a-\mu)\geq 0} t\cdot\Sigma^{-1}\cdot(x-a)\cdot f(x,a,t,\Sigma)\cdot f(a,t,\mu,\Theta)\times$$

$$\times f_1(t)\,dadt,$$

$$\frac{\partial p(x,\mu,\Sigma,\Theta)}{\partial \Sigma} = \int_0^{\infty} \int_{q\cdot\omega^{-1}\cdot(a-\mu)\geq 0} \left(-\Sigma^{-1} + t\cdot\Sigma^{-1}\cdot(x-a)\cdot(x-a)^T\cdot\Sigma^{-1}\right)\times$$

$$\times f(x,a,t,\Sigma)\cdot f(a,t,\mu,\Theta)\cdot f_1(t)\,dadt,$$

$$\frac{\partial p(x,\mu,\Sigma,\Theta)}{\partial \Theta} = \int_0^{\infty} \int_{q\cdot\omega^{-1}\cdot(a-\mu)\geq 0} \left(-\Theta^{-1} + t\cdot\Theta^{-1}\cdot(a-\mu)\cdot(a-\mu)^T\cdot\Theta^{-1}\right)\times$$

$$\times f(x,a,t,\Sigma)\cdot f(a,t,\mu,\Theta)\cdot f_1(t)\,dadt.$$

Denote the conditional density:

$$f(a,t,\mu,\Sigma,\Theta|x) = \frac{f(x,a,t,\Sigma)\cdot f(a,t,\mu,\Theta)\cdot f_1(t)\,dadt}{p(x,\mu,\Sigma,\Theta)}.$$

Using this definition, the derivatives of the likelihood function can be written in the following form:

$$\frac{\partial p(x,\mu,\Sigma,\Theta)}{\partial \mu} = E\left(t\cdot\Sigma^{-1}\cdot(x-a)|x\right)\cdot p(x,\mu,\Sigma,\Theta),$$

$$\frac{\partial p(x,\mu,\Sigma,\Theta)}{\partial \Sigma} = E\left(-\Sigma^{-1} + t\cdot\Sigma^{-1}\cdot(x-a)\cdot(x-a)^T\cdot\Sigma^{-1}|x\right)\times$$

$$\times p(x,\mu,\Sigma,\Theta),$$

$$\frac{\partial p(x,\mu,\Sigma,\Theta)}{\partial \Theta} = E\left(-\Theta^{-1} + t\cdot\Theta^{-1}\cdot(a-\mu)\cdot(a-\mu)^T\cdot\Theta^{-1}|x\right)\times$$

$$\times p(x,\mu,\Sigma,\Theta).$$

Let $\hat{\mu}, \hat{\Sigma} > 0, \hat{\Theta} > 0$ be the maximum likelihood estimates of parameters of $ST(\mu,\Sigma,\Theta)$. It is easy to see that now these estimates satisfy the equations:

$$\frac{1}{K}\sum_{i=1}^{K} E\left(t\cdot(X^i-a)\middle|X^i\right) = 0, \tag{5}$$

$$\hat{\Sigma} = \frac{1}{K}\sum_{i=1}^{K} E\left(t\cdot(X^i-a)\cdot(X^i-a)^T\middle|X^i\right), \tag{6}$$

$$\hat{\Theta} = \frac{1}{K}\sum_{i=1}^{K} E\left(t\cdot(a-\hat{\mu})\cdot(a-\hat{\mu})^T\middle|X^i\right), \tag{7}$$

where the conditional expectation is taken for $\hat{\mu}, \hat{\Sigma}, \hat{\Theta}$.

3 Monte–Carlo Markov Chain

Now it is convenient to calculate the estimates of parameters by an iterative method, starting from the initial values. Let us consider the EM – algorithm to solve equations (5)-(7). The recurrent EM relationships are as follows:

$$\mu_{k+1} = \mu_k + \frac{1}{K}\sum_{i=1}^{K} E\left(t\cdot(X^i - a)\middle| X^i\right), \tag{8}$$

$$\Sigma_{k+1} = \frac{1}{K}\sum_{i=1}^{K} E\left(t\cdot(X^i - a)\cdot(X^i - a)^T\middle| X^i\right), \tag{9}$$

$$\Theta_{k+1} = \frac{1}{K}\sum_{i=1}^{K} E\left(t\cdot(a - \mu_k)\cdot(a - \mu_k)^T\middle| X^i\right), \tag{10}$$

where conditional expectations are computed for $\mu_k, \Sigma_k, \Theta_k$, and $\mu_0, \Sigma_0, \Theta_0$ are some initial approximations, *k=0,1,2,....* The process is terminated, if the estimates during two current iterations differ insignificantly.

Since the integrals in the expressions obtained can be calculated analytically only in very simple cases, it is reasonable to apply the Monte-Carlo method. Say, random variables and vectors are generated:

$$B_j \sim \text{Gama}(\frac{b}{2}),$$

$$\eta_j \sim \text{N}(0, \Theta_k),$$

$$G_j = \begin{cases} \mu_k + \eta_j, \text{if } q\cdot\omega^{-1}\cdot\eta_j \geq 0, \\ \mu_k - \eta_j, \text{if } q\cdot\omega^{-1}\cdot\eta_j < 0, \end{cases}$$

where *j=0,1,2,...,N^k*, *N^k* is the Monte–Carlo simple size at the *k^th* step. Then

$$\mu_{k+1} = \mu_k + \frac{1}{K}\sum_{i=1}^{K}\frac{M_{i,k}}{P_{i,k}}, \tag{11}$$

$$\Sigma_{k+1} = \frac{1}{K}\sum_{i=1}^{K}\frac{S_{i,k}}{P_{i,k}}, \tag{12}$$

$$\Theta_{k+1} = \frac{1}{K}\sum_{i=1}^{K}\frac{T_{i,k}}{P_{i,k}}, \tag{13}$$

where the Monte–Carlo estimators are as follows:

$$P_{i,k} = \frac{1}{N^k}\sum_{j=1}^{N^k} f(X^i, G_j, B_j, \Sigma_k), \tag{14}$$

$$M_{i,k} = \frac{1}{N^k}\sum_{j=1}^{N^k}\left(X^i - G_j\right)\cdot B_j \cdot f(X^i, G_j, B_j, \Sigma_k), \tag{15}$$

$$S_{i,k} = \frac{1}{N^k}\sum_{j=1}^{N^k}\left(X^i - G_j\right)\cdot\left(X^i - G_j\right)^T \cdot B_j \cdot f(X^i, G_j, B_j, \Sigma_k), \tag{16}$$

$$T_{i,k} = \frac{1}{N^k}\sum_{j=1}^{N^k}(G_j - \mu_k)\cdot(G_j - \mu_k)^T \cdot B_j \cdot f(X^i, G_j, B_j, \Sigma_k). \quad (17)$$

The Monte-Carlo estimate of the log-likelihood function (4) is obtained using estimate (14):

$$L_k = -\sum_{i=1}^{K}\ln(P_{i,k}). \quad (18)$$

The statistical modelling error of the log-likelihood function can be evaluated for a large sample size N_k in the following way:

$$L_k = L(\mu_k, \Sigma_k, \Theta_k) - \sum_{i=1}^{K}\ln\left(1 + \frac{P_{i,k} - EP_{i,k}}{EP_{i,k}}\right) = \\ = L(\mu_k, \Sigma_k, \Theta_k) - \sum_{i=1}^{K}\left(\frac{P_{i,k} - EP_{i,k}}{EP_{i,k}}\right) + o(\frac{1}{\sqrt{N_k}}) \quad (19)$$

By virtue of (19) the variance of estimate (18) can be evaluated as:

$$D^2(L_k) = \sum_{i=1}^{K}\left(\frac{E\left(f\left(X^i, a, t, \Sigma\right)\right)^2}{\left(Ef\left(X^i, a, t, \Sigma\right)\right)^2} - 1\right) + o\left(\frac{1}{N^k}\right).$$

Hence, the 95% confidence interval of the estimate of the log-likelihood function can be also estimated by the Monte-Carlo method:

$$\left[L_k - \frac{2}{\sqrt{N^k}}\cdot\sqrt{\sum_{i=1}^{K}\left(\frac{P2_{i,k}}{P_{i,k}^{\ 2}} - 1\right)}, L_k + \frac{2}{\sqrt{N^k}}\cdot\sqrt{\sum_{i=1}^{K}\left(\frac{P2_{i,k}}{P_{i,k}^{\ 2}} - 1\right)}\right], \quad (20)$$

where $P2_{i,k} = \frac{1}{N^k}\sum_{j=1}^{N^k}\left(f(X^i, G_j, B_j, \Sigma_k)\right)^2$.

Note that there is no reason to generate large samples starting the estimation since it suffices only to approximately evaluate the direction leading to the solution of equations (5)-(7). Thus, large samples should be taken only at the moment of the decision on termination of the Monte–Carlo Markov chain. To this end, the next rule of sample size regulation is implemented:

$$N^{k+1} \geq Z_{\beta,p}\cdot\frac{N^k}{H^k}, \quad (21)$$

where $Z_{\beta,p}$ is the quantile of Fisher distribution, β is the significance level. In the general case, α may be coincident with β. As follows from (Sakalauskas, 2000), such a rule guarantees the convergence of procedure (11-13) to the solution of equations (5)-(7). The Monte–Carlo chain can be terminated at the k^{th} step if

$$\mu_{k+1} \approx \mu_k, \Sigma_{k+1} \approx \Sigma_k, \Theta_{k+1} \approx \Theta_k, \quad (22)$$

and Monte-Carlo estimates are presented with an admissible confidence interval. Since estimators (14)-(17) are averages of a large number of identically distributed random variables, their distribution is approximated by CLT. Hence, the statistical

criteria about the equality of sampling mean and covariance matrices to the given vector or matrices can be used for testing termination condition (22). Thus, the hypothesis on the termination condition is rejected, if

$$H^k = K \cdot N^k \cdot \left[-\ln\left(\frac{|\Sigma_{k+1}|}{|\Sigma_k|}\right) - \ln\left(\frac{|\Theta_{k+1}|}{|\Theta_k|}\right) + (\mu_{k+1} - \mu_k)^T \cdot (\Sigma_k)^{-1} \times \right. \tag{23}$$

$$\left. \times (\mu_{k+1} - \mu_k) + \mathrm{SP}\left(\Sigma_{k+1} \cdot (\Sigma_k)^{-1}\right) + \mathrm{SP}\left(\Theta_{k+1} \cdot (\Theta_k)^{-1}\right) - 2 \cdot d \right] > Z_{\alpha,p},$$

where $Z_{\alpha,p}$ is the quantile of Fisher distribution with $p = d \cdot (d+3)$ degrees of freedom, α is the significance level.

Thus, using the adaptive Monte-Carlo Markov chain approach developed, the Monte-Carlo estimators (11)-(18) are calculated changing the Monte-Carlo sample size according to (21) and terminating the chain when the confidence interval (20) becomes shorter than the prescribed admissible value and criteria (23) don't reject the hypothesis on condition (22).

4 Computer Simulation

Let us consider the numerical example with the following model data:

d=2, b=5, $\mu = (1 \quad 2)$, $\Sigma = \begin{pmatrix} 1.61 & 0.27 \\ 0.27 & 2.9 \end{pmatrix}$, $\Theta = \begin{pmatrix} 3.67 & 0.86 \\ 0.86 & 2.55 \end{pmatrix}$.

The random $ST(\mu, \Sigma, \Theta)$ sample with K=500 has been simulated to explore the approach developed. The maximum likelihood estimates (5)-(7), obtained from this sample by means of the subroutine *Minimize()* of *MathCad*, are as follows:

$\hat{\mu} = (1.04 \quad 2.17)$, $\hat{\Sigma} = \begin{pmatrix} 1.503 & 0.189 \\ 0.189 & 3.631 \end{pmatrix}$, $\hat{\Theta} = \begin{pmatrix} 4.081 & 0.096 \\ 0.096 & 1.912 \end{pmatrix}$.

The Monte-Carlo Markov chain of 100 estimators (11)-(18) has been computed with initial data: $\mu_0 = 1.5 \cdot \mu$, $\Sigma_0 = 1.5 \cdot \Sigma$, $\Theta_0 = 1.5 \cdot \Theta$.

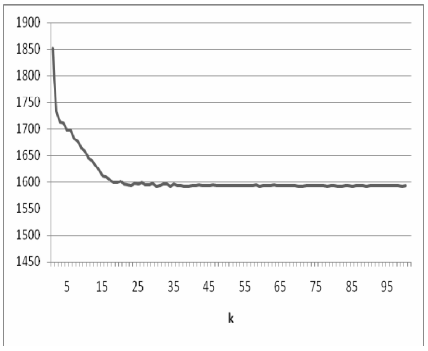


Fig. 1. Log-likelihood function

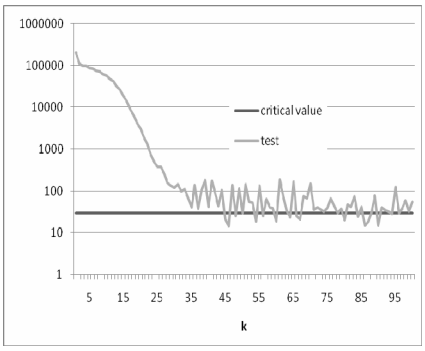


Fig. 2. Termination test

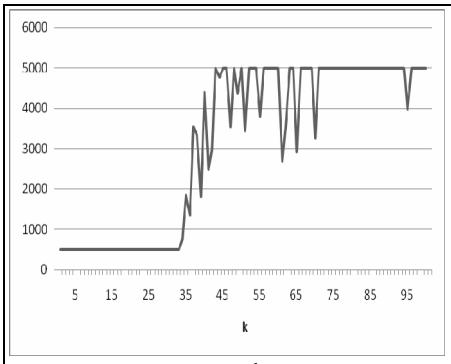


Fig. 3. Sample size N^k

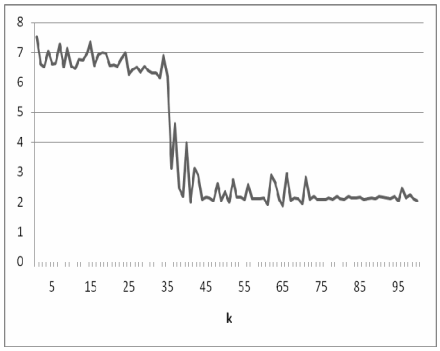


Fig. 4. Confidence interval

The changes of the log-likelihood function estimate (18), termination statistics (23), sample size (21) and the length of the confidence interval (20) are depicted in Fig's. 1-4. As we see in Fig. 1, the log-likelihood function is decreasing as long as the zone of possible solution is achieved. Correspondingly, the termination criteria in Fig. 2 are decreasing, too, until the critical value of termination is achieved. The sample size in Fig. 3 is changed so that it was small starting the chain and increase in the zone of possible solution. To avoid very small or very large sample sizes, the bounds were applied: $500 \le N^k \le 5000$. The length of the confidence interval in Fig. 4 and the error of estimates decreased as well. The termination conditions started to be valid at k=46. Thus, the following MCMC estimates have been obtained:

$$\mu_{46} = (1.13 \quad 2.20),\ \Sigma_{46} = \begin{pmatrix} 2.66 & -0.11 \\ -0.11 & 3.88 \end{pmatrix},\ \Theta_{46} = \begin{pmatrix} 2.61 & 0.15 \\ 0.15 & 1.53 \end{pmatrix}.$$

5 Conclusions

The Monte-Carlo Markov Chain (MCMC) method for estimation of the skew t-distribution has been developed in the paper. It distinguishes by adaptive regulation of Monte-Carlo sample size and treatment of the simulation error in the statistical manner. The computer counterexample has illustrated that numerical properties of the method correspond to the theoretical model.

References

1. Azzalini A., Genton M. G., Robust Likelihood Methods Based on the Skew-t and Related Distributions, *International Statistical Review*, **76**, 1, 106–129 (2008).

2. Cabral C. R. B., Bolfarine H., Pereira J. R. G., Bayesian density estimation using skew student-t-normal mixtures, *Computational Statistics and Data Analysis*, **52**, 12, 5075-5090 (2008).

3. Sakalauskas L., Nonlinear Stochastic Optimization by Monte-Carlo Estimators, *Informatica*, **11**, 4, 455-468 (2000).

Data exploration and analysis for obtaining mortality profiles from the animal bones remains found in archaeological sites *

A. Scaringella[1] and G. Siracusano[1]

Facoltà di Sociologia,
via Salaria 113, 00198 Roma, Italy
(e-mail: angela.scaringella@uniroma1.it})

Abstract. We discuss the problem of obtaining mortality profiles from some animal bones remains found in archaeological sites. The method takes account of different types of bones that can be related to different sets of possible ages of the animals. As a first preliminary step we obtain the likelihood function and derive the equations for the point of maximum likelihood.

Keywords: Mortality profile, Likelihood analysis, Archaeozoology.

1 Mortality profiles from bones data

The field of archaeozoolology deals with the study of animal remains in archaeological deposits as a result of human activity with the goal of reconstructing the relationship between man and animals. (see [3], [5], [6], [11], [12], [13], [15]). Among many factors widely used in archaeological investigation, the estimated age of death and the subsequent processing of the curves of mortality of animal populations are considered to be those of highest interest in order to identify the cultural implications in relation to farming systems. The analysis of animal bone remains helps archaeologists also to rebuild the cultural traditions of ancient peoples. helping to define the economy of subsistence of human settlements. The materials we are interested in come from the analysis of animal bones collected from levels VIII (Late Chalcolithic, 4300-4000 BC) and VID (Early Bronze Age III, 2500-2000 BC) of Arslantepe. After recovery of the remains on the excavation and examination of the feature concerning the archaeological context, and identification of taxonomic, anatomical of the fragments, the elements of diagnosis are considered, where and when its possible to ascertain, on the knitting of epiphyses, the fusion of sutures, the progress of the dentition, development and tooth abrasion, etc. .. These characters differ widely from individual to individual and this subject is a frequently discussed topic, as evidenced by the numerous studies to verify the presumed age of death. The age can only be determined

* the complete version appears in *Methodology and Computing in Applied Probability* (MCAP), Springer, 2009. DOI 10.1007/s11009-009-9145-3, http://www.springerlink.com/content/e86662j38168r247/

with some approximation, because it is almost impossible to state precisely at what age the bone development was interrupted.

The assessment of the age of death from the bones of animals collected during the archaeological excavations, is obtained from comparative studies with wild animals hunted, from whose anatomical and morphological evidence can be traced with some accuracy the age, or with animals slaughtered at a domestic note, of osteology collections. The variability of the stages of bone growth, both in the same individual and between different individuals of the same species, does not always allow a precise correlation between the periods of the development of various skeletal parts and the real age. In fact if we analyze the bones of an individual dead at known age and attribute an age to them according to the the state of skeletal development, we could easily obtain different results.

It is not easy to define exactly the age of death, as happens, for example, by observing the distal fused epiphysis of tibia of an animal, from which its possible to infer that this animal could certainly have more than two years, but it is impossible to say in the lack of other evidence, if it had three years or more. For this reason, despite the data on the fusion of the epiphysis and the eruption of teeth are reported in numerous osteological publications on the calculation of age of death, several questions still remain unanswered.

Normally one needs a large sample of bones to come up with a significant mortality curve, but it's not uncommon to find only a small number of samples. Discuss and interpret the data that may be few in number is sometimes hazardous, and their reliability is low. Whereas these data are only a fraction of the total and are not used together, a representation of populations of belonging may not be very representative.

In the determination of each researcher tries to identify with greater expertise as possible other factors, less obvious, but important in determining as far as possible the range of likely ages. Also one must take into account the historical contexts and on the biogeographic population of animals that judges consider most plausible. Inherent in the calculation of age of death of the animals based on the skeletal development there is a kind of uncertainty principle in its evaluation, which leads us to interpret and thus represent the strategies of hunting and breeding in an approximate way.

A method for taking account of all information contained in data is likelihood analysis based on Bayes formula (see f. e. [4], [7], [8], [9], [10]). The major advantage of this method is that it allows to take account at the same time of all the available data. We have started this analysis by writing the likelihood function and the equation for the point of maximum likelihood. Further work is required in order to devise and implement an efficient method to apply this kind of analysis to our data.

2 Likelihood analysis

We introduce a statistical model in order to obtain from the data information abot the death age distribution of the animals. Let the index $l = 0, 1, \ldots, N$ denote the possible values of the death age of the animals. Let f_l be the percentage of the animals at age l. Each animal has r types of "bones". The j-th type can be in one of s_j states. Let $p(i, m|l)$ be the probability for animal dead at age l that its bone of type j be in state m. The likelihood function for obtaining $k_m^{(j)}$ of bones of the j-th type in the m-th state is therefore

$$\prod_{j=1}^{r} \prod_{m=1}^{s_j} (\sum_{l=0}^{N} p(i, m|l) f_l)^{k_m^{(j)}}. \tag{1}$$

The log-likelihood is

$$\sum_{j=1}^{r} \sum_{m=1}^{s_j} k_m^{(j)} \log(\sum_{l=0}^{N} p(i, m|l) f_l). \tag{2}$$

The equations for the point of maximal likelihood are

$$\sum_{j=1}^{r} \sum_{m=1}^{s_j} k_m^{(j)} \frac{p(i, m|l)}{\sum_{l=0}^{N} p(i, m|l) f_l} = \lambda \tag{3}$$

for $l = 0, 1, \ldots, N$ and

$$\sum_{l=0}^{N} f_l = 1. \tag{4}$$

We plan to perform various methods of Bayes likelihood analysis including Monte Carlo Markov Chain simulation ([1], [10]) for obtaining reasonable mortality curves from our available data.

References

1. S. E. Ahmed, N. Reid. Empirical Bayes and Likelihood Inference. Lecture Notes in Statistics, Springer, 2000.
2. T. Amorosi. A postcranial guide to domestic neo-natal and juvenile mammals, *Oxford, BAR International series* 533, 1989.
3. D. Bullock, J. Rackham.1982. Ephiphyseal fusion and tooth erupting of feral goats from Moffaldale. Dumfries and Galloway, Scotland, In: Wilson B., Grigson c: Payne S (eds). , *Ageing and Sexing Animal Bones from Archaeological Sites BAR 109* Oxford, 55-72, 1982.
4. A. Chamberlain. Demography in Archaeology. Cambridge University Press. 2006.
5. R. E. Chaplin.The study of animal bones from archaeological sites. Seminar Press, London, 1971.

6. K. H. Habermehel. Altersbestimmung bei Haustieren, Pelztieren, und heim jagbaren Wildtieren. Verlag P. Parey, Berlin Hamburg, 1961.
7. L. W. Konigsberg, S. R. Frankenberg. Estimation of age structure in anthropological demography. *American Journal of Physical Anthropology* 89, 235-236, 1992.
8. L. W. Konigsberg, S. R. Frankenberg. Deconstructing death in paleodomography. *American Journal of Physical Anthropology* 117, 297-309, 2002.
9. L. W. Konigsberg, S. R. Frankenberg, R. B. Walker. Regress what on what? Paleodomographic age estimation as a calibration problem. In R. R. Paine (ed.) *Integrating Archaeological Demography: Multidisciplinary Approaches to Prehistoric Population* Carbondale: Southern Illinois University, 64-88.
10. L. W. Konigsberg, N. P. Herrmann. Markov chain Monte Carlo estimation of hazard model parameters in paleodemography. In R. D. Hoppa, J. W. Vaupel (eds.): *Paleodemography. Age Distributions from Skeletal Samples* Cambridge University Press, 222-242, 2002.
11. N. Moran, T. OConnor. Age attribution in in domestic sheep by skeletal and dental maturation: a pilot study of available sources. *International Journal Osteoarchaeozoology* 4, 267-285, 1994.
12. B. Noddle. Age of epiphyseal closure in feral and domestic goats and ages of dental eruption. *Journal of Archaeological Science* 12, 195-204, 1974.
13. S. Payne. Kill-off patterns in steep and goats: the mandibles from Aswan Kave. Anatolian Studies , vol. XXIII, 1973
14. I. A. Silver I.A. .The ageing of domestic animals, in D. Brothwell E. S. Higgs (Eds.) *Science in Archaeology*, Thames Hudson, London.283-302, 1969.
15. G. Siracusano. Spunti metodologici sui dati faunistici di Coppa Nevigata. In : *Atti 2 Convegno Nazionale di Archeozoologia*, Asti 1997. ABACO ed. Forl, 81-88, 2000.

A Hybrid Genetic Algorithm for Solving the Uncapacitated Multiple Allocation Hub Location Problem

Zorica Stanimirović[1], Maja Djukić[2] and Jozef Kratica[3]

[1] Faculty of Mathematics
University of Belgrade, Belgrade, Serbia
(e-mail: zoricast@matf.bg.ac.rs)
[2] Faculty of Mathematics
University of Belgrade, Belgrade, Serbia
(e-mail: mdjukic85@gmail.com)
[3] Mathematical Institute
Serbian Academy of Sciences and Arts, Belgrade, Serbia
(e-mail: jkratica@mi.sanu.ac.rs)

Abstract. In this paper, a hybrid genetic algorithm for solving the Uncapacitated Multiple Allocation Hub Location Problem is proposed. This NP-hard problem has significant application in designing modern transportation and telecommunication networks, such as road and railway systems, postal systems, systems of fast delivery, etc. In order to improve the efficiency, genetic algorithm is combined with the local search heuristic. The proposed hybrid method shows to be very successful in solving problems of large dimensions with up to $n = 120$ nodes. It is also tested on instances with $n = 130$ and $n = 200$ nodes for which no optimal solution is presented in the literature so far. Although the optimal solutions are not known, we believe that the proposed hybrid method provides high quality solutions on these problem instances unsolved to optimality before.
Keywords: Genetic Algorithms, Hub Location Problems, Combinatorial Optimization, Metaheuristics, Transportation and Telecommunication Networks.

1 Introduction

Hub location problems have become an important research field in location theory. They are frequently used as models in numerous applications including airline design, postal delivery, computer networks and various transportation and telecommunications networks

Hub nodes serve as concentration and consolidation points in the network by collecting traffic from one or more origins and distributing it to the destinations. Instead of serving each origin-destination pair directly, hub facilities consolidate flows in order to take advantage of economies of scale. Since there is an increase in the traffic between hub nodes, smaller transportation costs can be obtained, due to a discount factor for transport between the hubs.

According to the particular characteristics of the hub network, there are many variants of hub location problems. The number of hubs may or may

not be fixed in advance, there may be constraints on the capacities on the nodes or arcs in the network, fixed costs for establishing hub network may be assumed. In the hub network, each non-hub node is allocated to exactly one hub (single allocation scheme) or to several hubs (multiple allocation scheme). A general review of the different hub location problems can be found in the classical paper [5] or in more recent survey [2].

2 The Uncapacitated Multiple Allocation Hub Location Problem

Let $I = 1, ..., n$ be the set of nodes in the network exchanging some traffic. The distance between nodes i and j is C_{ij}, satisfying the triangle inequality [5]. The number of units of traffic to be sent from node i to j is denoted as W_{ij}. It is assumed that the hubs are fully interconnected and every non-hub node can send traffic to different hubs (multiple allocation scheme). The traffic between a pair of nodes i and j has to be routed through either one or two established hubs and it consists of three components: collection of flow from origin i at hub k, transfer between hubs k and m and distribution from hub k to destination j. If $k = m$, the traffic is routed just via one hub. Establishing a hub at potential location k causes fixed costs f_k. The decision variables are:

- $y_k = 1$, if node k is established as a hub, 0 otherwise,
- x_{ijkm} is the fraction of traffic from origin i to destination j that is routed via hubs k and m.

Parameters χ and δ denote the unit transportation cost for collection and distribution respectively, while factor α represents the reduced costs per unit between hubs due to increased traffic on the hub arcs.

Using the above notation, the UMAHLP can be modeled as a Mixed-Integer Linear Program MILP, as in [6]:

$$min \sum_{i,j,k,m} W_{ij} \cdot (\chi \cdot C_{ik} + \alpha \cdot C_{km} + \delta \cdot C_{mj}) \cdot x_{ijkm} + \sum_{k} f_k \cdot y_k \quad (1)$$

subject to:

$$\sum_{k,m} x_{ijkm} = 1 \quad for\ every\ i, j \quad (2)$$

$$\sum_{m} x_{ijkm} + \sum_{m,m \neq k} x_{ijmk} \leq y_k \quad for\ every\ i, j, k \quad (3)$$

$$y_k \in \{0, 1\} \quad for\ every\ k \quad (4)$$

$$x_{ijkm} \geq 0 \quad for\ every\ i, j, k, m. \tag{5}$$

The objective in (1) is to minimize the total cost, which is the sum of the transportation cost and the fixed cost. Constraints (2) guarantee that the total traffic for any origin-destination pair (i, j) is routed via some pair of hubs (k, m). Constraints (3) assure that traffic is routed only via opened hub locations. Variables y_k and x_{ijkm} are assumed to be binary and non-negative by (4) and (5) respectively.

The UMAHLP is NP-complete, with exception of some special cases that can be solved in poynomial time (e.g. when flow matrix is sparse). If the set of hubs is pre-determined, the Shorthest path Algorithm can provide solution for the UMAHLP in $O(n^3)$ computational time [9].

Several approaches for solving the UMAHLP are proposed in the literature. The dual-ascent technique within a branch-and-bound scheme in [12] was tested on ORLIB hub data set [3] with $n \leq 25$ nodes. Mayer and Wagner [14] developed a new branch-and-bound method that provides solutions for ORLIB instances with $n \leq 40$. Boland et al. in [4] developped preprocessing procedures and tightening constraints for the UMAHLP formuation and presentedl results on ORLIB data set with $n \leq 50$.

Canovás et al. proposed a heuristic based on a dual-ascent technique that was later implemented within a branch-and-bound algorithm in [7]. Through computational analysis using ORLIB CAB and AP data sets they were able to solve instances up to 120 nodes. These are the best computational results for the UMAHLP from the literature up to now.

3 Hybrid Genetic Algorithm for the UMAHLP

It is well known that classical genetic algorithms are usually less efficient in fine-tuning solutions in complex search-spaces [15]. For many hard combinatorial optimization problems combinations of genetic algorithms and local improvement techniques have been applied with success: [1], [16] ,[18], [8] and [17]. In the hybrid genetic algorithm (HGA), proposed in this paper, we have implemented local search heuristic within the GA framework. Local search is applied on the best individual in every generation of the genetic algorithm, if it has been changed comparing to the previous generation. The heuristic tries to interchange one opened and one closed facility location. The interchange process is performed when the first improvement is obtained. This step is repeated on the new, improved individual until it remains unchanged in two successive steps. The heuristic is applied before the selection, crossover and mutation operators. The basic scheme of the HGA implementation can be represented as:

```
Input_Data();
Population_Init();
```

```
while not Finish() do
  for i:=1 to Npop do
    obj[i] := Objective_Function(i);
    Improving_Heuristic (obj[i]);
  endfor
  Fitness_Function();
  Selection();
  Crossover();
  Mutation();
endwhile
Output_Data();
```

/* *Npop* denotes the number of individuals in a population and $obj[i]$ is objective value of i-th individual */

3.1 Encoding and Objective Function

Individuals in the HGA implementations for the UMAHMP encoded as binary strings of length n. One in the genetic code denotes that the current node is chosen as hub, zero if not. For example, from the genetic code 000100110111 for $n = 12$ we understand that hubs are located at nodes 3, 6, 7, 9, 10 and 11 (numbering is from 0 to $n - 1$).

The indices of established hubs are obtained from the genetic code. When the set of hubs is fixed, the problems of allocating non-hub nodes to hubs reduce to solving n^2 shortest paths problems, which can be done in $O(n^2p)$ time, where p is the number of located hubs. We use a modification of Floyd-Warshall algorithm, described in [9], to find a shortest path for each pair of nodes in the network. After assigning non-hub nodes to hubs, the objective value is then simply evaluated by summing distances origin-hub, hub-hub and hub-destination multiplied with corresponding flows and parameters χ, α and δ respectively and adding fixed costs for establishing hubs.

3.2 Genetic Operators and Other Aspects of the GA

As a selection method, we implemented the Fine Grained Tournament Selection (FGTS). Instead of having an integer parameter N_{tour}, which represents the tournament size in the classic tournament selection, the new operator's input is a real parameter F_{tour} representing an average tournament size [11]. The size of each of the tournaments is chosen so that this value is on average as close as possible to F_{tour}.

To recombine two individuals, we apply standard one-point crossover with the probability p_{cross}. A bit position i is randomly chosen in the genetic code (crossover point). Whole genes are exchanged after the chosen crossover point.

Mutating a randomly selected gene (0 to 1, 1 to 0) with certain mutation rate p_{mut} may be unnecessarily low, since usually few genes mutate, while the majority of them remains the same. After certain number of generations, it may happen that all individuals have the same bit value in a certain position in a gene (frozen bits), which may increase the possibility of premature convergence significantly. Therefore, frozen bits are mutated with 2.5 times higher rate comparing to basic mutation rate p_{mut}

We used traditional generational GA (containing N_{pop} individuals) with overlapping populations: N_{elite} is the number of elitist individuals that survive to the next generation. Since the evaluation of the objective value is a time-consuming operation, we store a certain amount of already calculated values in a cache table of size N_{cache}. Before the objective value for a certain individual is calculated, the table is checked. The Least Recently Used caching strategy, which is simple but effective, is used for that purpose [13].

In order to avoid premature convergence, multiple individuals are discarded from the population. If the individual with the same genetic code appears again in the population, its objective value is set to zero and the selection operator disables it to enter the next generation. The appearance of individuals with the same objective value, but different genetic codes is limited to a constant N_{rv}. This strategy helps in preserving the diversity of genetic material and in keeping the algorithm away from a local optima trap

4 Computational Results

The proposed hybrid genetic algorithm is tested on an AMD $K7/1.33$ GHz processor with 256 MB RAM memory. The code is written using C under Linux operation system. Computational experiments are carried out on the standard ORLIB AP data set, which is derived from a study of postal delivery system in Australia. The largest AP instance corresponds to 200 real-world postcode districts from the Australian Post and smaller problems are derived from it by aggregating the data. Fixed costs (loose -L- and tight -T) are included, according to [10]. The transportation cost parameters χ, α and δ take the same values: $\chi = \delta = 1$ and $\alpha = 0.1, 0.5, 0.75, 0.9$, as in [7].

The following setup of GA parameters was used for the HGA, as it proved to be robust in computational experiments: population size $N_{pop} = 150$, number of elitist individuals $N_{elite} = 100$; size of the cache $N_{cache} = 5000$; average group size for the fine grained tournament selection $F_{tour} = 5.4$, crossover probability $p_{cross} = 0.85$ and mutation parameter $p_{mut} = 0.1$. On each instance the proposed HGA was run 20 times. Each run was terminated after 1000 iterations or after the best individual was repeated 500 times. On all the instances we considered, this criterion allowed the HGA to converge so that only minor or no improvements in the quality of final solutions can be expected when prolonging the runs.

inst	Opt.sol	Best.sol.	t(s)	t_{tot} (s)	gen	agap(%)	σ(%)	eval	cache(%)
10L	221 032.734	opt	0.003	0.113	503	0.000	0.000	664	97.4
10T	257 558.086	opt	0.001	0.114	501	0.000	0.000	709	97.2
20L	230 385.454	opt	0.007	0.206	504	0.000	0.000	2547	89.9
20T	266 877.485	opt	0.010	0.204	506	0.000	0.000	2585	89.8
25L	232 406.746	opt	0.015	0.313	505	0.000	0.000	3401	86.6
25T	292 032.080	opt	0.014	0.295	506	0.000	0.000	3483	86.3
40L	237 114.749	opt	0.065	0.833	517	0.000	0.000	5302	79.6
40T	293 164.836	opt	0.017	0.792	501	0.000	0.000	5217	79.3
50L	233 905.303	opt	0.072	1.434	510	0.000	0.000	6650	74.1
50T	296 024.896	opt	0.072	1.339	512	0.000	0.000	6626	74.3
60L	225 042.310	opt	0.075	2.149	506	0.000	0.000	7248	71.5
60T	243 416.450	opt	0.130	2.417	516	0.000	0.000	7568	70.8
70L	229 874.500	opt	0.309	3.691	531	0.000	0.000	8980	66.3
70T	249 602.845	opt	0.152	3.629	513	0.000	0.000	8100	68.6
80L	225 166.	opt	0.809	5.119	565	0.000	0.000	9613	66.1
80T	268 209.406	opt	0.515	4.992	539	0.000	0.000	9488	65.0
90L	226 857.465	opt	0.368	6.693	518	0.000	0.000	10266	60.5
90T	277 417.972	opt	0.424	6.619	522	0.000	0.000	10017	61.9
100L	235 097.228	opt	1.205	8.381	561	0.000	0.000	10930	61.2
100T	305 097.949	opt	0.155	7.946	505	0.000	0.000	9746	61.6
110L	218 661.965	opt	0.557	9.695	517	0.000	0.000	10022	61.5
110T	223 891.822	opt	1.103	10.731	539	0.000	0.000	10877	59.8
120L	222 238.922	opt	0.885	12.609		524 0.000	0.000	10443	60.4
120T	229 581.755	opt	2.343	15.077	564	0.000	0.000	12188	57.0
130L	-	223 814.109	3.117	21.566	563	0.000	0.000	12198	56.9
130T	-	230 865.451	2.789	22.765	552	0.000	0.000	12651	54.4
200L	-	230 204.343	25.202	81.456	667	0.696	1.239	16374	51.2
200T	-	268 787.633	28.688	93.926	701	0.000	0.000	18778	46.5

Table 1. HGA results on AP instances with $\chi = 3, \delta = 2$ and $\alpha = 0.75$

A detailed report of computational experiments for HGA on AP problem instances is too large for this presentation. Therefore, in this paper in Table1 we present only computational results on AP instances with $n \leq 200$ nodes and parameter values $\chi = 3, \delta = 2$ and $\alpha = 0.75$. The exhaustive computational study on AP data set can be found on the website $http : //www.matf.bg.ac.yu/p/files/1269571675 - 10 - results.pdf$.

In the first column of Table1 instance's dimension is given, with mark "T" or "L", denoting "tight" or "loose" fixed costs respectively. The next column contains the optimal solution of the current instance *Opt.sol*, if it is previously known. If it is not, a dash "-" is written. The best value of HGA is given in the following column *Best.sol* with mark *opt* in cases where HGA reached the optimal solution known in advance. The average time needed to detect the best value is given in t column, while t_{tot} represents the total time

(in seconds) needed for finishing HGA. On average, HGA has finished after *gen* generations. The solution quality in all 20 executions ($i = 1, 2, ..., 20$) is evaluated as a percentage *agap* and standard deviation of the average gap σ. The average number of evaluations is given in the *eval* coloumn, while *cache* displays savings (in percent) achieved by using caching technique.

Comprehensive computational experiments on AP problem instances demonstrate the robustness of the proposed HGA with respect to the solution quality and running times. Implemented local search successfully leads the algorithm to the promising regions of the search space. Local search technique is employed only when it is necessary, so the running time of the HGA is relatively short.

From the presented results it can be seen that for all AP instances with up to 120 nodes the HGA reaches optimal solutions known in advance. The CPU time the HGA needed to detect the best (optimal) solution for the first time is less than 3.5 seconds. Total running time, until a finishing criterion is satisfied, is 12.866 seconds maximum.

Computational results on considered large scale AP instances show the appropriateness of applying proposed hybrid algorithm components. On four AP instances with 130 and 200 nodes, unsolved to optimality so far, the HGA also provides solutions in relatively short CPU time: $t \leq 28.688$ and $t_{tot} \leq 93.926$ seconds. Although the optimality can not be proved, we believe that the obtained solutions are of high-quality.

5 Conclusions

In this paper, we present a robust hybrid heuristics, named HGA, based on a genetic search framework for solving the UMAHLP. Binary representation of individuals, FGTS selection and one-point crossover are used in the HGA. Mutation with frozen genes increases the diversibility of genetic material and keeps the algorithm away from a local optima trap. Solution quality is improved by local search heuristic that is efficiently implemented in HGA. Caching technique additionally improves performance of the algorithm. Comprehensive computational experiments on ORLIB AP hub instances clearly demonstrate the robustness of the proposed HGA with respect to the solution quality and running times. Computational results show that the HGA outperforms other existing algorithms for this problem and also provides solutions for large-scale AP instances for which no optimal solution is presented in the literature up to now.

References

1. Abdinnour-Helm S., "A hybrid heuristic for the uncapacitated hub location problems",*European Journal of Operational Research*, 106, 489–499 (1998).

2. Alumur, S., and Kara, B. Y., "Network hub location problems: the state of the art", *European Journal of Operational Research*, 190, 1–21 (2008).
3. Beasley, J.E., "Obtaining Test Problems via Internet", *Journal of Global Optimization*, 8, 429–433 (1996). http://msmga.ms.ic.ac.uk/jeb/orlib/info.html
4. Boland, N., Krishnamoorthy, M., Ernst, A. T., and Ebery, J., "Preprocessing and Cutting for Multiple Allocation Hub Location Problems", *European Journal of Operational Research*, 155, 638–653 (2004)
5. Campbell, J. F., Ernst, A. T., and Krishnamoorthy, M., "Hub location problems", *Location Theory: Applications and Theory*, H. Hamacher and Z. Drezner, Eds, Springer-Verlag, 373–406 (2002).
6. Cánovas, L., Landete, M. and Marín, A., "Improved formulations for the uncapacitated multiple allocation hub location problem", 172(1), 274–292 (2006).
7. Cánovas, L., Garcia, S., and Marin, A., "Solving the Uncapacitated Multiple Allocation Hub Location Problem by Means of a Dual-ascent Technique", *European Journal of Operational Research*, 179, 990–1007 (2007).
8. Cunha, C.B., and Silva, M.R., " A genetic algorithm for the problem of configuring a hub-and-spoke network for a LTL trucking company in Brazil", *European Journal of Operational Research*, 179, 747-758 (2007).
9. Ernst, A. T., and Krishnamoorthy, M., "An Exact Solution Approach Based on Shortest-paths for p-hub Median Problem", *INFORMS Journal of Computing*, 10, 149–162 (1998).
10. Ernst, A.T., and Krishnamoorthy, M., "Solution algorithms for the capacitated single allocation hub location problem", *Annals of Operational Research*, 86, 141–159 (1999)
11. Filipović, V., "Fine-Grained Tournament Selection Operator in Genetic Algorithms", *Computing and Informatic*, 22(2), 143–161 (2003).
12. Klincewitz, J. G., "A Dual Algorithm for the Uncapacitated Hub Location Problem", *Location Science*, 4(3), 173–184 (1996).
13. Kratica J., "Improving Performances of the Genetic Algorithm by Caching", *Computers and Artificial Inteligence*, 18 (3), 271–283 (1999).
14. Mayer, G., and Wagner, B., "An exact solution method for the multiple allocation hub location problem", *Computers and Operational Research*, 29, 715–739 (2002)
15. Michalewicz, Z., *Genetic Algorithms + Data Structures = Evolution Programs*, Third Edition: Springer Verlag, Berlin Heideleberg (1996).
16. Ognjanović, Z., Kratica, J., and Milovanović, M., " A genetic algorithm for satisfiability problem in a probabilistic logic: A first report", *Lecture Notes in Artificial Intelligence - LNAI*, 2143, 805–816 (2001).
17. Stanimirović, Z., Kratica, J., and Dugošija, Dj., "Genetic algorithms for solving the discrete ordered median problem", *European Journal of Operational Research*, 182(3), 983–1001 (2007).
18. Topcuoglu, H., Corut, F., Ermis, M., and Yilmaz, G., " Solving the uncapacitated hub location problem using genetic algorithms", *Computers and OR*, 32 (4), 967-984.

A Wavelet Based Prediction Method for Time Series

Cristina Stolojescu[1,2] Ion Railean[1,3] Sorin Moga[1] Philippe Lenca[1] and Alexandru Isar[2]

[1] Institut TELECOM; TELECOM Bretagne, UMR CNRS 3192
Lab-STICC; Université europèenne de Bretagne, France
(e-mail: firstname.lastname@telecom-bretagne.eu)
[2] Politehnica University of Timisoara, Romania
Faculty of Electronics and Telecommunications
(e-mail: firstname.lastname@etc.upt.ro)
[3] Technical University of Cluj-Napoca, Romania
Faculty of Electronics, Telecommunications and Information Technology

Abstract. The paper proposes a wavelet-based forecasting method for time series. We used the multi-resolution decomposition of the signal implemented using trous wavelet transform. We combined the Stationary Wavelet Transform (SWT) with four prediction methodologies: Artificial Neural Networks, ARIMA, Linear regression and Random walk. These techniques were applied to two types of real data series: WiMAX network traffic and financial. We proved that the best results are obtained using ANN combined with the wavelet transform. Also, we compared the results using various types of mother wavelets. It is shown that Haar and Reverse biorthogonal 1 give the best results.
Keywords: time series, Stationary Wavelet Transform, forecasting.

1 Introduction

Forecasting, or prediction, is the process of estimation in unknown situations, based on the analysis of some factors that are believed to influence the future values, or based on the study of the past data behavior over time, in order to take decisions. Time-series forecasting is an important area of forecasting where the historical values are collected and analyzed in order to develop a model describing the behavior of the series. When the time series is non-stationary, it is very difficult to identify a proper global model, [3]. To overcome this problem, an efficient way is to use the wavelet decomposition technique in the preprocessing step. The Wavelet transform (WT) provides a useful decomposition of time series, in terms of both time and frequency, permitting us to effectively diagnose the main frequency component and to extract abstract local information from the time series.

WT has been frequently used for time series analysis and forecasting in the recent years, [1,2]. Models that accurately catch the statistical characteristics of the actual traffic play a significant role in studying the network, in understanding its dynamics, in designing and controlling the network. For

financial time series prediction, sales forecasts are very useful in the economic domain because they are used to optimize inventory levels. Several models have been proposed for time-series forecasting such as pure statistical or based on Artificial Neural Networks (ANN). Traditional linear time series models including ARIMA (Auto Regressive Integrated Moving Average) model proved to be good at capturing the behavior of the time series. To deal with the non-linear nature of time-series, the ANN model is probably the most popular method. It can capture any kind of relationship between the output and the input theoretically.

In this paper, we analyze the influence of different mother wavelets on the performance of forecasting. We compared the results trying to find out which is the best of the mother wavelets to be applied and, using this wavelet, which method gives the best forecasts. The rest of the paper is organized as follows: in Section 2 we present some theoretical considerations regarding WT and multi-resolution analysis. In Section 3 we describe the forecasting framework. The experimental results are presented in the fourth Section and finally, Section 5 is dedicated to the conclusions.

2 The wavelet analysis

As stated before one of our goals is to compare the forecasting accuracy by using the wavelet transform in the preprocessing step. The transform of a signal is just another form of representing it, which does not change the information content present in the signal. A linear time-frequency transform correlates the signal with a family of waveforms that are well concentrated in time and in frequency. Multi-resolution analysis (MRA) is a signal processing technique that takes into account the signal's representation at multiple time resolutions. Using wavelet MRA, the collected measurements can be smoothed until the overall long-term trend is identified. Fluctuations around the obtained trend are further analyzed at multiple time scales. The level of decomposition depends on the length of the data set (the number of values). At each temporal resolution two categories of coefficients are obtained: approximation coefficients and detail coefficients. Generally, the MRA is implemented based on Mallat's algorithm [7], which corresponds to the computation of the Discrete Wavelet Transform (DWT). The disadvantage of Mallat's algorithm is the decreasing of the length of the coefficient sequences with the increasing of the iteration index due to the utilization of the decimators. Another way to implement a MRA is the use of the trous methodology, also known as Shensa's algorithm [6], which corresponds to the computation of the Stationary Wavelet Transform (SWT). In this case the utilization of decimators is avoided, but at each iteration different low-pass and high-pass filters are used. There is a variety of mother wavelets [7] such as Daubechies, Symlet, Meyer, Morlet, etc., and the choice of the mother wavelets depends on the characteristics of data. The Daubechies wavelet transforms have been in-

creasingly adopted by signal processing researchers. Haar wavelet transform, which is also the simples Daubechies wavelet is a good choice to detect time localized information. In this work we propose to use some mother wavelets belonging to Daubechies family, but also other orthogonal wavelet families such as Symmlets, also known as the Daubechies least asymmetric mother wavelets, and Coiflets also designed by Ingrid Daubechies to be more symmetrical than the Daubechies mother wavelet, and biorthogonal respective reverse biorthogonal wavelets. Biorthogonal wavelets exhibit the property of linear phase, which is needed for signal reconstruction. If, instead of a single wavelet, two wavelets are used (one for decomposition and the other for reconstruction), interesting properties are derived, [7]. Different types of mother wavelets will be used in the data preprocessing step of our forecasting framework presented in the next Section.

3 Forecasting framework

The main idea of the prediction method using wavelets is to decompose the original signal into a range of frequency scales and then to apply the forecasting methods to these individual components. Our forecasting framework, which belongs to the supervised paradigm, is presented in Figure 1 and implies the following steps:

1. Preprocessing step which includes data clearing, such as identification of the potential errors in data sets, handling missing values, and removal of noises or other unexpected results that could appear during the acquisition process. At this stage the input data is also analyzed in order to find if it contains large spikes and valleys indicating periodicities.
2. Use the SWT to decompose the data separately for the training set and the test set. Each component represents the real data in a frequency range that is easier to predict than the original series. A good predictor should be able to identify the separate scale-related components of the series, in order to produce models that give accurate forecasts. So, our approach is to decompose the original time series into scale or frequency related components and model each component separately, in order to obtain more accurate models.
3. After obtaining the wavelet decomposition, we select the information from each level of decomposition for building the model.
4. In the training phase we design predictive models for each of the decomposed components of the original series. In the test phase the developed forecasting models are used to predict future values for each component. The Inverse SWT is used in the testing phase in order to obtain the forecasted signal from the predictions of the components.

The four models used in this work are presented below:

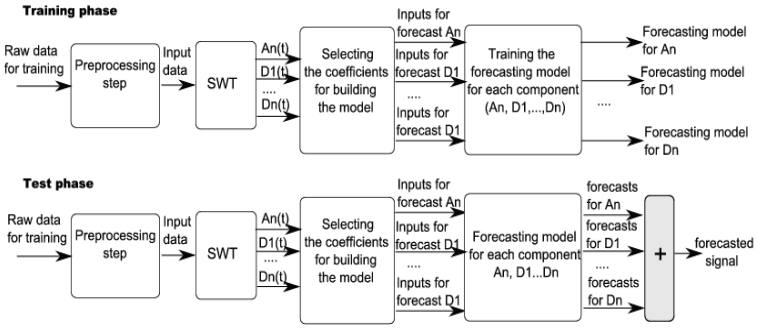


Fig. 1. The forecasting framework.

1. *ANN* models [5] represent a wide class of flexible nonlinear models which have been very used recently in the area of forecasting. The main advantage of an ANN that makes it suitable for various applications is that it learns from the past experiences. So, the basic idea is to train the ANN with past data and then use it to predict future values. Although many types of architectures have been proposed, the most popular one for time series forecasting is the feed-forward neural network [9]. In this work we used a Time-delayed neural networks(TDNN), detailed in [4].
2. *ARIMA* processes [8,10], are the natural generalizations of standard ARMA processes. This class of models is based on Box-Jenkins methodology [10] which is used to build the time series model in a sequence of steps which are repeated until the optimum model is achieved. More details about this method and how it was applied in our case are presented in [11].
3. *Linear regression (LR)* [8] is a simple statistical tool for modeling the output as a linear combination of inputs. The model's parameters are usually estimated using the least-squares method.
4. *Random walk (RW)* method [8] is based on the hypothesis that from one period to the next, the original time series takes a random "step" away from its last recorded position. The prediction of the future values is based on the previous values plus a constant that represents the average change between the two periods.

4 Experiments

4.1 Data sets

In this work we used historical data obtained by monitoring the traffic from 67 Base Stations (BS) composing a WiMAX network. The period of collection is of eight weeks, from March 17th till May 11th, 2008. Each BS has its own data set which is composed of numerical values representing the total number of packets from the uplink channel. Each value is recorded

every 15 minutes. It can be easily deduced that for a given BS we have the following number of samples: 96 samples/day, 672 samples/week, and a total number of 5376 samples. So, the WiMAX data base can be seen as formed by 67 matrices (one for every BS) that have eight columns (the number of weeks) and 672 lines (the moments of time when the number of packets are recorded in a week).We also used one time series of financial data representing the total number of EUR-USD currency exchanges (the volume of data is similar to the number of packets from WiMAX. The period of collection is of fifteen weeks and the values are recorded every 15 minutes. We will have 96 samples/day, 672 samples/week and a total number of 10080 samples. In this case only one matrix will correspond to each of the two sets and it will be formed by fifteen columns (the number of weeks) and 672 lines. The objective of our work is to compare the influence of different mother wavelets used in the preprocessing step on the prediction accuracy. Also, using the best mother wavelets, we propose to evaluate some prediction models, such as pure statistical or based on neural networks.

4.2 Evaluation criteria

In order to evaluate the performance of prediction using different types of wavelets, we considered the most used statistical measures of error: the Mean absolute error (MAE), the Mean Square Error (MSE), the analysis of variance (ANOVA), the Symmetric Mean Absolute Percent Error (SMAPE), and the Root Mean Square Error (RMSE). We have also calculated SMAPE L, MAPE L and MAE L, between the mean of the original signal and the mean of the forecasted signal, because ARIMA and LR cannot be used to obtain forecasts for every moment of time as ANN and RW can. For linear models the trajectory of the forecasts is represented through sloping line which represents the weekly increase.

4.3 Results and discussions

Regarding the WT, we propose various types of mother wavelets such as Daubechies (db), Coiflet (coif), Symlet (sym), Biorthogonal (bior), and Reverse Biorthogonal (rbio). In Table 1 and Table 2, for each type of mother wavelets and every type of error, excepting SMAPE L, MAPE L and MAE L, we present the average value corresponding to the three types of ANN and RW. SMAPE L, MAPE L and MAE L correspond to all the proposed methods. We do not take into consideration the results given by using the linear regression because the wavelet transform does not have any influence on the predictions. In this case the mean value of the details is zero, and the prediction obtained for the details will be also zero. In the case of WiMAX traffic (Table 1), the results are represented as the average value for all 67 BS. According to Table 1, the wavelet of Haar (db1), which is the simplest of the Daubechies family and rbio1.1 give the best prediction performance. The

results also indicate that with the increase of the filters' length (support of the mother wavelets), the performance of the wavelet transform deteriorates. The results represent the mean values for all the forecasting methods and all the 67 BSs with the observation that in the case of ARIMA only SMAPE L, MAPE L and MAE L could be calculated.

Wavelet	RSQ	SMAPE	MAPE	MSE	RMSE	MAE	SMAPE L	MAPE L	MAE L
coif 1	1.445	1.09	**0.2113**	11.72	2.80	1.0304	0.890	0.0020	0.9599
coif 2	1.493	1.22	0.2285	12.95	2.83	0.8748	0.837	0.0019	**0.7191**
db 1	**1.168**	**1.08**	0.2367	**8.06**	**2.43**	**0.7685**	**0.812**	**0.0016**	0.7327
db 2	1.364	1.15	0.2451	10.52	2.69	0.8408	0.855	0.0019	0.7768
db 3	1.358	1.12	**0.2117**	**9.76**	**2.64**	0.8193	0.857	0.0018	0.7678
db 4	1.490	1.11	0.2159	10.61	2.58	0.7985	0.834	0.0018	0.7563
db 5	1.435	1.11	0.2190	12.56	2.75	0.8339	0.823	0.0019	0.7730
bior 3.1	0.695	1.13	0.3152	9.86	2.52	84020	0.860	0.0018	**0.7071**
rbio 1.1	**1.200**	**1.08**	0.2215	10.00	2.61	**0.7948**	**0.820**	**0.0017**	0.8947
rbio 2.2	1.482	1.19	0.3202	10.29	2.71	0.8747	0.891	0.0018	0.7690
rbio 3.3	1.952	1.21	0.2623	10.33	2.88	0.9509	0.907	0.0022	1.0690
sym 2	1.365	1.26	0.2146	13.20	2.89	0.8854	0.895	0.0019	0.7412

Table 1. Comparison between wavelets, WiMAX traffic.

In the case of financial data, for the set containing the EUR-USD exchange currency, the results are shown in Table 2. We can observe that the best forecasting performance is obtained using the mother wavelets coif2 and sym2.

For the purpose of forecasting methods comparison, we propose the following variants: three types of methods based on ANN (ANN No Sliding, ANN Known Sliding, and ANN UnKnown Sliding), ARIMA, LR, and RW model for two weeks prediction. The first method using ANNs, ANN No Sliding, is the simplest one: we train the ANN once for each decomposition level. For inputs, we have the first (n-2k) weeks, where n is the total number of weeks, and k is the number of weeks we want to forecast. The target

Wavelet	RSQ	SMAPE	MAPE	MSE	RMSE	MAE	SMAPE L	MAPE L	MAE L
coif 1	0.688	0.556	0.1152	1.627	1.220	0.3586	0.516	0.0821	0.4240
coif 2	0.455	0.522	**0.0793**	**1.089**	**1.038**	**0.2967**	**0.453**	0.0732	0.3799
db 1	0.625	0.520	**0.0839**	1.356	1.126	0.3175	0.454	**0.0713**	**0.3690**
db 2	0.715	0.578	0.1088	1.610	1.219	0.3550	0.497	0.0812	0.4199
db 3	0.586	0.585	0.1156	1.499	1.188	0.3618	0.531	0.0864	0.4461
db 4	**0.871**	0.600	0.1114	1.604	1.239	0.3745	0.527	0.0863	0.4459
db 5	**0.808**	0.587	0.1121	1.557	1.225	0.3700	0.546	0.0912	0.4710
bior 3.1	0.628	0.534	0.1137	1.552	1.173	0.3390	0.433	**0.0712**	**0.3681**
rbio 1.1	0.615	**0.519**	0.0958	1.286	1.096	0.3208	0.457	0.0716	0.3704
rbio 2.2	0.541	0.555	0.1087	1.440	1.149	0.3406	0.491	0.0790	0.4081
rbio 3.3	0.772	0.595	0.0993	1.418	1.167	0.3480	0.455	0.0716	0.3705
sym 2	0.476	**0.499**	0.0890	**1.117**	**1.037**	**0.2962**	**0.453**	0.0736	0.3813

Table 2. Comparison between wavelets, EUR-USD currency exchanges.

consists of the data taken from the weeks (n-2k+1) to (n-k). The data used for ANN's inputs during the testing phase is the information from the weeks (k+1) to (n-k). The output signal is compared to the real data of the last k weeks. The next method (ANN Known Sliding) uses sliding, retraining the network with the real information. The entire signal is divided into smaller parts. Each of these sequences will predict a small part of the final forecasted signal. The information for ANNs retraining is always taken from the real data. The last method, ANN UnKnown Sliding, proposes a forecasting using sliding with unknown data. The only difference consists in the fact that the information used for the next simulation and retraining is taken not from the original signal, but from the previously predicted one. For more details see [4]. The use of ARIMA is detailed in [11].

In the case of WiMAX traffic, the comparison was made using db1 mother wavelets. The results presented in Table 3 prove that ANN performs better than the other prediction techniques. Also the linear regression model gives very good forecasting results.

Forecasting Model	SMAPE L	MAPE L	MAE L
ANN No Sliding	**0.472**	0.0011	0.4428
ANN Known Sliding	0.509	**0.0009**	0.4241
ANN UnKnown Sliding	0.722	0.0017	0.6681
ARIMA	0.772	0.0027	0.9990
Linear Regression	0.523	0.0031	**0.3868**
Random Walk using Wavelets	4.440	0.0030	1.3633

Table 3. Forecasting techniques comparison for WiMAX traffic.

For the financial data base, we used the coif2 mother wavelets. The results are presented in Table 4. We found that the suitable model is as well the one using ANNs.

Forecasting Model	SMAPE L	MAPE L	MAE L
ANN No Sliding	0.169	0.020	0.1079
ANN Known Sliding	**0.153**	**0.0178**	**0.957**
ANN UnKnown Sliding	0.267	0.0344	0.1812
ARIMA	1.135	0.3245	1.7054
Linear Regression	0.191	0.0243	0.1109
Random Walk using Wavelets	0.940	0.2670	1.4173

Table 4. Forecasting techniques comparison for financial data.

5 Conclusion

Regarding the Wavelet transform, our results show that Haar, which is the simplest of Daubechies family, and Reverse biorthogonal 1 improve the performance of the prediction technique.

An important conclusion is that as much the support of the mother wavelets increases, the performance of the wavelet transform deteriorates. In addition, using the best mother wavelets in data preprocessing step, we proved that ANN outperforms the other forecasting methods. Also, our results confirms the results in [Papagiannaki, et al, 2005] and point out that if we are interested in tendency prediction, for more than one month ahead, than linear models are suitable for this type of forecasting. We should also point out that we have applied our algorithm on two different data sets which are not comparable. The financial data (the EUR-USD currency exchanges) exhibit an almost constant tendency, while WiMAX traffic presents a strong variability and its tendency (long term trend) represents a sloping line. However, our algorithm is applicable to both types of data and the obtained predictions are accurate. As a future work we propose to apply our algorithm on other time series, for example transportation data, including highway traffic, aircraft flights, traffic data of cars in tunnels, traffic at automatic payment systems on highways, traffic of individuals on subway systems, etc.

References

1. X. Wang, X. Shan, A wavelet-based method to predict Internet traffic, in *Communications, Circuits and Systems and West Sino Expositions*, vol.1, pp. 690-694, (2002).
2. K. Papagiannaki, et al, Long-term forecasting of Internet backbone traffic, in *IEEE* Transactions on Neural Networks, vol.16, pp. 1110-1124,(2005).
3. Zhang et al, Multiresolution Forecasting for Futures Trading Using Wavelet Decompositions,in *IEEE* Transactions on neural networks, vol. 12, no. 4, (2001).
4. I.Railean et al, WIMAX Traffic Forecasting based on Neural Networks in Wavelet Domain, submitted to RCIS 2010 (2010).
5. P. Mehra and B.W.Wah, Artificial Neural Networks: Concepts and Theory in *IEEE* Computer Society Press Tutorial, Los Alamitos, CA, (1992).
6. M.J.Shensa, *Discrete Wavelet Transform. Wedding the a trous and Mallat algorithms, IEEE* Transactions and Signal Processing, 40, pp. 2464-2482,(1992).
7. S. Mallat, *A Wavelet Tour of Signal Processing*, Second Edition, (1999).
8. B. Abraham and J. Ledolter, Statistical Methods for forecasting, in *Wiley Series in Probability and Mathematical Statistics*, (1983).
9. G. P. Zhang, M. Qi, Neural network forecasting for seasonal and trend time series, in the *European Journal of Operational Research* 160, pp. 501-514,(2005).
10. G. Box, G. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, CA, (1970).
11. C. Stolojescu et al, Forecasting WiMAX BS Traffic by Statistical Processing in the Wavelet Domain, in *Proceedings of the International Symposium on Signals, Circuits and Systems*, Iasi, Romania, pp. 177-183, (2009).

FUZZY MARKOV SYSTEMS FOR THE DESCRIPTION OF OCCUPATIONAL CHOICES IN GREECE

M. A. Symeonaki[1] and R. N. Filopoulou[2],
Department of Social Policy,
Panteion University of Social and Political Sciences,
136 Syggrou Av., 176 71, Athens, Greece.

Abstract

In this paper the theory of fuzzy non homogeneous Markov systems is applied for the first time to the study of occupational choices in Greece. This is an effort to deal with the uncertainty introduced in the estimation of the transition probabilities and the difficulty of estimating their real values. In the case of studying the occupational choices of children, the traditional methods for estimating the probabilities can not be used due to lack of data. The introduction of fuzzy logic into Markov systems provides us with a powerful tool, taking advantage of the heuristic knowledge that the experts of the system posses. The proposed model uses the symbolic knowledge of the occupational choices of children and focuses on the important factors which derive from the family environment and affect those choices. The aim is to develop a fuzzy expert system which best simulates the real conditions affecting the process of occupational choices in Greece.

Keywords: occupational choices, family factors, Markov systems, Fuzzy logic, Fuzzy Inference System

1. Introduction

The present paper is concerned with constructing a model that reflects the occupational choices of children in Greece using fuzzy Markov systems and symbolic knowledge. The purpose of the paper is to take into account the theoretical discussion and the empirical facts that exist about occupational choices and create a fuzzy expert system which provides essential information about this social process.

[1] msimeon@panteion.gr

[2] celestfilopoulou@msn.com

The occupational choices people make has been a rather challenging area for scientists over the past years (Parsons [1909], Ginsberg [1951], Super [1951], Hoppock [1976], Holland [1976], Parsons [1909], Ginsberg [1951], Super [1951], Hoppock [1976], Holland [1976]). Occupation has gradually attracted scientific interest as a measure which can provide useful information about people concerning their social, economical and cultural state. Additionally, the use of occupation as a means of categorizing people with similar characteristics avoids the theoretical and ideological friction that such classifications bring about when they are based merely on the concept of class (Kasimati [2004]).

Most scientists agree that people conclude to a certain occupation through a process of decision-making that starts from the early years of their life and ends when their professional development is completed. Furthermore, this process appears to be influenced by different factors coming from individual characteristics, as well as from the broader environmental surroundings of the individual. In the first case special attention is given to the personality, the interests, the preferences and the abilities of people which lead them to express their orientation to a certain occupation or a group of similar occupations (Parsons [1909], Ginsberg [1951], Super [1951], Hoppock [1976], Holland [1976]). Apart from these basically psychological theories a number of theories concerning non-individual factors were also developed. These theories focused on the interaction between the existing social, economical and cultural environment and the individual (Lipset [1962], Blau et al [1967]). We could argue that the occupational choices are mainly a sum of both individual and non-individual factors as, on the one hand, they directly or indirectly affect one's decisions and on the other hand the influence cannot always be strictly separated as they coexist and shape each other.

The family arises as a significantly important factor which influences the occupational choices people make. Firstly, the family plays an active role in the development of the personal characteristics of the children. Moreover, it is the family that sets the starting point of a person into society, through its social, economical and cultural state. Especially in Greece, where the family continues to be an important element of the social structure, it appears that factors such as the *education level of parents*, their *occupation*, their *socio-economical status*, the *overall family environment* and the *family aspirations* affect the children's capabilities and achievements at school, their entry into higher education, their occupational orientation, their opportunities and

personal goals and aspirations (Lampiri-Dimaki [1974], Fragoudaki [1985], Kintis [1980], Kontogiannopoulou [1995, 1996], Kasimati [1998, 2001], Viki and Papanis [2007]).

In this paper a new technique of studying occupational choices is proposed for the first time. More specifically, Fuzzy Markov Systems, firstly introduced in Symeonaki et al [2000, 2002], are used in order to deal with common problems arising in the study and the analysis of population systems, especially when they refer to the measurement and the estimation of social phenomena.

The theory of Markov systems (Bartholomew [1982], McClean [1976, 1978, 1980], Bartholomew, Forbes and McClean [1991], Vassiliou [1982]) is very important for the description of population systems. Different applied probability population models can be adapted in this general framework since Markov systems provide one of the most significant tools for describing them. Therefore, Markov systems are applied in a numerous of different domains such as Operation Research, Ecology, Social Policy, etc. One of the basic problems in the theory of Non-Homogenous Markov Systems (NHMS) and one of the main reasons for their impracticability is the uncertainty that is inherent in the estimation of the transition probabilities. The uncertainty due to lack of data and measurement errors is overcome with the use of fuzzy reasoning in the population system which allows us to take into account the symbolic, heuristic knowledge of the experts. In this way it is possible to adopt all the linguistic elements which are essential to the estimation of the transition probabilities. Thus, a Fuzzy Inference System is proposed, which will take advantage of the complex but also rich and important information that exists about occupational choices.

The paper is organized as follows: in Section 2 the need for the use of Fuzzy Logic and Fuzzy Reasoning in the Markov Systems is described and the basic elements of the fuzzy Markov Systems are introduced. In Section 3 a Fuzzy Inference System concerning the occupational choices of the children in Greece is developed, based on the empirical knowledge of the experts of the subject. Finally, in Section 4 the conclusions resulting from this paper are given and the potential future work regarding Fuzzy Markov Systems and the occupational choices are discussed.

2. Fuzzy Markov Systems

It is well known that in the Aristotelian theory something is true or not, i.e. an element belongs to a set or not ($x \in A$, or $x \notin A$). The theory of Fuzzy Logic was introduced by L. A. Zadeh [1965] as an antipode to Aristotelian Logic. Zadeh introduced the concept of *fuzzy sets*, where the participation to a set is expressed by the *membership function* of the fuzzy set. Therefore, what really matters is not whether an element belongs to a set or not, but its degree of membership to the set (for example, one could be POOR with a membership grade equal to 0.3 and simultaneously be AVERAGE with a grade equal to 0.7). There are different kinds of membership functions depending on the fuzzy sets being studied and the specific problem that they are applied to, e.g. triangular, trapezoidal, sigmoidal, etc (Symeonaki, et al [2002]).

As previously mentioned, the contribution of Markov systems to the study of population systems is determinant. However, the estimation of the transition probabilities from one state of the system to another implies the uncertainty of a precise estimation, due to lack of data and measurement errors (Symeonaki [2006]). Therefore, there is a strong need to develop a different method in order to estimate these probabilities more accurately.

Here, we introduce the theory of Fuzzy Logic to the theory of Markov Systems. A fuzzy expert system is developed which contributes to the estimation of the transition probabilities and deals with the uncertainty that they imply. In this way, with the use of Fuzzy Logic and the symbolic knowledge that the experts posses, the gradual transitions realized in the system can be estimated.

A basic feature of the proposed fuzzy system is that it concerns a methodology based on knowledge. Existing knowledge on a given topic is therefore central to the development of such an expert system. This knowledge is reflected by a set of empirical, linguistic rules. Finally, the conclusions are drawn based on these rules and the existing data.

More specifically, there is a number of *population parameters* (Symeonaki et al [2002]) that are used to estimate the transition probabilities. Thus, each transition probability is a function of the population parameters of the system:

$$p_{ij}(t) = f_{ij}(pp_1, pp_2, ..., pp_l) \quad (1)$$

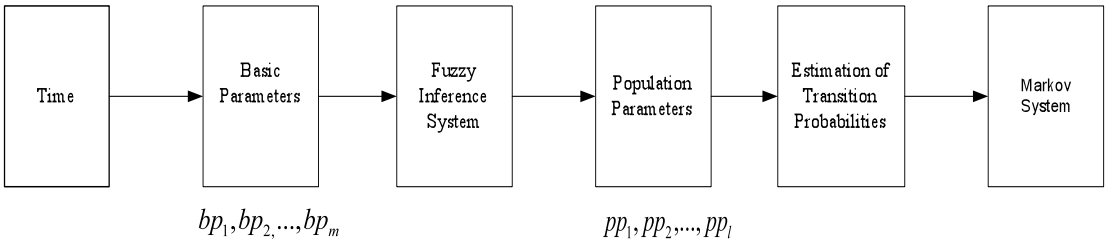
where l the number of the population parameters.

For each value of the parameters it is true that:

$$\sum_{j \in S} f_{ij}(pp_1, pp_2, ..., pp_l) = 1 \quad (2)$$

Furthermore, each population parameter depends on a number of parameters, which are called *basic parameters* of the system (Symeonaki et al [2002]).

We will use a Fuzzy Inference System (FIS) in order to find the population parameters of the system based on the basic parameters of the system. The structure of the FIS is given in Figure 1.



*Symeonaki M. A., Stamou G. B., Tzafestas S. G., *Fuzzy Non-Homogenous Markov Systems*, Applied Intelligence, 17(2), 2002, 203-214.

Figure 1 The structure of the Fuzzy Markov System

Then we define the fuzzy partitions $A^{(i)}$ and $B^{(i)}$ for the basic parameters and the population parameters respectively. These fuzzy partitions have linguistic substance, thus they include linguistic elements, such as “LOW”, “MEDIUM”, “HIGH” etc. The possible fuzzy partitions for the basic parameters bp_i and the population parameters pp_i appear below (FIGURE 2, FIGURE 3).

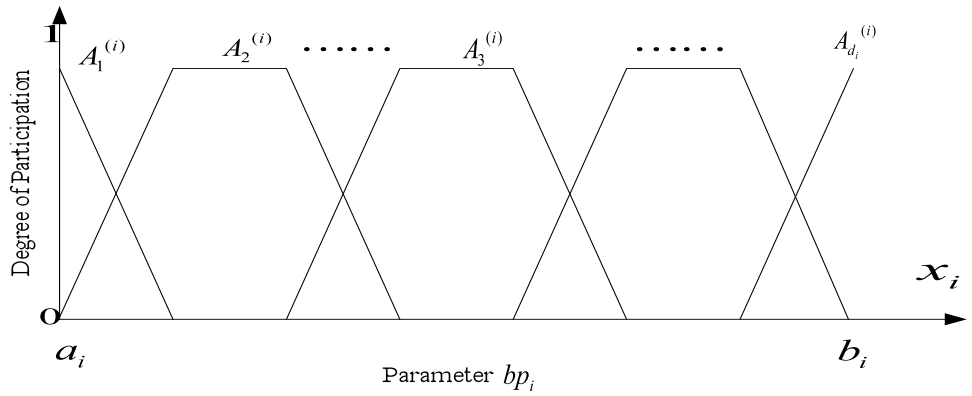


Figure 2 A possible fuzzy partition of the basic parameters bp_i

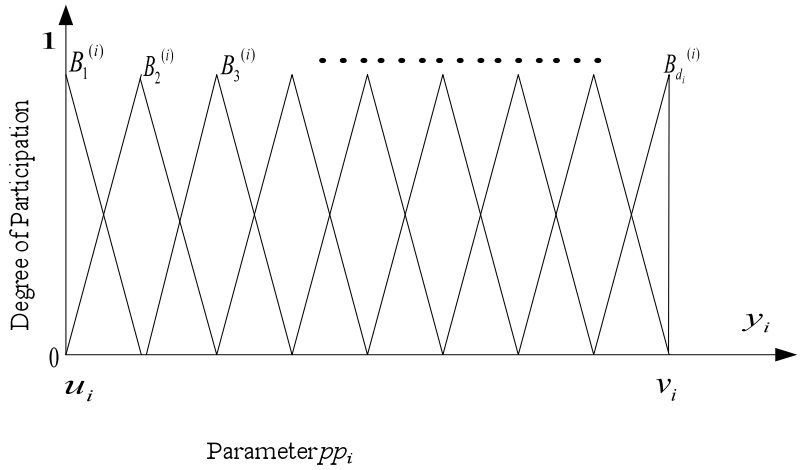


Figure 3 A possible fuzzy partition of the population parameters pp_i

As the elements of $A^{(i)}$ and $B^{(i)}$ have a linguistic form, empirical, verbal rules can be used in order to describe the association that exists between them. Based on the heuristic knowledge that the experts posses around the occupational choices linguistic rules that describe the relations in the system are created.

3. Description of the Fuzzy Inference System for the occupational choices in Greece

Based on the theory of Fuzzy Markov Systems and the knowledge that exists concerning the occupational choices of children in Greece, a Fuzzy Inference System

is created, that indicates all the social processes which are essential to the decisions made.

The present paper is focused on the factors deriving from the family. Thus, it is assumed that the factors (i.e. population parameters) that influence the occupational choices of the children are the following:

pp_1 : family environment,

pp_2 : social state of the parents,

pp_3 : father's occupation (income),

pp_4 : mother's occupation (income).

More specifically, the concept "family environment" refers to the environment in which the individual evolves and acts. The family environment can be a rather determinant factor to the choices made. The educational and cultural experiences, the sense of stability and safety, the level of motivation and the stimulation for personal growth can have a serious impact on this decision-making process.

The "social state of the parents" represents the position they occupy in the society. This position introduces a number of components, such as the educational level, the financial and social status, the social networks and the lifestyle of the parents. It is clear that these factors which are related with the social state of the parents can be inherited to the children and can play a powerful role in their future occupational choices.

The profession exercised by the parents, in terms of income, is indicative of the potential that an individual has. Moreover, children develop economic needs and professional standards similar to those of their parents. Thus, it is common that the occupation of the parents leads the children towards certain types of occupations that can provide the desirable financial earnings.

It is assumed that the above population parameters depend on the basic parameters of the system which are the following:

bp_1 : father's education (in years of education),

bp_2 : mother's education (in years of education) and

bp_3 : parents' aspirations.

Parents' educational level (PL) influences family in different ways. It appears that PL is relevant to all population parameters. However, the third basic parameter (related

with the desires, the demands and the plans that parents have for their children) appears to influence only the existing family environment.

Moreover, the fuzzy partitions $A^{(1)}$, $A^{(2)}$, $A^{(3)}$ in the domain of the three basic parameters are defined respectively. For simplicity reasons, it is assumed that the system under study has the following categories:

1. High demand occupations (Engineering Schools, Medical and Law Schools, etc),
2. Intermediate demand occupations (Earth Sciences, Schools of Mathematics or Physics, Schools of Political or Social Sciences, Primary Education Schools, etc),
3. Non-privileged occupations.

Thus, the following transition probability matrix must be estimated:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

Moreover, the fuzzy partitions $B^{(i)}$, where $i = 1,2,3,...,9$ in the domain of $p_{11}, p_{12}, p_{13}, p_{21}, p_{22}, p_{23}, p_{31}, p_{32}, p_{33}$ are defined.

As the elements of $A^{(i)}$ and $B^{(i)}$ are linguistic we can use empirical verbal rules in order to describe the input and output relation of the system. For example, the empirical knowledge informs us that:

- When the "family environment" is "positive", then the transition probability from the high demand occupations to the non-privileged sectors (p_{13}) is "small".
- When the "social state of the parents" is "high", then the transition probability from the intermediate demand occupations to the high demand occupations is "medium".

In this way, heuristic expert knowledge is concentrated in rules with the above form.

Finally, d_i represents the class of the fuzzy partition $A^{(i)}$, which concerns the cardinality of each partition. Simplifying, we would say that the class d_i of $A^{(i)}$ is the number of the fuzzy subsets which we define in order to partition the domain of $A^{(i)}$ (Klir and Yuan [1995], Stamou and Tzafestas [1999]). Thus, we conclude that

$d_1 = d_2 = d_3 = 3$. The number of all different rules in the system is denoted by k and it is obvious that $k = d_1 \cdot d_2 \cdot d_3 = 27$.

For each population parameter of the system (pp_1, pp_2, pp_3, pp_4) a Fuzzy Inference System is established and the linguistic rules governing the system are formulated. This way, it is possible to estimate the transition probabilities from one state to another based on the heuristic knowledge of the experts. For example, the regulations of the system for the population parameter are given in FIGURE 4.

We denote by w_i the degree in which the rule i fires. Each rule corresponds to a transition matrix $\mathbf{P}_i$ and it can easily be proved by induction that if we use as *t*-norm the product, then:

$$\sum_{i=1}^{27} w_i = 1.$$

Therefore:

$$\mathbf{P} = \sum_{i=1}^{27} w_i \mathbf{P}_i$$

with:

$$\mathbf{P}_i \cdot \mathbf{1}' = \mathbf{1}'$$

where:

$$\mathbf{1}' = \left[1,1,...,1\right]'.$$

If we also assume that the matrices $\mathbf{P}_i$ are irreducible, regular, stochastic matrices, then the asymptotic behaviour of the system could be studied. If $\lim_{t\to\infty} T(t) = T$, then:

$$\lim_{t\to\infty} \mathbf{N}\left(t\right) = \mathbf{N}\left(\infty\right) = T\mathbf{p}^*$$

where $\mathbf{p}^*$ is the row vector of the matrix:

$$\mathbf{P}^* = \lim_{t\to\infty} \mathbf{P}^t = \lim_{t\to\infty} \left(\sum_{i=1}^{27} w_i \mathbf{P}_i \right)^t.$$

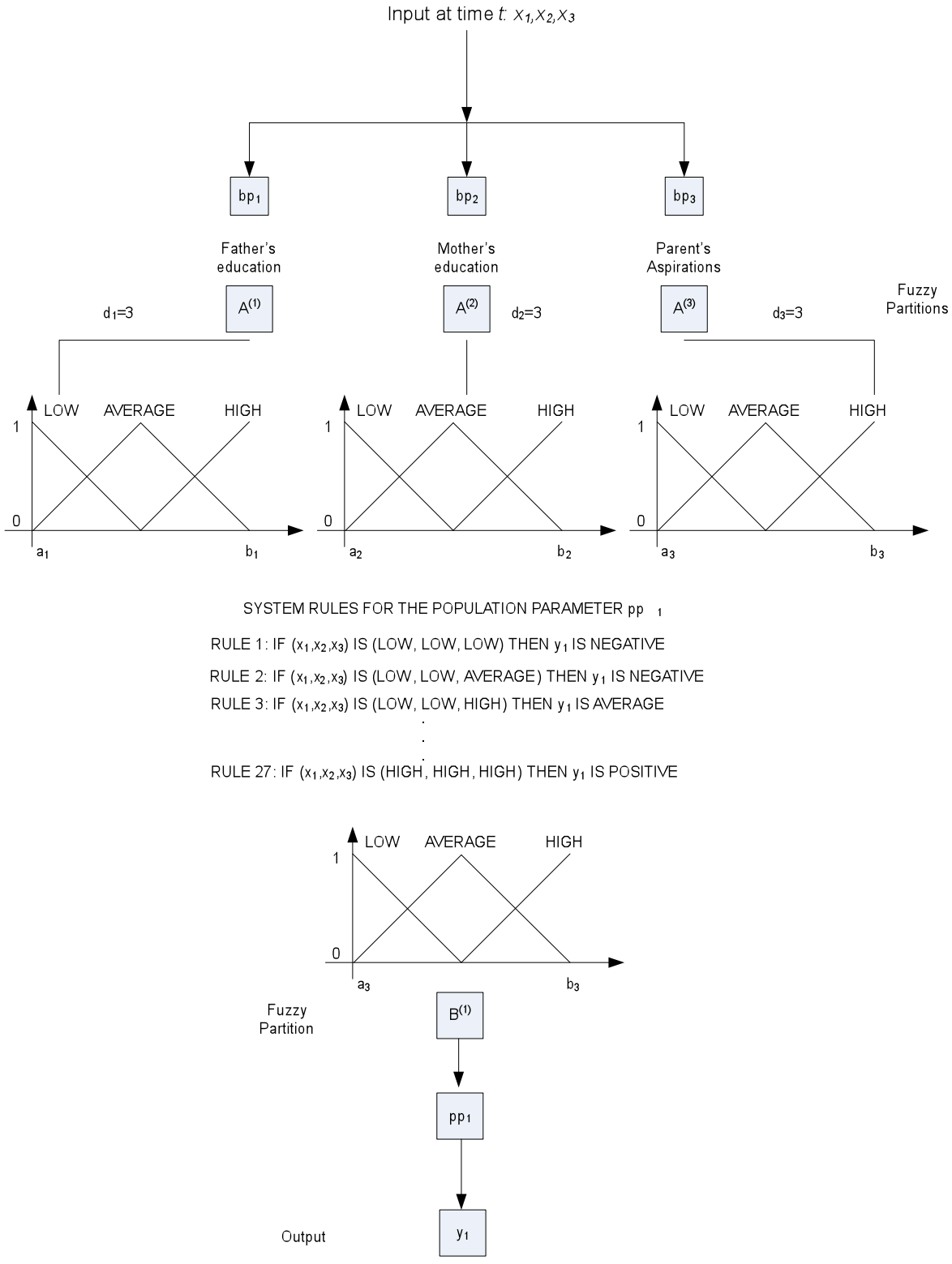


Figure 4 Regulations of the system for the population parameter pp_1.

4. Conclusions – Future Work

In this paper, Fuzzy Markov Systems are proposed as a new technique in order to study the occupational choices that people make. In response to the weakness that traditional methods face concerning the need of a precise estimation of the transition probabilities, we propose the use of Fuzzy Logic and Fuzzy Reasoning in the theory of Markov Systems. Thus, we take into account the theoretical and empirical knowledge existing about the occupational choices and we create a fuzzy expert system (Fuzzy Inference System) consisting of linguistic rules. Especially in a topic such as occupational choices, the heuristic knowledge of the experts could be a powerful tool, providing us with essential information about the system.

More specifically, as the family generally appears to be a significant factor influencing the occupational choices of the children, we focus on factors deriving from the family that seem to be important in Greece. This is a first attempt to take advantage of all the linguistic elements that contribute to the understanding of this issue and moreover an effort to incorporate the real social conditions. We, thus, envisage the development of a realistic and useful tool.

Future work is needed in order to provide a better inside to the system by considering the time as non homogeneous. It is more realistic to observe the system at time $t = 0,1,2,...$ and to consider the state of the system as a function of time *t*, i.e. $\mathbf{N}(t) = \left[N_1(t), N_2(t), \ldots, N_k(t)\right]$. The sequence of the transition matrices will be represented by $\left\{\mathbf{P}(t)\right\}_{t=0}^{\infty} = \left\{p_{ij}\left(t\right)\right\}_{i,j \in S}$, where $p_{ij}\left(t\right)$ denotes the transition probability from state *i* to state *j* in the time interval $\left(t-1, t\right]$.

Implementing the system to cohort data is also a next step towards the improvement of the system proposed.

References

1. Bartholomew D. J., Stochastic models for Social Processes, 3rd Ed., *Willey*, *Chichester*, (1982).
2. Bartholomew D. J., A. Forbes and S. McClean, Statistical techniques for manpower planning, *Willey*, *Chichester*, (1991).
3. Blau, P. M., Gustad, J. W., Jessor, R., Parries, H. S., & Wilcock, R. C. Occupational choice: A conceptual framework. *Industrial Labor Relations Review*, 9, pp 531-543, (1956).
4. Blau P. and Duncan O., The American Occupational Structure, *John Willey and Sons*, *New York*, (1967).
5. Fragoudaki A., Sociology of education, Theories about social inequality in school, *Papazisis*, *Athens*, (1985).
6. Ginsberg E., Career Guidance, *McGraw-Hill*, *New York*, (1951).
7. Holland J. L., A new synthesis for an old method and a new analysis of some old phenomena, *The Counseling Psychologist*, Vol. 6, pp 12-15, (1976).
8. Hoppock R., Occupational Information, McGraw-Hill, (1976).
9. Kasimati K., Research about the social characteristics of occupation, Study I, "The occupational choice, Theoretical references and empirical research", 2nd Ed., *National Centre for Social Research*, *Athens*, (1998).
10. Kasimati K., Structures and Flows, The phenomenon of social and occupational mobility, *Gutenberg*, *Athens*, (2004).
11. Kintis A. A., The higher education in Greece, Anatomy and thoughts for its reconstruction, *Gutenberg*, *Athens*, (1980).
12. Klir G. J. and B. Yuan, Fuzzy sets and fuzzy logic: theory and applications, *Prentice Hall PTR*, USA, (1995).
13. Kontogiannopoulou – Polydoridi G., Educational policy and practice, Sociological analysis, *Ellinika Grammata*, *Athens*, (1995).
14. Kontogiannopoulou – Polydoridi G., Sociological analysis of the performance and of the evaluation, The entrance examinations, Establishment of the performance, Integration in a hierarchical higher education, Vol. II, *Gutenberg*, *Athens*, [1996],

15. Lampiri-Dimaki I., Towards a Greek sociology of Education, Greek Students: Origins and Perspectives, *National Centre for Social Research*, *Athens*, (1974).
16. McClean S. I., A continuous time population model with Poisson recruitment *J. Appl. Prob.*, 15, pp. 26-32, (1976).
17. McClean S. I., Continuous time stochastic models of a multigrade population. *J. Appl. Prob.*, 16, pp. 28-39, (1978).
18. McClean S. I., A semi-Markov model for a multistage population with Poisson recruitment. *J. Appl. Prob.*, 17, pp. 846-852, (1980).
19. Stamou G. B. and S. Tzafestas, Fuzzy Relation Equations and Fuzzy Inference Systems: an Inside Approach, *IEEE Trans. on Systems*, *Man and Cybernetics*, Vol. 99, Num. 6, pp. 694-702, (1999).
20. Super D. E., Vocational Adjustment: Implementing of self Concept, Occupations, No. 30, pp 88-92, (1951).
21. Symeonaki M. A., G. B Stamou and S. G. Tzafestas, Fuzzy Markov Systems for the Description and Control of Population Dynamics, *Computational Intelligence in Systems and Control of Design and Applications*, *Kluwer Academic Publishers*, pp. 301-310, Boston, (2000).
22. Symeonaki M. A., G. B. Stamou, S. G. Tzafestas, Fuzzy non-homogenous Markov systems, *Applied Intelligence*, 17(2), pp. 203-214, (2002).
23. Symeonaki M. A., Modeling Conceptual Change in Physics using Probabilistic, Fuzzy and Symbolic Approaches, *International Mathematical Forum*, 1, Num. 12, pp. 577-591, (2006).
24. Vassiliou P.-C. G., Asymptotic behavior of Markov systems, *J. Appl. Prob.*, 19, pp. 851-857, (1982).
25. Viki A. and Papanis E., The occupation of parents and the occupational choise of the children, Greek Social Survey, September 9th, 2007, http://epapanis.blogspot.com/2007/09/blog-post_3528.html .
26. Viki A. and Papanis E., The effect of students' gender and of the educational level of their parents in shaping occupational types, Greek Social Survey, September 1st, 2007, http://epapanis.blogspot.com/2007/09/blog-post_726.html .
27. Zadeh L. A., Fuzzy Sets, *Information and Control*, 8(3), pp. 338-353, (1965).

On deviations of the sample mean for Markov chains

Zbigniew S. Szewczak

Faculty of Mathematics and Computer Science, Nicolaus Copernicus University,
ul. Chopina 12/18, 87-100 Toruń, Poland,
(e-mail: zssz@mat.uni.torun.pl)

Abstract. For homogeneous Markov chains satisfying the uniform recurrence condition Bahadur-Rao's expansions are established.
Keywords: large deviation, Markov chains.

1 Introduction and results

Let $\{\xi_k\}_{k\in\mathbb{N}}$ be a homogeneous Markov chain with a regular transition probability $P(x,A)$ (cf. [2]) and an initial distribution $\mu(A)=P[\xi_0\in A]$, where A is a Borel set, $x\in\mathbb{R}$. In the case where $\{\xi_k\}$ is an independent identically distributed sequence of random variables Bahadur and Rao (cf. [1]) derived the full asymptotic expansion for deviations of the sample mean. For Markov chains the investigation of this problem was continued in [4], however the corresponding expansions were not carried out. This note aims to fill this gap.

Let $L^\infty(\mu)$ denote a Banach algebra of complex valued Borel functions with the essential supremum norm $\big|\!\cdot\!\big|$. Define the following uniform recurrence (cf. [4]) condition (Ψ)

$$0<a=\inf_{\{A\,|\,\mu(A)>0\}}\operatorname*{ess\,inf}_x\frac{P(x,A)}{\mu(A)}\le\sup_{\{A\,|\,\mu(A)>0\}}\operatorname*{ess\,sup}_x\frac{P(x,A)}{\mu(A)}=b<\infty.$$

Under the above condition the Perron-Frobenius theorem holds (cf. [9], Proposition 1) so that $\{\xi_k\}$ is uniformly ergodic, i.e. there exist $C>0$, $\gamma,\ 0\le|\gamma|<1$ and stationary probability π such that

$$\Big|\sup_A|P^n(x,A)-\pi(A)|\Big|\le C|\gamma|^n.$$

Define the class $\mathfrak{C}(\mu)$ of Borel functions such that for any $f\in\mathfrak{C}(\mu)$ there exist, depending on f, non-empty interval $\mathbb{S}=(s_-,s_+)$ such that for any $s\in\mathbb{S}$ we have

$$\int e^{sf(y)}\mu(dy)<\infty. \tag{1}$$

For $s\in\mathbb{S}$ and $f\in\mathfrak{C}(\mu)$ define a linear operator

$$\mathbf{L}_s h=\int e^{sf(y)}h(y)P(\,\cdot\,,dy).$$

Let $\rho(s)$ be the spectral radius, u_s be the density of the eigenmeasure with respect to μ (eigendensity) and v_s be the eigenfunction of $\mathbf{L}_s$, respectively. We choose them normalized so that $\int v_s(y)u_s(y)\mu(dy) = 1$. In view of (Ψ) we have $\operatorname{ess\,inf}_x v_s(x) > 0$, so we may define the regular conjugate conditional distribution $P_s(x,y)$ by

$$\frac{P_s(x,dy)}{P(x,dy)} = \frac{e^{sf(y)}v_s(y)}{\rho(s)v_s(x)}.$$

For any $f \in \mathfrak{C}(\mu)$ we have uniformy ergodic chains $\{\xi_k^s\}$ defined on the same phase space as $\{\xi_k\}$ and governed by the conjugate transition probabilities $P_s(x,A)$ (cf. [9], Proposition 1). The stationary distribution π_s for $P_s(x,dy)$ is defined by $\pi_s(A) = \int I_A(x)v_s(x)u_s(x)\mu(dx)$. Let $X_k^s = f(\xi_k^s)$ and $S_n^s = \sum_{k=1}^n X_k^s$, while $X_k = X_k^0$ and $S_n = S_n^0$. We say that $\mathcal{L}(X_0)$ satisfies Cramér's condition (C) if:

$$\limsup_{|\theta|\to\infty} |E[e^{i\theta X_0}]| < 1. \tag{2}$$

Put, for short, $\bar{f}(\cdot) = f(\cdot) - E_{\pi_s}[f]$. Assuming (Ψ) to hold we have (cf. [9], Lemma 3)

$$\sigma_s^2 = E_{\pi_s}[\bar{f}^2(\xi_0^s)] + 2\sum_{n=1}^{\infty} E_{\pi_s}[\bar{f}(\xi_0^s)\bar{f}(\xi_n^s)] > 0$$

for $f \in \mathfrak{C}(\mu)$ and $s \in \mathbb{S}$. Let $f^- = \operatorname{ess\,inf}_x f(x)$, $f^+ = \operatorname{ess\,sup}_x f(x)$. The following theorems generalize results in [1],[5], [4] and [9] (see also [7]).

Theorem 1. *Assume (Ψ). Suppose (2) holds for some $f \in \mathfrak{C}(\mu)$, $E_\pi[f] = 0$. Then $\sigma_s^2 > 0$ for any $s \in \mathbb{S}$ and for $0 \le h \in L^\infty(\mu)$ we have:*
1) if $f^+ > F^+ > \epsilon > 0$, then for $k \ge 0$

$$\sup_{\epsilon\le t\le F^+} \left| \frac{E[I_{[S_n\ge nt]}h(\xi_n)\,|\,\xi_0 = \cdot\,]}{\frac{v_s}{\sigma_s\sqrt{2\pi n s}}\left(\frac{\rho(s)}{e^{st}}\right)^n \sum_{\nu=0}^k \frac{1}{n^\nu}\mathbf{C}_{\nu s}(\frac{h}{v_s})} - 1 \right| = O(\frac{1}{n^{k+1}})$$

2) if $f^- < F^- < -\epsilon < 0$, then for $k \ge 0$

$$\sup_{F^-\le t\le -\epsilon} \left| \frac{E[I_{[S_n< nt]}h(\xi_n)\,|\,\xi_0 = \cdot\,]}{\frac{v_s}{\sigma_s\sqrt{2\pi n s}}\left(\frac{\rho(s)}{e^{st}}\right)^n \sum_{\nu=0}^k \frac{1}{n^\nu}\mathbf{C}_{\nu s}(\frac{h}{v_s})} - 1 \right| = O(\frac{1}{n^{k+1}})$$

where $\mathbf{C}_{\nu s}$ are bounded linear operators.

Theorem 2. *Assume (Ψ). Suppose $\mathcal{L}(X_0)$ is on a lattice $\{\nu d\}$, $\nu \in \mathbb{Z}$, for $f \in \mathfrak{C}(\mu)$, $E_\pi[f] = 0$. Then $\sigma_s^2 > 0$ for any $s \in \mathbb{S}$ and for $0 \le h \in L^\infty(\mu)$ we have:*
1) if nt takes values of the form νd then, for $f^+ > F^+ > \varepsilon > 0$, $k \ge 0$

$$\sup_{\epsilon\le t\le F^+} \left| \frac{E[I_{[S_n\ge nt]}h(\xi_n)\,|\,\xi_0 = \cdot\,]}{\frac{v_s}{\sigma_s\sqrt{2\pi n}}\left(\frac{\rho(s)}{e^{st}}\right)^n \sum_{\nu=0}^k \frac{1}{n^\nu}\mathbf{C}_{\nu s}(\frac{h}{v_s})} - \frac{d}{1-e^{-sd}} \right| = O(\frac{1}{n^{k+1}})$$

2) if nt takes values of the form νd then, for $f^- < F^- < -\varepsilon < 0$, $k \geq 0$

$$\sup_{F^- \leq t \leq -\epsilon} \left| \frac{E[I_{[S_n<nt]}h(\xi_n)\,|\,\xi_0 = \cdot\,]}{\frac{v_s}{\sigma_s\sqrt{2\pi n}}\left(\frac{\rho(s)}{e^{st}}\right)^n \sum_{\nu=0}^{k} \frac{1}{n^\nu}\mathbf{C}_{\nu s}(\frac{h}{v_s})} - \frac{d}{1-e^{-sd}} \right| = O(\frac{1}{n^{k+1}})$$

where $\mathbf{C}_{\nu s}$ *are bounded linear operators.*

The operator coefficients $\mathbf{C}_{\nu s}$ are given explicitly (cf. [8]).

2 Proofs

2.1 Proof of Theorem 1

Write

$$H_{ns}(y) = E_s[I_{[S_n^s - nt<y]}h_s(\xi_n)\,|\,\xi_0 = \cdot\,],$$

where $h_s = \frac{h}{v_s}\cdot$ Denote by $\mathfrak{Q}_{1s}, \mathfrak{Q}_{2s}, \mathfrak{Q}_{2s}, \ldots, \mathfrak{Q}_{ks}$ the linear operators in the asymptotic expansions for the conjugate transition operator. Then, by Theorem 1 and 2 in [8], we obtain

$$H_{ns}(y\sigma_s\sqrt{n}) = \mathfrak{N}(y)\mathbf{\Pi}_s h_s + \mathfrak{n}(y)\left(\frac{\mathfrak{Q}_{1s}(y)h_s}{n^{\frac{1}{2}}} + \cdots + \frac{\mathfrak{Q}_{ks}(y)h_s}{n^{\frac{k-2}{2}}}\right) + o(\frac{1}{n^{\frac{k-2}{2}}}) \quad (3)$$

where $\mathbf{\Pi}_s h = E_{\pi_s}[h]$. Replacing k by $2k+5$ and y by $\frac{y}{\sqrt{n}s\sigma_s}$ in (3) we can decompose uniformly in y

$$H_{ns}(\frac{y}{s}) = G^{(k)}_{h_s}(\frac{y}{\sqrt{n}s\sigma_s}) + O(\frac{1}{n^{k+\frac{3}{2}}}) \quad (4)$$

where

$$G^{(k)}_{h_s}(\frac{y}{\sqrt{n}s\sigma_s}) = \mathfrak{N}(\frac{y}{\sqrt{n}s\sigma_s})\mathbf{\Pi}_s h_s + \mathfrak{n}(\frac{y}{\sqrt{n}s\sigma_s}) \sum_{m=1}^{2k+2} \frac{\mathfrak{Q}_{ms}(\frac{y}{\sqrt{n}s\sigma_s})h_s}{n^{\frac{m}{2}}}.$$

The Fourier-Stieltjes transform of the above is

$$\hat{G}^{(k)}_{h_s}(\theta) = e^{-\frac{\theta^2}{2}}\Big(\sum_{m=0}^{2k+2} \frac{1}{(\sqrt{n})^m} \sum_{j=0}^{m} \frac{1}{j!}(\frac{i\theta}{\sigma_s})^j \mathfrak{P}_{m-j}(i\theta)\hat{\mathbf{P}}^{(j)}_{1s}h_s\Big) \quad (5)$$

where

$$\mathfrak{P}_\nu(\zeta) = \sum_{\substack{(k_1,k_2,\ldots,k_\nu) \\ k_i \geq 0, \sum_{i=1}^{\nu} i k_i = \nu}} \prod_{m=3}^{\nu+2} \frac{1}{k_{m-2}!}\left(\frac{\gamma_{ms}\zeta^m}{m!\sigma_s^m}\right)^{k_{m-2}}$$

and $\mathfrak{P}_0(\zeta) = 1$ for $\zeta \in \mathbb{C}$. Here $\gamma_{ms} = \frac{1}{i^m}\frac{\partial^m}{\partial\theta^m}\log\lambda_s(\theta)|_{\theta=0}$ and $\lambda_s(\theta)$ is the principal eigenvalue of characteristic operator

$$\hat{\mathbf{P}}_s(\theta)(g)(x) = \int_{-\infty}^{\infty} e^{i\theta(f(y)-E_{\pi_s}[f])}g(y)P_s(x,dy)$$

$g \in L^\infty(\mu)$. Linear operators $\hat{\mathbf{P}}_{1s}^{(j)}$ in (5) are derivatives at zero of the projections $\hat{\mathbf{P}}_{1s}(\theta)$ on the eigenspace corresponding to $\lambda_s(\theta)$ (cf. [9]). Note that

$$\int_0^\infty e^{-sy}H_{ns}(dy) = s\int_0^\infty e^{-sy}(H_{ns}(y) - H_{ns}(0))dy.$$

Taking into account this and (4) we get

$$\begin{aligned}\int_0^\infty e^{-sy}H_{ns}(dy) &= \int_0^\infty e^{-y}(H_{ns}(\frac{y}{s}) - H_{ns}(0))dy\\ &= \int_0^\infty e^{-y}(G_{h_s}^{(k)}(\frac{y}{\sqrt{n}s\sigma_s}) - G_{h_s}^{(k)}(0))dy + O(\frac{1}{n^{k+\frac{3}{2}}})\cdot\end{aligned}$$

Hence, integrating by parts

$$\begin{aligned}\int_0^\infty e^{-sy}H_{ns}(dy) &= \int_0^\infty e^{-y}G_{h_s}^{(k)}(\frac{dy}{\sqrt{n}s\sigma_s}) + O(\frac{1}{n^{k+\frac{3}{2}}})\\ &= \int_0^\infty e^{-y}\frac{\partial}{\partial y}G_{h_s}^{(k)}(\frac{y}{\sqrt{n}s\sigma_s})dy + O(\frac{1}{n^{k+\frac{3}{2}}})\cdot\end{aligned}$$

Now, using Parseval's formula

$$\begin{aligned}\int_0^\infty e^{-sy}H_{ns}(dy) &= \int_{-\infty}^\infty e^{-y}I_{(0,\infty)}(y)\frac{\partial}{\partial y}G_{h_s}^{(k)}(\frac{y}{\sqrt{n}s\sigma_s})dy + O(\frac{1}{n^{k+\frac{3}{2}}})\\ &= \frac{1}{2\pi}\int_{-\infty}^\infty \hat{G}_{h_s}^{(k)}(\theta\sqrt{n}s\sigma_s)\overline{\left(\frac{1}{1-i\theta}\right)}d\theta + O(\frac{1}{n^{k+\frac{3}{2}}})\cdot\end{aligned}$$

Thus, writing out the truncated MacLaurin expansion

$$\int_0^\infty e^{-sy}H_{ns}(dy) = \frac{1}{2\pi\sqrt{n}s\sigma_s}\sum_{l=0}^{2k+1}\int_{-\infty}^\infty\left(\frac{-i\theta}{\sqrt{n}s\sigma_s}\right)^l\hat{G}_{h_s}^{(k)}(\theta)d\theta + O(\frac{1}{n^{k+\frac{3}{2}}})\cdot$$

In view of this and (5)

$$\begin{aligned}\int_0^\infty e^{-sy}H_{ns}(dy) = \frac{1}{\sqrt{2\pi n}s\sigma_s}&\sum_{l=0}^{2k+1}\sum_{m=0}^{2k+1-l}\sum_{j=0}^{m}\hat{\mathbf{P}}_{1s}^{(j)}h_s\\ &\times\int_{-\infty}^\infty\left(\frac{-i\theta}{\sqrt{n}s\sigma_s}\right)^l\frac{1}{(\sqrt{n})^m}\frac{1}{j!}(\frac{i\theta}{\sigma})^j\mathfrak{P}_{m-j}(i\theta)\mathfrak{N}(d\theta) + O(\frac{1}{n^{k+\frac{3}{2}}})\cdot\end{aligned}$$

Polynomial $(i\theta)^{l+j}\mathfrak{P}_{m-j}(i\theta)$ is even [odd] polynomial if $l+m$ is even [odd]. Since

$$\int_{-\infty}^{\infty}\theta^{2j+1}\mathfrak{N}(d\theta)=0,$$

therefore,

$$\int_0^\infty e^{-sy}H_{ns}(dy)=\frac{1}{\sqrt{2\pi n}s\sigma_s}\sum_{\nu=0}^{k}\frac{1}{n^\nu}\sum_{l+m=2\nu}\sum_{j=0}^{m}\frac{(-1)^l}{j!s^l\sigma_s^{l+j}}\hat{\mathbf{P}}_{1s}^{(j)}h_s \\ \times\int_{-\infty}^{\infty}(i\theta)^{l+j}\mathfrak{P}_{m-j}(i\theta)\mathfrak{N}(d\theta)\left(1+O(\frac{1}{n^{k+1}})\right).$$

Consequently,

$$\int_0^\infty e^{-sy}H_{ns}(dy)=\frac{1}{\sqrt{2\pi n}s\sigma_s}\sum_{\nu=0}^{k}\frac{1}{n^\nu}\sum_{l=0}^{2\nu}\sum_{j=0}^{2\nu-l}\frac{(-1)^l}{j!s^l\sigma_s^{l+j}}\hat{\mathbf{P}}_{1s}^{(j)}h_s \\ \times\int_{-\infty}^{\infty}(i\theta)^{l+j}\mathfrak{P}_{2\nu-l-j}(i\theta)\mathfrak{N}(d\theta)\left(1+O(\frac{1}{n^{k+1}})\right).$$

Thus

$$\int_0^\infty e^{-sy}H_{ns}(dy)=\frac{1}{\sqrt{2\pi n}s\sigma_s}\sum_{\nu=0}^{k}\frac{1}{n^\nu}\mathbf{C}_{\nu s}(\frac{h}{v_s})\left(1+O(\frac{1}{n^{k+1}})\right)$$

where

$$\mathbf{C}_{\nu s}=\sum_{l=0}^{2\nu}\sum_{j=0}^{2\nu-l}\frac{(-1)^l}{j!s^l\sigma_s^{l+j}}\hat{\mathbf{P}}_{1s}^{(j)}\sum_{\substack{(k_1,k_2,\dots,k_{2\nu-l-j})\\ k_i\ge 0,\ \sum_{i=1}^{2\nu-l-j}ik_i=2\nu-l-j}}(-1)^{\nu+\sum_{i=1}^{2\nu-l-j}k_i}\mathfrak{m}_{2\nu+2\sum_{i=1}^{2\nu-l-j}k_i} \\ \times\prod_{m=3}^{2\nu-l-j+2}\frac{1}{k_{m-2}!}\Big(\frac{\gamma_{ms}}{m!\sigma_s^m}\Big)^{k_{m-2}}$$

$\mathfrak{m}_{2k}=1\cdot 3\cdots(2k-3)\cdot(2k-1)$ and the third sum equals $(-1)^\nu\mathfrak{m}_{2\nu}$ when $2\nu=l+j$. Now, by Lemma 4 in [9]

$$E[I_{[S_n\ge nt]}h(\xi_n)\,|\,\xi_0=x]=(e^{-st}\rho(s))^n v_s(x)\int_0^\infty e^{-sy}H_{ns}(dy),$$

whence Theorem 1 is proved since all the estimates hold uniformly for $t\in[\epsilon,F^+]$.

2.2 Proof of Theorem 2

By Lemma 4 in [9]

$$E[I_{[S_n\geq nt]}h(\xi_n)\,|\,\xi_0=x]$$
$$=(e^{-st}\rho(s))^n v_s(x)\sum_{\nu\geq 0}e^{-s\nu d}E_s[I_{[S_n^s-nt=\nu d]}h_s(\xi_n)\,|\,\xi_0=x].$$

Let $y_{n\nu}=\frac{\nu d}{\sigma_s\sqrt{n}}\cdot$ It can be proved that for $2k+4$ we have expansion (cf. [3], Theorem 1 on p.241; [6])

$$\frac{\sigma_s\sqrt{n}}{d}E_s[I_{[S_n^s-nt=\nu d]}h_s(\xi_n)\,|\,\xi_0=\,\cdot\,]$$
$$=\mathfrak{n}(y_{n\nu})\mathbf{\Pi}_s h_s+\mathfrak{n}(y_{n\nu})\sum_{m=0}^{2k+1}\frac{\mathfrak{R}_{ms}(y_{n\nu})h_s}{n^{\frac{m}{2}}}+O(\frac{1}{n^{\frac{2k+2}{2}}}),$$

where $\mathfrak{R}_{ms}(y)=\frac{\partial}{\partial y}\mathfrak{Q}_{ms}(y)$. Further,

$$\mathfrak{n}(y_{n\nu})\mathbf{\Pi}_s h_s+\mathfrak{n}(y_{n\nu})\sum_{m=0}^{2k+1}\frac{\mathfrak{R}_{ms}(y_{n\nu})h_s}{n^{\frac{m}{2}}}$$
$$=\frac{1}{2\pi}\int e^{-i\theta y_{n\nu}}e^{-\frac{\theta^2}{2}}\Big(\sum_{m=0}^{2k+1}\frac{1}{(\sqrt{n})^m}\sum_{j=0}^{m}\frac{1}{j!}(\frac{i\theta}{\sigma_s})^j\mathfrak{P}_{m-j}(i\theta)\hat{\mathbf{P}}_{1s}^{(j)}h_s\Big)d\theta.$$

Therefore,

$$\sum_{\nu\geq 0}e^{-s\nu d}\Big(\mathfrak{n}(y_{n\nu})\mathbf{\Pi}_s h_s+\mathfrak{n}(y_{n\nu})\sum_{m=0}^{2k+1}\frac{\mathfrak{R}_{ms}(y_{n\nu})h_s}{n^{\frac{m}{2}}}\Big)$$
$$=\frac{1}{\sqrt{2\pi}}\int(1-e^{-sd-\frac{i\theta d}{\sigma_s\sqrt{n}}})\Big(\sum_{m=0}^{2k+1}\frac{1}{(\sqrt{n})^m}\sum_{j=0}^{m}\frac{1}{j!}(\frac{i\theta}{\sigma_s})^j\mathfrak{P}_{m-j}(i\theta)\hat{\mathbf{P}}_{1s}^{(j)}h_s\Big)\mathfrak{N}(d\theta).$$

Now let for $l\geq 0$

$$b_l=\frac{\partial^l}{\partial y^l}\frac{1-e^{-sd}}{1-e^{-sd}e^{-y}}\Big|_{y=0}.$$

Whence,

$$
\begin{aligned}
&(1-e^{-sd})\sum_{\nu\geq 0}e^{-s\nu d}\Big(\mathfrak{n}(y_{n\nu})\mathbf{\Pi}_s h_s+\mathfrak{n}(y_{n\nu})\sum_{m=0}^{2k+1}\frac{\mathfrak{R}_{ms}(y_{n\nu})h_s}{n^{\frac{m}{2}}}\Big)\\
&=\frac{1}{\sqrt{2\pi}}\sum_{l=0}^{2k+1}b_l\frac{d^l}{\sigma_s^l}\frac{1}{(\sqrt{n})^l}\\
&\qquad\times\int\frac{(i\theta)^l}{l!}\Big(\sum_{m=0}^{2k+1}\frac{1}{(\sqrt{n})^m}\sum_{j=0}^{m}\frac{1}{j!}\big(\frac{i\theta}{\sigma_s}\big)^j\mathfrak{P}_{m-j}(i\theta)\hat{\mathbf{P}}_{1s}^{(j)}h_s\Big)\mathfrak{N}(d\theta)\\
&\qquad\times\Big(1+O(\frac{1}{n^{k+1}})\Big)\\
&=\frac{1}{\sqrt{2\pi}}\sum_{l=0}^{2k+1}\sum_{m=0}^{2k+1-l}\sum_{j=0}^{m}b_l\frac{d^l}{\sigma_s^l}\frac{1}{(\sqrt{n})^l}\frac{1}{(\sqrt{n})^m}\frac{1}{j!}\\
&\qquad\times\int\frac{(i\theta)^l}{l!}\big(\frac{i\theta}{\sigma_s}\big)^j\mathfrak{P}_{m-j}(i\theta)\hat{\mathbf{P}}_{1s}^{(j)}h_s\mathfrak{N}(d\theta)\Big(1+O(\frac{1}{n^{k+1}})\Big).
\end{aligned}
$$

Polynomial $(i\theta)^{l+j}\mathfrak{P}_{m-j}(i\theta)$ is even [odd] polynomial if $l+m$ is even [odd]. Consequently,

$$
\begin{aligned}
&(1-e^{-sd})\sum_{\nu\geq 0}e^{-s\nu d}\Big(\mathfrak{n}(y_{n\nu})\mathbf{\Pi}_s h_s+\mathfrak{n}(y_{n\nu})\sum_{m=0}^{2k+1}\frac{\mathfrak{R}_{ms}(y_{n\nu})h_s}{n^{\frac{m}{2}}}\Big)\\
&\qquad=\frac{1}{\sqrt{2\pi}}\sum_{\nu=0}^{k}\frac{1}{n^\nu}\sum_{l+m=2\nu}\sum_{j=0}^{m}\frac{d^l b_l}{j!l!\sigma_s^{l+j}}\hat{\mathbf{P}}_{1s}^{(j)}h_s\\
&\qquad\qquad\times\int_{-\infty}^{\infty}(i\theta)^{l+j}\mathfrak{P}_{m-j}(i\theta)\mathfrak{N}(d\theta)\left(1+O(\frac{1}{n^{k+1}})\right)\\
&\qquad=\frac{1}{\sqrt{2\pi}}\sum_{\nu=0}^{k}\frac{1}{n^\nu}\sum_{l=0}^{2\nu}\sum_{j=0}^{2\nu-l}\frac{d^l b_l}{j!l!\sigma_s^{l+j}}\hat{\mathbf{P}}_{1s}^{(j)}h_s\\
&\qquad\qquad\times\int_{-\infty}^{\infty}(i\theta)^{l+j}\mathfrak{P}_{2\nu-l-j}(i\theta)\mathfrak{N}(d\theta)\left(1+O(\frac{1}{n^{k+1}})\right)\\
&\qquad=\frac{1}{\sqrt{2\pi}}\sum_{\nu=0}^{k}\frac{1}{n^\nu}\mathbf{C}_{\nu s}(\frac{h}{v_s})\left(1+O(\frac{1}{n^{k+1}})\right)
\end{aligned}
$$

where

$$\mathbf{C}_{\nu s}=\sum_{l=0}^{2\nu}\sum_{j=0}^{2\nu-l}\frac{d^l b_l}{j!l!\sigma_s^{l+j}}\hat{\mathbf{P}}_{1s}^{(j)}\sum_{\substack{(k_1,k_2,\dots,k_{2\nu-l-j})\\ k_i\geq 0,\ \sum_{i=1}^{2\nu-l-j} ik_i=2\nu-l-j}}(-1)^{\nu+\sum_{i=1}^{2\nu-l-j}k_i}\mathfrak{m}_{2\nu+2\sum_{i=1}^{2\nu-l-j}k_i}$$
$$\times\prod_{m=3}^{2\nu-l-j+2}\frac{1}{k_{m-2}!}\Big(\frac{\gamma_{ms}}{m!\sigma_s^m}\Big)^{k_{m-2}}$$

and $\mathfrak{m}_{2k}=1\cdot 3\cdots(2k-3)\cdot(2k-1)$. The proof is completed.

3 Remarks

The first three of linear operators $\mathbf{C}_{\nu s}$ in the non-lattice case are

$$\mathbf{C}_{0s}=\mathbf{\Pi}_s,$$

$$\mathbf{C}_{1s}=\{\frac{\gamma_{4s}}{24\sigma_s^4}\mathfrak{m}_4-\frac{\gamma_{3s}^2}{72\sigma_s^6}\mathfrak{m}_6-\frac{1}{s\sigma_s}\frac{\gamma_{3s}}{6\sigma_s^3}\mathfrak{m}_4\}\mathbf{\Pi}_s$$
$$+\{\frac{1}{\sigma_s}\frac{\gamma_{3s}}{6\sigma_s^3}\mathfrak{m}_4+\frac{1}{s\sigma_s^2}\mathfrak{m}_2\}\hat{\mathbf{P}}_{1s}^{(1)}-\{\frac{1}{2\sigma_s^2}\mathfrak{m}_2\}\hat{\mathbf{P}}_{1s}^{(2)},$$

$$\begin{aligned}\mathbf{C}_{2s}=\{&-\frac{\gamma_{6s}}{720\sigma_s^6}\mathfrak{m}_6+\frac{\gamma_{5s}\gamma_{3s}}{720\sigma_s^8}\mathfrak{m}_8+\frac{\gamma_{4s}^2}{1152\sigma_s^8}\mathfrak{m}_8-\frac{\gamma_{4s}\gamma_{3s}^2}{1728\sigma_s^{10}}\mathfrak{m}_{10}\\
&+\frac{\gamma_{3s}^4}{31104\sigma_s^{12}}\mathfrak{m}_{12}-\frac{1}{s\sigma_s}(-\frac{\gamma_{5s}}{120\sigma_s^5}\mathfrak{m}_6+\frac{\gamma_{4s}\gamma_{3s}}{144\sigma_s^7}\mathfrak{m}_8-\frac{\gamma_{3s}^2}{1296\sigma_s^9}\mathfrak{m}_{10})\\
&+\frac{1}{s^2\sigma_s^2}(-\frac{\gamma_{4s}}{24\sigma_s^4}\mathfrak{m}_6+\frac{\gamma_{3s}^2}{72\sigma_s^6}\mathfrak{m}_8)-\frac{1}{s^3\sigma_s^3}(-\frac{\gamma_{3s}}{6\sigma_s^3})\mathfrak{m}_6+\frac{1}{s^4\sigma_s^4}\mathfrak{m}_4\}\mathbf{\Pi}_s\\
&+\{\frac{1}{\sigma_s}(-\frac{\gamma_{5s}}{120\sigma_s^5}\mathfrak{m}_6+\frac{\gamma_{4s}\gamma_{3s}}{144\sigma_s^7}\mathfrak{m}_8-\frac{\gamma_{3s}^2}{1296\sigma_s^9}\mathfrak{m}_{10})\\
&-\frac{1}{s\sigma_s^2}(-\frac{\gamma_{4s}}{24\sigma_s^4}\mathfrak{m}_6+\frac{\gamma_{3s}^2}{72\sigma_s^6}\mathfrak{m}_8)+\frac{1}{s^2\sigma_s^3}(-\frac{\gamma_{3s}}{6\sigma_s^3})\mathfrak{m}_6-\frac{1}{s^3\sigma_s^4}\mathfrak{m}_4\}\hat{\mathbf{P}}_{1s}^{(1)}\\
&+\{\frac{1}{2\sigma_s^2}(-\frac{\gamma_{4s}}{24\sigma_s^4}\mathfrak{m}_6+\frac{\gamma_{3s}^2}{72\sigma_s^6}\mathfrak{m}_8)-\frac{1}{2s\sigma_s^3}(-\frac{\gamma_{3s}}{6\sigma_s^3})\mathfrak{m}_6+\frac{1}{2s^2\sigma_s^4}\mathfrak{m}_4\}\hat{\mathbf{P}}_{1s}^{(2)}\\
&+\{\frac{1}{6\sigma_s^3}(-\frac{\gamma_{3s}}{6\sigma_s^3})\mathfrak{m}_6-\frac{1}{6s\sigma_s^4}\mathfrak{m}_4\}\hat{\mathbf{P}}_{1s}^{(3)}+\{\frac{1}{24\sigma_s^4}\mathfrak{m}_4\}\hat{\mathbf{P}}_{1s}^{(4)},\end{aligned}$$

where $\mathfrak{m}_k$ is k-th moment of standard normal random variable. For the lattice case denote

$$b=\frac{e^{-sd}}{1-e^{-sd}}.$$

Then every expression $(-\frac{1}{s})^l$ in the above operators need to be replaced by $\frac{d^l b_l}{l!}$ where

$$b_0 = 1,$$

$$b_1 = -b,$$

$$b_2 = b^2 + \frac{1}{2}b,$$

$$b_3 = -b^3 - b^2 - \frac{1}{6}b,$$

$$b_4 = b^4 + \frac{3}{2}b^3 + \frac{7}{12}b^2 + \frac{1}{24}b,$$

$$b_5 = -b^5 - 2b^4 - \frac{5}{4}b^3 - \frac{1}{4}b^2 - \frac{1}{120}b.$$

The above calculations can be implemented using *Maple* package.

References

1. Bahadur, R. R., and Rao, R. Ranga, "On deviations of the sample mean", *Ann. Math. Statist.* 31, 1015–1027 (1960).
2. Breiman, L., *Probability*, Addison-Wesley, Reading, Mass. (1968).
3. Gnedenko, B. V., and Kolmogorov, A. N., *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading, Mass. (1968).
4. Iscoe, I., Ney, P., and Nummelin, E., "Large deviations of uniformly reccurent Markov additive processes", *Adv. Appl. Math.* 6, 373–412 (1985).
5. Petrov, V. V., "On the probabilities of large deviations for sums of independent random variables", *Theory Probab. Appl.* 10 2, 287–298 (1965).
6. Szewczak, Z. S., "Lattice Edgeworth expansions in operator form", preprint.
7. Szewczak, Z. S., "A remark on large deviation theorem for Markov chain with finite number of states", *Theory Probab. Appl.* 50 3, 518–528 (2006).
8. Szewczak, Z. S., "Edgeworth expansions in operator form", *Statist. Probab. Lett.* 78 12, 1583–1592 (2008).
9. Szewczak, Z. S., "Large deviations in operator form", *Positivity* 12 4, 631–641 (2008).

A Dynamical Recurrent Neuro-Fuzzy Algorithm for System Identification

Dimitris C.Theodoridis[1], Yiannis S. Boutalis[1], and Manolis A. Christodoulou[2]

[1] Democritus University of Thrace, Xanthi, Greece, 67100
(e-mail: dtheodo@ee.duth.gr)
[2] Technical University of Crete, Chania, Crete, Greece, 73100
(e-mail: manolis@ece.tuc.gr)

Abstract. In this paper we discuss the identification problem which consists of choosing an appropriate identification model and adjusting its parameters according to some adaptive law, such that the response of the model to an input signal (or a class of input signals), approximates the response of the real system to the same input. For identification models we use fuzzy-recurrent high order neural networks. High order networks are expansions of the first order Hopfield and Cohen-Grossberg models that allow higher order interactions between neurons. In the present approach the HONN's used as approximators of the underlying fuzzy rules. New learning laws are proposed which ensure that the identification error converges to zero exponentially fast. There is a core idea in the proposed method: Several high order neural networks are specialized to work around fuzzy centers separating in this way the system in *neuro-fuzzy* subsystems which are associated with a number of fuzzy rules.
Keywords: Neuro-Fuzzy Systems, Identification, Gradient Descent, Pure Least Lquares.

1 Introduction

The purpose of this paper is to present the design, analysis, and simulation of algorithms that can be used for online parameter identification of continuous time plants.

It has been established that neural networks and fuzzy inference systems are universal approximators [2], [4], [9],i.e., they can approximate any nonlinear function to any prescribed accuracy provided that sufficient hidden neurons and training data or fuzzy rules are available. Recently, the combination of these two different technologies has given rise to fuzzy neural or *neuro-fuzzy* approaches, that are intended to capture the advantages of both fuzzy logic and neural networks. Numerous works have shown the viability of this approach for system modeling [3],[6].

In this paper, we present new adaptive algorithms for identifying nonlinear systems using neural networks and fuzzy logic. The neural architecture that is used is of the high order neural network (HONN) form and the fuzzy logic contribution to the algorithm is an estimate of the output fuzzy centers.

In the present approach the HONN's are used as approximators of the underlying fuzzy rules. Therefore, the required a-priori information obtained by linguistic information or data is very limited. The parameter identification is then easily addressed by Center-HONN's, based on the linguistic information regarding the structural identification of the output part and from the numerical data obtained from the actual system to be modeled.

The paper is organized as follows. Section 2 gives the overall scheme of the *neuro-fuzzy* model while Section 3 presents its approximation capabilities. The learning algorithms for parameter identification and the weight updating laws derived from them are demonstrated in Section 4. Finally, Section 5 presents simulations and comparisons with adaptive neural network representations, while Section 6 concludes the work.

2 Neuro-Fuzzy Model

Let us consider a nonlinear function $f(x,u)$, where $f : R^{n+m} \rightarrow R^n$ is a smooth vector field defined on a compact set $\Psi \subset R^{n+m}$, with input space $u \in U \subset R^m$ and state - space $x \in X \subset R^n$. Also, we assume that its i/o relation being governed by the following equation

$$\dot{x}_i(t) = f_i(x(t), u(t)) \tag{1}$$

where $f_i(\cdot)$, $i = 1, 2, ..., n$, is a continuous function and t denotes the temporal variable.

Assumption 1. Notice that since $\Psi \subset \Re^{n+m}$ then Ψ is closed and bounded set. Also, it is noted that even if Ψ is not compact we may assume that there is a time instant T such that $(x(t), u(t))$ remain in a compact subset of Ψ for all $t < T$; i.e. if $\Psi_T := \{(x(t), u(t)) \in \Psi, t < T\}$. The interval Ψ_T represents the time period over which the approximation is to be performed.

Following the notation of [8] the above Eq. (1) can be approximated by

$$\hat{f}_i(x(t), u(t)) = -a_i \hat{x}_i + \sum_{p=1}^{q} \bar{x}_{f_i}^p \cdot \left(\sum_{l=1}^{k} w_{f_i}^{pl} \cdot s_l(x(t), u(t)) \right) \tag{2}$$

where $a_i > 0$, $\bar{x}_{f_i}^p$ is the $p-th$ fuzzy center of the $i-th$ state variable and the summation is carried over all the available fuzzy rules. The above equation can be rewritten in a more compact form including all the state dynamics as

$$\dot{\hat{x}} = A\hat{x} + X_f W_f s_f(x, u) \tag{3}$$

where A is a $n \times n$ stable matrix which for simplicity can be taken to be diagonal as $A = diag[-a_1, -a_2, ..., -a_n]$, X_f is a matrix containing the centers of the partitions of every fuzzy output variable of $f(x,u)$ and $s_f(x,u)$ is a vector containing high order combinations of sigmoid functions of the state

x and control input u. Also, W_f is a matrix containing respective neural weights according to (2) and (3). For notational simplicity we assume that all output fuzzy variables are partitioned to the same number, q, of partitions. The exact definition of the above matrices are given detailed in [8].

From the above definitions and Eq. (2) it is obvious that the accuracy of the approximation of $f(x,u)$ depends on the approximation abilities of HONN's and on an initial estimate of the centers of the output membership functions. These centers can be obtained by experts or by off-line techniques based on gathered data.

3 Approximation capabilities

The approximation problem consists of determining whether by allowing enough high order connections and fuzzy centers, there exist weights W_f, such that the F-RHONN model could approximate the input-output behavior of an arbitrary dynamical system of the form (1).

In order to have a well-posed problem, we assume that f_i is continuous and satisfies a local Lipschitz condition such that (1) has a unique solution in the sense of Caratheodory [1]. Based on the above assumptions we obtain the following theorem.

Theorem 1. *Suppose that the system (1) and the model (3) are initially at the same state $\hat{x}(0) = x(0)$, then for any $\varepsilon > 0$ and any finite $T > 0$, there exists an integer k, a matrix $W_f^\star \in R^{n\times q\times k}$ and appropriately selected fuzzy output modified centers $\bar{x}_{f_i}^p$ such that the state $\hat{x}(t)$ of the Fuzzy-RHONN model (3) with k high order connections, weight values $W_f = W_f^\star$ and center values X_f satisfies*

$$\sup_{0\le t\le T} |\hat{x}(t) - x(t)| \le \varepsilon.$$

Proof. The dynamic behavior of the Fuzzy-RHONN model is described by (3). Adding and subtracting Ax, (1) is rewritten as

$$\dot{x} = Ax + g(x,u) \tag{4}$$

where $g(x,u) = f(x,u) - Ax$. Since $\hat{x}(0) = x(0)$, the state error $e = \hat{x} - x$ satisfies the differential equation

$$\dot{e} = Ae + X_f W_f s_f - g(x,u) \tag{5}$$

where $e(0) = 0$. By assumption, $(x(t), u(t)) \in \Psi$ for all $t \in [0,T]$, where Ψ is a compact subset of R^{n+m}.

Let $\Psi_e = \{(x,u) \in R^{n+m} : |(x,u) - (x_y, u_y)| \le \varepsilon, (x_y, u_y) \in \Psi\}$. It can be seen readily that Ψ_e is also a compact subset of R^{n+m} and $\Psi \subset \Psi_e$. In simple words Ψ_e is ε larger than Ψ, where ε is the required degree of approximation.

Since s_f is a continuous function, it satisfies a Lipschitz condition in Ψ_e, i.e. there is a constant l such that for all $(x_1, u), (x_2, u) \in \Psi_e$

$$|s_f(x_1, u) - s_f(x_2, u)| \leq l\,|x_1 - x_2|\,. \tag{6}$$

According to [7], we can prove that the function $X_f W_f s_f$ satisfies the conditions of Stone-Weirstrass Theorem and can approximate any continuous function over a compact domain. In what follows, we consider the learning problem of adjusting the weights adaptively, such that the Fuzzy-RHONN model identifies general dynamic systems.

4 Learning Algorithms for parameter identification

In this section we develop weight adjustment laws under the assumption that the unknown system is modeled exactly by a Fuzzy-RHONN architecture of the form (3).

4.1 Gradient descent

In developing this identification scheme we start again from the differential equation that describes the unknown system with no modeling error,

$$\dot{x}_i = -a_i x_i + \bar{x}_{f_i} W_{f_i}^{\star} s_f(x, u). \tag{7}$$

Based on (7), the identifier is now chosen as

$$\dot{\hat{x}}_i = -a_i \hat{x}_i + \bar{x}_{f_i} W_{f_i} s_f(x, u) \tag{8}$$

where W_{f_i} is again the estimate of the unknown matrix $W_{f_i}^{\star}$. In this case the state error $e_i = \hat{x}_i - x_i$ satisfies

$$\dot{e}_i = -a_i e_i + \bar{x}_{f_i} \tilde{W}_{f_i} s_f(x, u) \tag{9}$$

where $\tilde{W}_{f_i} = W_{f_i} - W_{f_i}^{\star}$.

The next theorem gives the filtered error Fuzzy-RHONN model with the gradient method for adjusting the weights.

Theorem 2. Consider the filtered error Fuzzy-RHONN model given by (8) whose weights are adjusted according to equation

$$\dot{W}_{f_i} = -\bar{x}_{f_i}^T e_i s_f^T P_i. \tag{10}$$

Then for $i = 1, 2, ..., n$, guarantees the following properties

1. $e_i, \tilde{W}_{f_i} \in L_\infty, \quad e_i \in L_2$
2. $\lim_{t\to\infty} e_i(t) = 0$
3. $\lim_{t\to\infty} \dot{W}_{f_i}(t) = 0$

Remark 1. The above theorem does not imply that the weight estimation error $\tilde{W}_{f_i} = W_{f_i} - W_{f_i}^*$ converges to zero. In order to achieve convergence of the weights to their correct value the additional assumption of persistent excitation needs to be imposed on the vector $s_f(x,u)$. In particular, $s_f \in R^k$ is said to be *persistently exciting* if there exist positive scalars β_1, β_2 and T such that for all $t \geq 0$

$$\beta_1 I \leq \int_t^{t+T} s_f(\tau) s_f^T(\tau)\ d\tau \leq \beta_2 I, \tag{11}$$

where I is the $k \times k$ identity matrix.

4.2 Pure Least Squares

The method is simple to apply and analyze in the case where the unknown parameters appear in a linear form, such as in eq. (7). The pure LS algorithm can be thought as a gradient algorithm with a time-varying learning rate and could be written as follows

$$\dot{W}_{f_i} = -\frac{\bar{x}_{f_i}^T e_i z_i^T P_i}{|\bar{x}_{f_i}|^2}, \quad W_{f_i}(0) = W_{f_0} \tag{12}$$

$$\dot{P}_i = -\frac{P_i z_i z_i^T P_i}{n_s^2}, \quad P_i(0) = P_0 \tag{13}$$

where z_i is a filtered version of s_f as will be described below, $n_s^2 \geq 1$ is a normalization signal designed to guarantee that $\frac{z_i}{n_s}$ is bounded. The property of n_s is used to establish the boundedness of the estimated parameters even when z_i is not guaranteed to be bounded. A straightforward choice for n_s in this paper is $n_s^2 = 1 + \alpha z_i^T z_i$, $\alpha > 0$. If z_i is bounded, we can take $\alpha = 0$. The following lemma is useful in the development of the adaptive identification algorithm which will be presented in this subsection.

Lemma 1. *The system described by Eq. (7) can be expressed as*

$$\dot{z}_i = -a_i z_i + s_f,\ z_i(0) = 0, \tag{14}$$

$$x_i = \bar{x}_{f_i} W_{f_i}^\star z_i + e^{-a_i t} x_i(0). \tag{15}$$

Proof. The proof described similarly in [7].

Using the above lemma the dynamical system is described by the following equation

$$x_i = \bar{x}_{f_i} W_{f_i}^\star z_i + \varepsilon_i, \tag{16}$$

where $\varepsilon_i = e^{-a_i t} x_i(0)$ is an exponentially decaying term which appears when a non zero initial state is applied. After ignoring the exponentially decaying term ε_i [7], the fuzzy-RHONN model can be written as

$$\hat{x}_i = \bar{x}_{f_i} W_{f_i} z_i. \tag{17}$$

The state error equation, after substituting (16), (17) becomes

$$e_i = \bar{x}_{f_i} \tilde{W}_{f_i} z_i - \varepsilon_i. \tag{18}$$

The cost function $J(W_{f_i})$ is chosen as

$$J(W_{f_i}) = \frac{\sum_{i=1}^{n} e_i^2}{2} = \frac{\sum_{i=1}^{n} \left[\left(\bar{x}_{f_i} W_{f_i} z_i - \bar{x}_{f_i} W_{f_i}^{\star} z_i \right) - \varepsilon_i \right]^2}{2}. \tag{19}$$

Depending on the optimization method we result to least squares method described by (12) and (13). A problem that may be encountered in the application of the LS's algorithm is that P_i may become arbitrarily small and thus slow down adaptation in some directions. This so-called problem can be prevented by using one of various modifications which prevent $P_i(t)$ of going to zero. One such modification is the so-called, where if the smallest eigenvalue of $P_i(t)$ becomes smaller than ρ_1 then $P_i(t)$ is reset to $P_i(t) = \rho_0 I$, where $\rho_0 \geq \rho_1 > 0$ are some design constants. In the following theorem we present the stability proof of the method.

Theorem 3. The pure LS algorithm given by (12), (13) guarantees the following properties

1. $e_i, \dot{W}_{f_i} \in L_2 \cap L_\infty, \quad W_{f_i}, P_i \in L_\infty$.
2. $\lim_{t\to\infty} W_{f_i}(t) = \bar{W}_{F_i}$,
   $where \bar{W}_{F_i}\, is\, a\, constant\, matrix.$
3. $If\ \frac{z_i}{n_s}\, is\, PE,\, then\ W_{f_i}(t) \to W_{f_i}^{\star}\, as\ t \to \infty.$

5 Simulation results

For simulation purposes, we are using the following dynamic equations appearing in the well known problem of the control of an inverted pendulum.

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = \frac{g \sin x_1 - \frac{mlx_2^2 \cos x_1 \sin x_1}{m_C+m}}{l\left(\frac{4}{3} - \frac{m\cos^2 x_1}{m_C+m}\right)} + \frac{\frac{cosx_1}{m_C+m}}{l\left(\frac{4}{3} - \frac{m\cos^2 x_1}{m_C+m}\right)} u \tag{20}$$

where $x_1 = \theta$ and $x_2 = \dot{\theta}$ are the angle from the vertical position and the angular velocity respectively. where $g = 9.8\ m/s^2$ is the acceleration due to gravity, m_c is the mass of the cart, m is the mass of the pole, and l is the half-length of the pole. We choose $m_c = 1\ kg$, $m = 0.1\ kg$, and $l = 0.5\ m$ in the following simulation.

In the simulations carried out the aim is not the control of the system but only to test the identification performance of the proposed scheme. Therefore, we use Eq. (20) as a means for deriving training data.

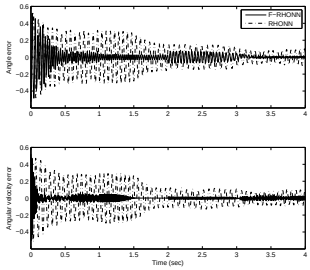


Fig. 1. Approximation errors of angle and angular velocity for RHONN (dashed line) and F-RHONN (solid line) approach.

It is our intention to compare the approximation abilities of the proposed dynamic neuro-fuzzy network (8) with RHONN's, [5] in approximating Eq. (20). For the RHONN's, we use the adaptive law which is described in, [5] (page 19) and for the proposed F-RHONN model we use the adaptive law which is described by Eq. (10). Numerical training data were obtained by using Eq. (20) with initial conditions $\left[x_1(0)\ x_2(0)\right] = \left[\frac{\pi}{6}\ -\frac{\pi}{6}\right]$, and the input signal has the following form

$$u(k) = 0.3\sin\left(\pi k/25\right) + 0.1\sin\left(\pi k/32\right) + 0.6\sin\left(\pi k/10\right) \tag{21}$$

The proposed neuro-fuzzy representation was chosen to use 4 output partitions of f_i, $\bar{x}_{f_1} = \left[1\ 2\ 4\ 5\right]$, $\bar{x}_{f_2} = \left[-9\ -5\ 8\ 9\right]$ and the number of high order sigmoidal terms (HOST) used in HONN's were chosen to be 3 $(s(x_1), s(x_2), s(u))$ up to first order. The initial weights are $W_{f_i}(0) = 0$ and the recurrent constant $a_i = 0.28$. Also, the parameters of the sigmoidal terms were chosen to be $\alpha_1 = 4.86$, $\beta_1 = 2$, $\gamma_1 = -0.75$ and the adaptive learning rate as $P_i = 7$, with $i = 1, 2, ..., n$.

The RHONN model given from [5] is constructed with the same initial weights, number of high order terms with these of F-RHONN approach and $a_i = 0.05$. The parameters of the sigmoidal terms were chosen to be $\alpha_2 = 2.89$, $\beta_2 = 2.95$, $\gamma_2 = -4.45$ and the adaptive learning rate as $P_i = 6.48$, with $i = 1, 2, ..., n$. Fig. 1 gives the evolution of identification errors for RHONN and F-RHONN models while and 2 presents the evolution of random weights for every approach, respectively. One can see that the dynamic neuro-fuzzy networks are more powerful than the simple neural networks.

6 Conclusion

This paper presents a new approach for *neuro-fuzzy* dynamical system identification based on high order neural network approximators (HONN's) and the fuzzy output partition. Instead of having a big RHONN we separate it

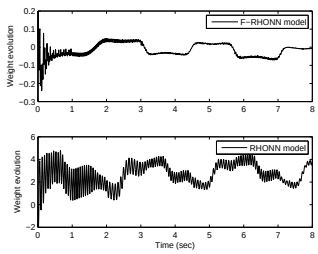


Fig. 2. Evolution of random weights in RHONN and F-RHONN approaches.

to smaller HONN's which finally gives a Fuzzy-RHONN. This leads to lower order of magnitude for high order terms while the weights energy becomes very restricted and difficult to drift into infinite. Future trends, are the expansion of the algorithm to robust systems and the better fuzzy selection of a certain number of HONN's determining the system.

References

1. J.K. Hale. *Ordinary Differential Equations*. Wiley-Interscience, New York, 1969.
2. K. Hornik, M. Stinchcombe, and H. White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2:359–366, 1989.
3. D. Kukolj and E. Levi. Identification of complex systems based on neural and takagi-sugeno fuzzy model. *IEEE Trans. Systems Man. Cybern. Part B: Cybernetics*, 34(1):272–282, 2004.
4. K. Passino and S. Yurkovich. *Fuzzy Control*. Addison, 1998.
5. G. A. Rovithakis. Performance of a neural adaptive tracking controller for multi-input nonlinear dynamical systems in the presence of additive and multiplicative external disturbances. *IEEE Trans. SMC-Part A*, 30:720–730, 2000.
6. G. A. Rovithakis, I. Chalkiadakis, and M. E. Zervakis. High-order neural network structure selection for function approximation applications using genetic algorithms. *IEEE Trans. SMC-Part B*, 34(1):150–158, February 2004.
7. G.A. Rovithakis and M.A.Christodoulou. *Adaptive Control with Recurrent High Order Neural Networks (Theory and Industrial Applications), in Advances in Industrial Control*. Springer Verlag London Limited, 2000.
8. D. C. Theodoridis, Y. S. Boutalis, and M. A. Christodoulou. A new neuro-fuzzy dynamical system definition based on high order neural network function approximators. In *European Control Conference ECC-09*, Budapest, Hungary, August 2009.
9. L. Wang. *Adaptive Fuzzy Systems and Control: Design and Stability Analysis*. Prentice Hall, NJ, 1994.

BERNSTEIN - VON MISES THEOREM IN BAYESIAN ANALYSIS OF COX MODEL

Jana Timková

timkova@karlin.mff.cuni.cz

Department of Statistics, Charles University, Prague
Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague

Summary: Although not well-known, the Bernstein-von Mises theorem (BvM) is a so-called bridge between bayesian and frequentist asymptotics. Basically, it states that under mild conditions the posterior distribution of the model parameter centered at the maximum likelohood estimator (MLE) is asymptotically equivalent to the sampling distribution of the MLE. This is a powerful tool especially when the classical asymptotics is tedious or impossible to conduct while bayesian asymptotic properties can be obtianed via MCMC. However, in semiparametric setting with presence of infinite-dimensional parameters, as is e.g. Cox model for survival data, the results regarding BvM are more difficult to establish but still not impossible. The proposed poster gives short overview of BvM results found in the survival analysis context.

Cox's regression model

Let us observe **a dataset** of following type:

- $X_i, i = 1, ..., n$, survival times of n independent individuals, $X_i \sim F_i$, F_i is a distribution function
- $C_i, i = 1, ..., n$ censoring random variables independent on X_i's
- $\boldsymbol{Z}_i \in \mathbb{R}^p, i = 1, ..., n$ a set of covariates describing the individuals

$\Longrightarrow$ the actual observed dataset is a *right-censored* set of triplets $(T_i, \delta_i, \boldsymbol{Z}_i)_{i=1}^n$ where $T_i = \min(X_i, C_i)$, $\delta_i = I(T_i = X_i)$. Denote $\tau = \max\{T_1, ..., T_n\}$.

We specify the Cox model is via particular form of **the hazard rate** which is assumed to satisfy

$$\Lambda_i(t) = \Lambda(t, \boldsymbol{Z}_i) = \int_0^t \exp\{\boldsymbol{\beta}^\top \boldsymbol{Z}_i\} d\Lambda(s), \quad i = 1, .., n, t \in [0, \tau],$$

with **two unknown parameters**:

- $\boldsymbol{\beta}$ is an unknown p-dimensional regression parameter
- $\Lambda(t), t \in [0, \tau]$ is an unknown cumulative hazard rate of a survival time of an individual with covariate being equal to 0

$\Longrightarrow$ with Λ being an functional (so, an infinitely-dimensional) parameter and $\boldsymbol{\beta}$ a finite-dimensional parameter inference on Cox model falls among **the semiparametric problems**.

Traditional approach to estimate the unknown parameters β and Λ ...

... is based on **partial likelihood** theory. Let $\boldsymbol{\beta}_0$ and Λ_0 be the true parameters. The estimator $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ is defined as a solution to the vector equation

$$\sum_{i=1}^n \left\{ \boldsymbol{Z}_i - \frac{\sum_{j:T_j \geq T_i} \boldsymbol{Z}_j \exp\{\boldsymbol{\beta}^\top \boldsymbol{Z}_j\}}{\sum_{k:T_k \geq T_i} \exp\{\boldsymbol{\beta}^\top \boldsymbol{Z}_k\}} \right\} = 0$$

The cumulative baseline hazard function $\Lambda(t)$ is estimated using the Breslow estimator

$$\hat{\Lambda}(t) = \sum_{i:T_i \leq t} \frac{\delta_i}{\sum_{j \in R_i} \exp\{\hat{\boldsymbol{\beta}}^\top \boldsymbol{Z}_j\}}$$

Theorem 1 (Frequentist asymptotics for Cox model, [1]) *Let the conditions A-D in [2] be fulfilled. Then the following is true:*

1.

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \Sigma(\boldsymbol{\beta}_0, \tau)^{-1})^\dagger$$

2.

$$\mathscr{L}(\sqrt{n}(\hat{\Lambda}(\cdot) - \Lambda_0(\cdot)) | \sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) = x) \xrightarrow{\mathscr{D}} W(V_0(\cdot) - xE_0(\cdot))^\dagger$$

on the space of functions continuous to the right and with limits to the left, $D[0, \tau]$. W denotes the standard Brownian motion.

$\dagger$ Terms $\Sigma(\boldsymbol{\beta}_0, \tau)$, $V_0(\cdot)$ and $E_0(\cdot)$ in limiting ditributions are matrix functions of unknown parameters $\boldsymbol{\beta}_0$ and Λ_0 and can be consistently estimated (for details see [1]).

Bayesian approach: Beta process as a prior for Λ

A prior process on the baseline d.f. F is a **process neutral to the right** if corresponding $\Lambda(t) = \int_0^t \frac{dF(s)}{1-F(s_-)}$ is a positive nondecreasing independent increment process (most common examples are Beta process and Gamma process) such that $\Lambda(0) = 0, 0 \leq \Delta\Lambda(t) \leq 1$ for all t w.p. 1 and either $\Delta\Lambda(t) = 1$ for some $t > 0$ or $\lim_{t\to\infty} \Lambda(t) = \infty$ w.p. 1.

PRIOR:

- Let the baseline c.d.f. F be, a priori, a process neutral to the right, such that the corresponding prior process for Λ is **a Beta process with mean Λ_{pr} and scale parameter c** that possesses following Lévy measure

$$\nu(dt, dx) = c(t)x^{-1}(1-x)^{c(t)-1}dx \; d\Lambda_{pr}(t)$$

- **let $\pi(\boldsymbol{\beta})$ be prior distribution for $\boldsymbol{\beta}$** which is continuous at $\boldsymbol{\beta}_0$ with $\pi(\boldsymbol{\beta}_0) > 0$, where $\boldsymbol{\beta}_0$ is true value of $\boldsymbol{\beta}$.

POSTERIOR:

- the posterior of Λ given β is a Lévy process with Lévy measure given in [4]

Theorem 2 (Bernstein - von Mises for Cox model, [2,3]) *Under certain conditions following hold*

1.

$$\lim_{n\to\infty} \int_{\mathbb{R}^p} |f_n(x) - \phi(x)| dx = 0$$

with probability 1, where f_n is the marginal posterior density of $x = \sqrt{n}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})$ and ϕ is the normal density with mean 0 and variance $\Sigma(\boldsymbol{\beta}_0, \tau)^{-1}$.

2.

$$\mathscr{L}(\sqrt{n}(\Lambda(\cdot) - \hat{\Lambda}(\cdot)) | \sqrt{n}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) = x, (T_i, \boldsymbol{Z}_i, \delta_i)_{i=1}^n) \xrightarrow{\mathscr{D}} W(V_0(\cdot) - xE_0(\cdot)) \qquad (1)$$

on the space of functions continuous to the right and with limits to the left, $D[0, \tau]$, with probability 1, as $n \to \infty$. W denotes the standard Brownian motion.

As a direct result of Theorem 2 we have the convergence of the joint posterior distribution

$$\mathscr{L}(\sqrt{n}(\Lambda(\cdot) - \hat{\Lambda}(\cdot), \boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) | (T_i, \boldsymbol{Z}_i, \delta_i)_{i=1}^n) \xrightarrow{\mathscr{D}} (W(V_0(\cdot) - XE_0(\cdot)), X)$$

with probability 1, as $n \to \infty$ on $D[0, \tau]$. X represents p-dimensional multivariate normal distribution with mean 0 and variance $\Sigma(\boldsymbol{\beta}_0, \tau)^{-1}$.

$\Longrightarrow$ Then similar result can be obtained for $\sqrt{n}(A(\boldsymbol{\beta}, \Lambda) - A(\hat{\boldsymbol{\beta}}, \hat{\Lambda}))$, where A is an arbitrary **Hadamard-differentiable functional of model parameters** $(\Lambda, \boldsymbol{\beta})$ (apply the functional delta method, see e.g. [5], Section 20.2).

Remark: Useful examples of Hadamard-differentiable functionals:

- baseline survival function $S(t) = \prod_{s \leq t}\{1 - d\Lambda(s)\}$
- survival function for i-th individual $S(t; \boldsymbol{Z}_i) = \prod_{s \leq t}\{1 - d\Lambda_i(s)\} = \prod_{s \leq t}\{1 - \exp\{\boldsymbol{\beta}^\top \boldsymbol{Z}_i\} d\Lambda(s)\}$
- median residual life $\eta_{t_0}^i$ so that $S(\eta_{t_0}^i; \boldsymbol{Z}_i)/S(t_0; \boldsymbol{Z}_i) = 0.5$, for $t_0 \in (0, \tau)$ and$i = 1, .., n$.

RESULT:

- Under mild conditions **the posterior distribution** of the model parameter **centered at the maximum likelohood estimator (MLE) is asymptotically equivalent to the sampling distribution of the MLE.**
- **So, what does it all mean?** Practically, the limit posterior distribution does not depend on prior and it is equivalent to frequential limiting distribution. In general, for functionals of parameters where asymptotic distribution is tedious to derive, the Bayesian approach gives a solution.

Illustration

We illustrate the model on simulated data from the hazard rate of form $\lambda(t; z) = 0.1t \exp\{1.5z\}$ where z was randomly generated from the normal distribution with mean 2 and standard deviation 1. For the prior we chose Beta process with parameters $\Lambda_{pr}(t) = 0.05t$ and $c(t) = 10e^{-0.05t}$, similarly as in [4]. We ran 15000 repetitions and used last 3000 for analysis of posterior.

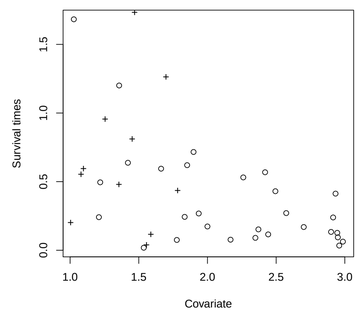


Figure 1.
Upper row from left to right:

1. The histogram of posterior sample of $\boldsymbol{\beta}$ with theoretical limiting distribution from Theorem 1 in red line.

2. Several iterations from the posterior sample of the baseline cumulative hazard rate $\Lambda(\cdot)$.

3. The posterior mean and 95% pointwise credibility band for the cumulative hazard rate with frequentists' Breslow estimator in red line alongside.

Bottom row from left to right:

4. Taking in mind the result in (1), following is true

$$\mathscr{L}\left(\sup_{t\in[0,\tau]} \frac{\sqrt{n}}{[\hat{V}_0(\tau) + \hat{E}_0(\tau)^\top \hat{\Sigma}(\hat{\beta}, \tau)\hat{E}_0(\tau)]^{1/2}} \left(\Lambda(\cdot) - \hat{\Lambda}(\cdot)\right) \middle| (T_i, \boldsymbol{Z}_i, \delta_i)_{i=1}^n \right) \xrightarrow{\mathscr{D}} \sup_{x\in[0,1]} W(x)$$

Several iterations from the posterior sample such transformation of the baseline cumulative hazard rate are plotted with dashed lines in black giving y such that $Pr\{\sup_{t\in[0,\tau]} C(t) > y\} = 0.05$, where with we $C(\cdot)$ denote the transformated process. Theoretical K-S confidence bands given by Brownian motion are in red dashed lines.

5. Several iterations from the posterior sample of the baseline survival function $S(t) = \prod_{s \leq t}\{1 - d\Lambda(s)\}$.

6. The posterior mean and 95% pointwise credibility bands for the baseline survival function $S(t)$.

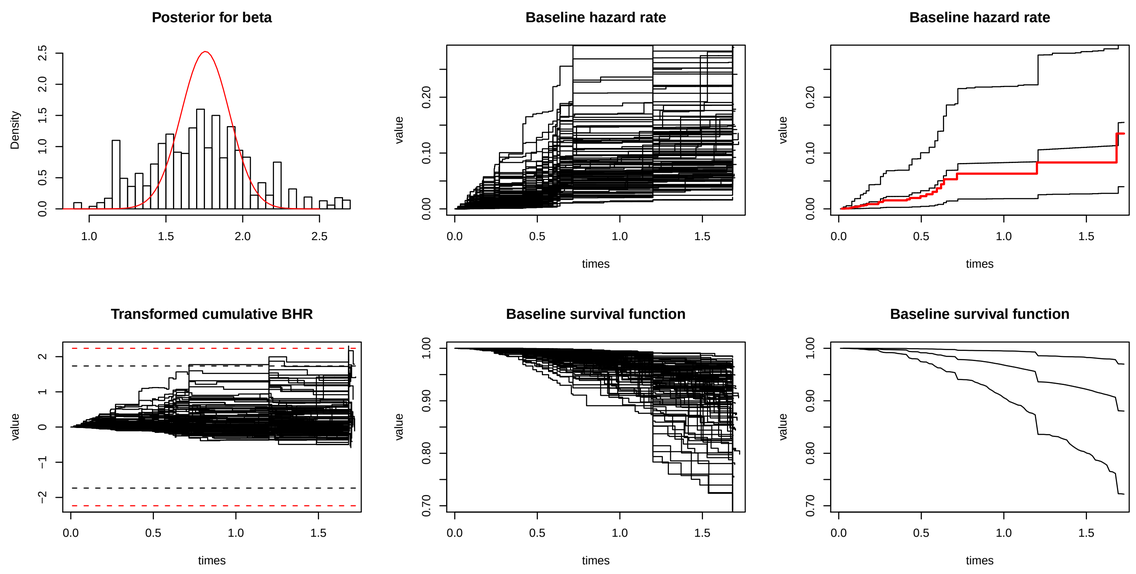


Figure 1: Results of Bayesian analysis of the simulated data using Beta process prior with parameters $\Lambda_{pr}(t) = 0.05t$ and $c(t) = 10e^{-0.05t}$. Posterior summaries on regression parameter $\boldsymbol{\beta}$ is $mean(\boldsymbol{\beta}) = 1.723$ and $sd(\boldsymbol{\beta}) = 0.33$. The frequentist' estimator is 1.757 with $sd = 0.39$.

Acknowledgement. The poster was supported by grants GA ČR 201/05/H007 and by GA AV IAA101120604.

References.

[1] Andersen, P. K., Gill R. D. (1982): Cox's regression model for counting processes: A large sample study, *Ann. Statist.* 10, pp. 1100 – 1120.
[2] De Blasi, P., Hjort, N. L. (2009): The Bernstein-von Mises theorem in semiparametric competing risks models, *J. Stat. Plan. and Infer.* Vol. 34, No. 4, pp. 1678 – 1700.
[3] Kim, Y.(2006): The Bernstein-von Mises theorem for the proportional hazard model, *Ann. Statist.* 34, no. 4, pp. 1678 – 1700.
[4] Laud, P. W., Damien, P., Smith, A. (1998): Bayesian nonparametric and covariate analysis of failure time data. In: *Dey, D., Muller, P., Sinha, D. (Eds.), Practical Nonparametric and Semiparametric Bayesian Statistics, Lecture Notes in Statistics* Vol. 133. Springer.
[5] Vaart, A. W. van der (2000): *Asymptotic statistics (Cambridge Series in Statistical and Probabilistic Mathematics)* . Cambridge University Press.

SMTDA 2010: Stochastic Modeling Techniques and Data Analysis
International Conference, *Chania, Crete, Greece, 8 - 11 June 2010*

Branching process and Monte Carlo simulation for solving Fredholm integral equations

Kianoush Fathi Vajargah, Fatemeh Kamalzade

and Farshid Mehrdoust

Department of statistics, Islamic Azad University, North Branch, Tehran, Iran

fathi_kia10@yahoo.com, fa.kamalzadeh@gmail.com

Department of Mathematics, University of Guilan, Rasht, Iran

fmehrdoust@guialn.ac.ir

Abstract: In this paper we establish a new method for solving nonlinear Fredholm integral equations; however the type of one dimension nonlinear Fredholm integral equations were solved previously by Albert [1]. Now thinking about high dimension of integral equation is a cause of finding a marvelous relationship between branching process and Monte Carlo simulation, although this method require the optimum probability, integral simulation, and Monte Carlo algorithm.

Key words: Monte Carlo simulation, Markov chain, Fredholm integral equations, Branching process

1 Introduction

So many different ways exist for solving Fredholm integral equations, but the main point is that they have an acceptable answer only in low dimensional integral equations specially in one dimension, the major problem is started just in high dimensional of integral equation therefore accompanied by this problem the error will increase. Therefore it seems

that Monte Carlo method considered is true [1]. In probability theory, a branching stochastic process is a Markov process that models a population in which each individual in generation n produces some random number of individuals in n+1.
Consider the following function

$$j(u) \equiv (g,u) = \int_G g(x)u(x)dx \qquad (1)$$

Where the domain $G \subset R^d$ and the point $x \equiv (x_1, x_2, ..., x_n) \in G$ is a point Euclidean space and g(x), u(x) belong to Banach space, mark that g(x) has been supposed Dirac-delta due to the special usage of this function is sampling from probability density function .
The full model of Fredholm integral equation as follow:

$$u(\mathrm{x}) = f(x) + \lambda \iint_G \ldots \int_G k(x, y_1, y_2 \ldots, y_m) \prod_{i=1}^{m} u(x_i) \prod_{i=1}^{m} dx_i \qquad (2)$$

Now we consider this model for double integral equation of the second kind

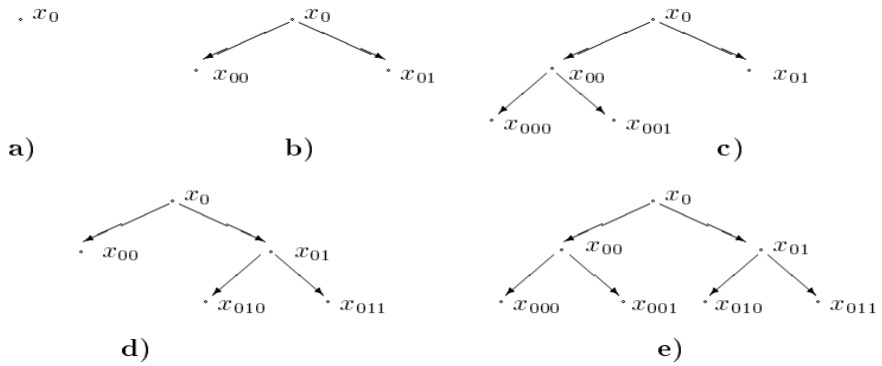
$$u(\mathrm{x}) = f(x) + \lambda \iint_G k(x, y, z)u(y)u(z)dydz \qquad (3)$$

The equation
(3) converges when satisfying:

2 A Monte Carlo method for solving nonlinear Fredholm integral equations

We will explain that finding a relation between branching process, and simulating it to solving nonlinear integral equation is the solution of this problem. Consider the model of branching process that just has been divided

to two branches. It begins with one point (x_0) then x_0 generate the next generation x_{00}, x_{01}, the generating particles behave at the next moment as the initial one and etc, and this process continue until generating all points. The particle from the k-th generation has the following form
$v[k+l]=(v[k],L)$, $L=0,1,...,m-1$.



Each particle begins generating with probability $p_m(x_0)$ and die out with probability h(x) such that

$$p_m(x_0)=1-h(x_0)$$

This probability is initial probability of each step and it has direct relation with die out probability, Moreover transition probability is

1. $p_m(x)\geq 0$

2. $p(x_0,x_{00},...,x_{0m-1})\geq 0$

3. $\int...\int p(x_0,x_{00},...,x_{0m-1})\prod_{i=0}^{m} dx_{0i}=1$

Associated with the sample path $x_0 \longrightarrow x_1 \longrightarrow x_2 \longrightarrow \ldots \longrightarrow x_n$ where n is given integer number, the Markov chain is being defined by:

$$\Gamma_n(h) = \frac{h(x_0)}{p(x_0)} \sum_{m=0}^{n} w_m f(x_m)$$

where

$$w_m = w_{m-1} \frac{k(x_{m-1}, x_m)}{p(x_{m-1}, x_m)}$$

Now we fit the branching process as we mention in table # to Fredholm integral equation .This process such an iterative process, therefore we choose importance sampling in this case.

$$u_0(x_0) = f(x_0)$$
$$u_1(x_0) = f(x_0) + \iint k(x_0, x_{00}, x_{01}) f(x_{00}) f(x_{01}) dx_{00} dx_{01}$$
$$u_2(x_0) = f(x_0) + \iint k(x_0, x_{00}, x_{01}) f(x_{00}) f(x_{01}) dx_{00} dx_{01} +$$
$$\iint k(x_0, x_{00}, x_{01}) f(x_{01}) \times \iint k(x_{00}, x_{000}, x_{001}) f(x_{000}) f(x_{001}) dx_{00} dx_{01} +$$
$$+ \iint k(x_0, x_{00}, x_{01}) f(x_{01}) \iint k(x_{01}, x_{010}, x_{011}) f(x_{010}) f(x_{011}) dx_{010} dx_{011}$$
$$\times \iint k(x_{00}, x_{000}, x_{001}) f(x_{000}) f(x_{001}) dx_{00} dx_{01}$$

Definition: Full tree with L generations is called the tree Γ_L where the dying out of particles is not visible from zero to L-1-sth generation, but all the generation particles of the L-th generation die out.

If the process has been stopped at the initial point then $u_0(x_0) = f(x_0)$, Therefore the Monte Carlo estimation is

$$\theta_g(\gamma_0) = \frac{g(x_0) f(x_0)}{p_0(x_0) h(x_0)}$$

The full model of Monte Carlo estimator is

$$\theta_g(\gamma_0)=\frac{g(x_0)}{p_0(x_0)}\times\frac{k(x_0,x_{00},x_{01})}{p_2(x_0)p(x_0,x_{00},x_{01})}\times\frac{k(x_{00},x_{000},x_{001})}{p_2(x_{00})p(x_{00},x_{000},x_{001})}$$
$$\times\frac{k(x_{01},x_{010},x_{011})}{p_2(x_{01})p(x_{01},x_{010},x_{011})}\times\frac{f(x_{000})f(x_{001})f(x_{010})f(x_{011})}{h(x_{000})h(x_{000})h(x_{000})h(x_{000})}$$

Theorem: The mathematical expectation of the random variable $\theta_g(\gamma_0)$ is equal to function $J(u_L)$.

$$E\,\theta_{[g]}(\frac{\gamma}{\Gamma_L})=J(u_L)=(g,u_L)$$

Proof: See [4].

Lemma1: The transition frequency function

minimizes u(x) $p(x,y_{1,\dots},y_m)=\dfrac{\left|k(x,y_{1,\dots},y_m)\prod_{i=1}^{m}u(y_i)\right|}{\int\dots\int\left|k(x,y_{1,\dots},y_m)\prod_{i=1}^{m}u(y_i)\right|\prod_{i=1}^{m}dy_i}$

$u(x)=\hat{u}(x)$ for any x from G, and min

Proof: Refer to [2]

Lemma2: The initial frequency function

$$p_0(x_0)=\frac{|g(x_0)\phi(x_0)|}{\int g(x_0)\phi(x_0)dx_0}$$

Minimizes the functional $\int g^2(x_0)u^2(x_0)p^{-1}(x_0)$ the minimizes of this functional is equal to $(\int|g(x_0)u(x_0)|dx_0)^2$.

Consider the Fredholm integral equation (3), It can suppose that the kernel of integral is separable without losing any information.

Proof: See to [2].

4 Numerical Example

Here, we present the performances of the above algorithm to obtain the unique solution $u(x)=1$ of the following integral equation. The experimental results for three different transition density function $(0.25, 0.5, x/4)$, are outlined in Table 1.

$$u(x)=1-0.1667x+0.0094\iint_D x(8y-z)u(y)u(z)dydz$$

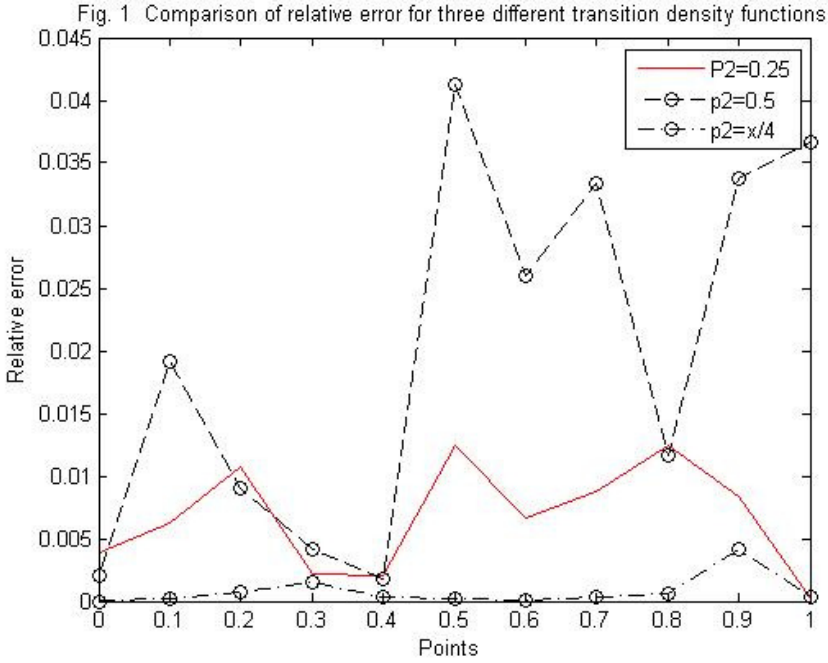
where $D=[0,1]\times[0,1]$ and $x\in$ □ .

Table 1
MC solution and relative for three transition density functions $(N=5000)$

Transition density functions	$p_2(x)=0.25$		$p_2(x)=0.5$		$p_2(x)=x/4$	
x	MC solution	Relative error	MC solution	Relative error	MC solution	Relative error
0.0	0.9917	0.0083	0.9980	0.0020	1.00	0.00
0.1	1.0063	0.0063	1.0192	0.0192	0.9999	1.459e-4
0.2	0.9893	0.0107	1.0090	0.0090	1.0007	7.417e-4
0.3	1.0022	0.0022	1.0041	0.0041	1.0015	0.0015
0.4	1.0021	0.0021	0.9982	0.0018	1.0003	2.647e-4
0.5	1.0125	0.0125	1.0412	0.0412	0.9998	1.869e-

						4
0.6	1.0067	0.0067	1.0260	0.0260	0.9999	1.314e-4
0.7	1.0088	0.0088	1.0334	0.0334	1.0004	3.550e-4
0.8	1.0125	0.0125	1.0116	0.0116	1.0006	5.672e-4
0.9	1.0084	0.0084	1.0337	0.0337	1.0042	0.0042
1.0	1.0002	1.92e-4	1.0367	0.0367	0.9996	3.741e-4

Fig. 1 Comparison of relative error for three different transition density functions



5 Conclusion and future study

We have proposed that it is possible to solve Fredholm and by extension, Volterra and other related equations of the second kind by using branching process and an appropriately defined distribution.

6 References

1. Albert, G.E , A general theory of stochastic estimations of the Newmann series for solution of certain Fredholm integral equations and related series in : M.A. Meyer(Ed.), Symposium of Monte Carlo method, Wiely, New York, 37-46 (1956).

2. Dimov, I.T., Minimization of the probable error for Monte Carlo methods. , Application of Mathematics in Technology, Differential equations and applications, 161-164 (1987).

3. Dimov, I.T., Gurov T., Monte carlo algorithm for solving integral equations with polynomial non-linearity. Parallel implemention , Pliska Studia Mathemaica Bulgaria ,13 , 117-132 (2000).

4. Hammersley, J.M., and Handscomb, D.C., Monte Carlo methods, John Wiely and Sons inc., New York ,London, Sydney ,Methuen,(1964).

5. Mikhailov, G.A., some problems in the theory of the Monte Carlo methods, Wiley, New York, 37-46 (1956).

6. Sobol, J.M., Computational methods of the Monte Carlo, Nauka, Moscow,(1973).

On Energy Based Cluster Stability Criterion

Z. Volkovich[1], D. Toledano Kitai[2] and R. Avros[3]

[1] Ort Braude College of Engineering
Software Engineering Department, Karmiel, Israel
(e-mail: vlvolkov@braude.ac.il)
[2] Ort Braude College of Engineering
Software Engineering Department, Karmiel, Israel
(e-mail: dvora@braude.ac.il)
[3] Ort Braude College of Engineering
Software Engineering Department, Karmiel, Israel
(e-mail: r_avros@braude.ac.il)

Abstract. This article proposes an approach for the cluster stability evaluation. This method adopts a physical point of view where a physical magnitude of samples mixing within clusters is considered. Samples closeness is quantified by the relative potential energy between items belonging to different samples inside the clusters. Actually, the two-sample energy test statistic of Zech and Aslan[20] based upon this perception is employed. The partition merit is represented by the worst cluster corresponding to the maximal potential energy value. To ensure readiness of the proposed model and to decrease the uncertainty in the model, we draw many pairs of samples for each given number of clusters and construct an empirical distribution of the inner clusters potential energy corresponding to the partitions created within the samples. Among all those distributions, one can expect that the true number of clusters can be characterized by the empirical distribution which is most concentrated at the origin. Numerical experiments, provided by means of the proposed methodology, demonstrate high ability of the approach.
Keywords: Cluster analysis, Clustering, Partitioning, Unsupervised learning, Cluster stability, Two-sample energy test.

1 Introduction

The cluster analysis encompasses algorithms and methods intended to categorize similar kind of objects in a manner that the connection measure between two objects is maximal if they fit in the same group and minimal otherwise. Cluster methodology is used in many fields such as machine learning, data mining and bioinformatics. The basic premise is that there is a similarity between elements of data collection which reflects a natural division into groups. An iterative clustering algorithm, proposed to divide a datum into groups, commonly requires as an input a suggested number of clusters. The problem emerging difficulty is associated with a number of meaningful clusters. Particularly, determining the true number of clusters is known as "ill posed" cluster problem (Jain and Dubes[11] and Gordon[8]). For instance,

the "right" number of clusters can be affected by the scale of data measurement (see e.g., Chakravarthy and Ghosh[4]). A procedure anticipated to estimate the true number of clusters is typically applied to compare the possible number of clusters in a given area. The amount of clusters that yields the best specified score is accepted as a true number. Many methods are known to handle the cluster validation problem. Until now, none of them has been agreed as superior. Two well known methodologies can be reminded here.

For example, the so-called "elbow" criterion is used in many variations to indicate the "true" number of clusters (see e.g., Dunn[6], Hubert and Scultz[10], Calinski and Harabasz[3], Hartigan[9], Krzanowski and Lai[12], Sugar and James [17], Gordon[7], Milligan and Cooper[16] and Tibshirani *et al.*[18]). Within the framework of the stability concept, partitions goodness is estimated by low variability amid the repeated cluster solutions obtained for the same dataset (see e.g. Lange *et al.*[14], Cheng and Milligan[5], Levine and Domany[15], Ben-Hur *et al.*[1] and Ben-Hur and Guyon[2]).

In the current article, we purpose a method for the study of cluster stability. This method adopts a physical point of view. Such standpoint suggests using a physical magnitude of samples mixing within clusters constructed by means of a clustering algorithm. We quantify samples closeness by the relative potential energy between items belonging to different samples for each one of the clusters. This potential energy is closely linked with a "gravity" force between two samples. If the samples within each cluster are well mingled, this quantity is sufficiently small. As known from electrostatics, if the sizes of the samples grow to infinity, then the total potential energy of the pooled samples, tends to zero, in the case of the samples drawn from the same population. The Two-Sample Energy test has been constructed based upon this perception (Zech and Aslan[20]). The statistic of the test measures the potential energy of the combined samples. Actually, we use this function as a characteristic of clustered samples similarity. The partition merit is represented by the worst cluster corresponding to the maximal potential energy value. To ensure readiness of the proposed model and to decrease the uncertainty of the model, we draw many pairs of samples for a given number of clusters and construct an empirical distribution of the potential energy corresponding to the partitions created within the samples. Among all those distributions, one can expect that the true number of clusters can be characterized by the empirical distribution which is most concentrated at the origin. Numerical experiments, provided by means of the proposed methodology, demonstrate high ability of the approach.

2 The Two-Sample Energy Test

Let $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots, y_m\}$ be two samples of independent random elements belonging to the Euclidean space $\mathbb{R}^d$, distributed

according to F and G, respectively. The two-sample problem examines the hypothesis

$$H_0 : F(x) = G(x)$$

against the general alternative

$$H_1 : F(x) \neq G(x),$$

when the distributions F and G are unknown.

The *Two-Sample Energy test* for this problem considers the sample X as a system of positive charges equal in value to $1/n$ and the second sample Y as a system of negative charges equal in value to $-1/m$. These charges provide the total normalized charge of each sample equal to 1. According to the one-over-distance law in electrostatics, it can be concluded that if the samples have the same distribution then the total energy of the united sample asymptotically neglects.

Let us denote by $\Phi_{X,R}$ the energy of a charged sample X. This value is calculated:

$$\Phi_{X,R} = \frac{1}{|X|^2} \sum_{i<j}^{|X|} R\left(|x_i - x_j|\right),$$

where the function R is suggested to be a continuous, monotonic decreasing function of the Euclidean distance between the charges. Correspondingly, the interaction energy of two samples X and Y is:

$$\Phi_{X,Y,R} = -\frac{1}{|X|\,|Y|} \sum_{i=1}^{|X|} \sum_{j=1}^{|Y|} R\left(|x_i - y_j|\right)$$

The test statistic $\Psi_{X,Y,R}$ is defined as:

$$\Psi_{X,Y,R} = \Phi_{X,R} + \Phi_{Y,R} + \Phi_{X,Y,R}. \tag{1}$$

The common cases of R are:

- $R\left(r\right) = -\ln\left(r\right)$;
- $R\left(r\right) = \frac{1}{r^\alpha}$, $\alpha > 0$;
- $R\left(r\right) = e^{-r^\alpha}$, $\alpha > 0$.

The choice $R\left(r\right) = -\ln\left(r\right)$ ensures that the test is scale invariant.

3 Method Description

Let us consider a finite subset $X = \{x_1, x_2 \ldots, x_n\}$ of the d-dimensional Euclidean space $\mathbb{R}^d$. For a given set $S \subset X$, we create a partition $\Pi_k(S)$ of S by splitting it into k nonempty and disjointed sub-groups $\{\pi_j\left(S\right)\}_{j=1}^{k}$ called *clusters*. The union of these clusters is S.

As mentioned earlier, we are going to describe cluster stability by means of the sampling procedure steadiness. We assume that there is a cluster stable structure which can be reflected by an appropriate clustering algorithm $\Delta(X,k)$ where X is the clustered dataset, and k is the suggested number of clusters. The algorithm output is a partition of the set X into k clusters. For this purpose, we draw pairs of samples $S_{j,1}, S_{j,2}$, having the same size m to assess the potential energy between their elements within the clusters. To do this we should know the occurrences of the samples within the clusters. Due to the fact that a cluster structure is latent, these occurrences have to be simulated. This task, as usual, meets the so-called cluster coordination problem. Namely, the same cluster could be differently assigned in various algorithms outcomes. In this paper we use a simulation method proposed by Volkovich *et al.*[19].

Let us consider the pool set

$$S_j = S_{j,1} \cup S_{j,2}$$

along with three partitions:

$$\begin{aligned} \Pi_k(S_j) &= \Delta(S_j,k), \\ \Pi_k^{(1)}(S_{j,1}) &= \Delta(S_{j,1},k), \\ \Pi_k^{(2)}(S_{j,2}) &= \Delta(S_{j,2},k). \end{aligned}$$

Each item $x \in S_j$ is located in the partition $\Pi_k(S_j)$ and in one of the other two partitions $\Pi_k^{(i)}(S_{j,i}), i = 1,2$. A correspondence of cluster tags between the partitions can be provided according to the minimal misclassification rate calculated over all data points. So, the labels in the clustered samples $\Pi_k^{(i)}(S_{j,i}), i = 1,2$ are altered, aiming at maximal coincidence with $\Pi_k(S_j)$:

$$\sigma_i^* = \underset{\sum_k}{\arg\min} \sum_{x \in X} I(\sigma(\alpha_{k,i}(x)) \neq \alpha_k^{(i)}(x)),\ i = 1,2,$$

where $I(\bullet)$ denotes the indicator function and $\alpha_{k,i}, \alpha_k^{(i)}$ are assignments defined by $\Pi_k(S_j)$ and $\Pi_k^{(i)}(S_{j,i}), i = 1,2$, correspondantly. $\sum_k$ is the set of all possible permutations of the set $\{1,...,k\}$. The Hungarian method [13] meets this goal by $O(k^3)$ complexity. After changing the cluster labels in $\Pi_k^{(i)}(S_{j,i}), i = 1,2$ with respect to σ_i^*, $i = 1,2$ the sets

$$S_{j,i}^{(t)} = \{x \in S_{j,i} | \alpha_{k,i}(x) = t),\ i = 1,2,\ t = 1,...,k,$$

can be interpreted as sub-samples creating the clusters. Following the two sample test energy methodology, we calculate in each group its inner potential energy $\Psi_{S_{j,1}^{(t)}, S_{j,2}^{(t)}, R};\ t = 1,...,k$, according to (1). This value characterizes the

cluster quality. Subsequently, the partition quality is represented by its worst cluster having maximal inner potential energy:

$$\Psi^*_{S_{j,1},S_{j,2},R} = \max_t \Psi_{S^{(t)}_{j,1},S^{(t)}_{j,2},R}. \tag{2}$$

Another possibility would be to determine the average of (2) among all k occurred clusters. However, the first possibility seems to be more stable.

Next, we consider the distributions of (2) constructed by a multitude of samples for number of clusters in the range $k = 2, \ldots, k^\star$, where $k^\star$ is the maximal number of tested clusters. In order to view those distributions in the same scale, we perform a normalization. In our approach, we divide the range of the distances values into g equal range subgroups and characterize the concentration by the frequency $N_{k,g}$ of the lowest subgroup which is expected to provide the greatest value of $N_{k,g}$ in the case of the true number of clusters.

An algorithm which implements this procedure, consists of the following steps:

1. Choose the parameters:
   (a) $k^\star$: maximal number of clusters to be tested,
   (b) J : number of the drawn sample pairs,
   (c) m : the samples size,
   (d) X : the data to be clustered,
   (e) Δ : a clustering algorithm,
   (f) R : a distance function,
   (g) g : number of range subgroups.

2. For $k = 2$ to $k^\star$
3. for $j = 1$ to J do
4. $S_{j,1} = sample(X, m), S_{j,2} = sample(X \backslash S_{j,1}, m)$;
5. Clustering $S_{j,1}$ by means of Δ;
   Clustering $S_{j,2}$ by means of Δ;
   Clustering $S_{j,1} \bigcup S_{j,2}$ by means of Δ;
   Solve the coordination problem;
6. Measure the potential energy inside the clusters;
7. Calculate the partition quality according to (2);
8. end for j;
9. Calculate $N_{k,g}$;
10. End for k.
11. The "true" $\widehat{k}$ is selected as the one which yields the maximal value of $N_{k,g}$.

$Sample(S, m)$ is a procedure which selects a random sample of size m from the set S, without replacement.

4 Numerical Experiments

We exemplify the described approach by means of the numerical experiments on synthetic and real datasets provided for 3 functions $R(r)$ mentioned in section 2. We choose $k^* = 7$, $J = 200$ and $m = 100$ in all tests and perform 10 trials for each experiment. The results are presented via the error-bar plots of $N_{k,10}$ within the trials. The sizes of the error bars equal to two standard deviations, found inside the trials. The spherical $k-$means algorithm is employed.

4.1 Synthetic Data

In the first example the datum is simulated as a mixture of 5 two-dimensional Gaussian distributions with independent coordinates owning the same standard deviation $\sigma = 0.35$. The components means are placed on the unit circle with the angular neighboring distance $2\pi/5$. The dataset contains 4000 items. The results shown in Fig. 1 demonstrate that for this combination of parameters and kernels, a five clusters structure is clearly indicated.

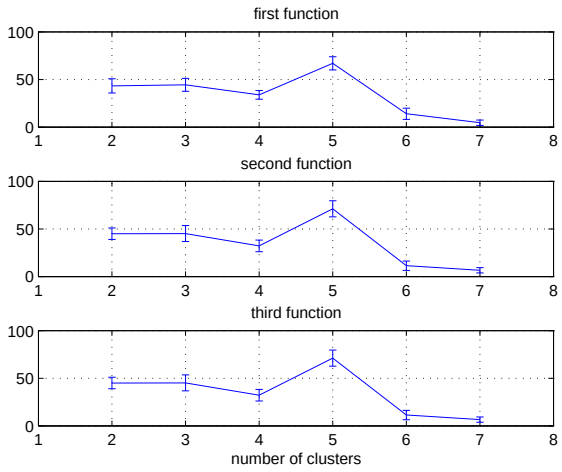


Fig. 1. Error-bar plots of $N_{k,g}$ for the five components simulated data.

4.2 Real-World Data

In the second example a real dataset is chosen from the text collection *http : //ftp.cs.cornell.edu/pub/smart/*.

This set includes 3 sub-collections, consisting of 1033 medical, 1460 information science and 1400 aerodynamics abstracts, correspondingly. We select the 600 "best" terms, following the common "bag of words" method and used the data representation by means of two leading principal components. The results presented in Fig. 2 show that the number of clusters is properly determined for all functions $R(r)$.

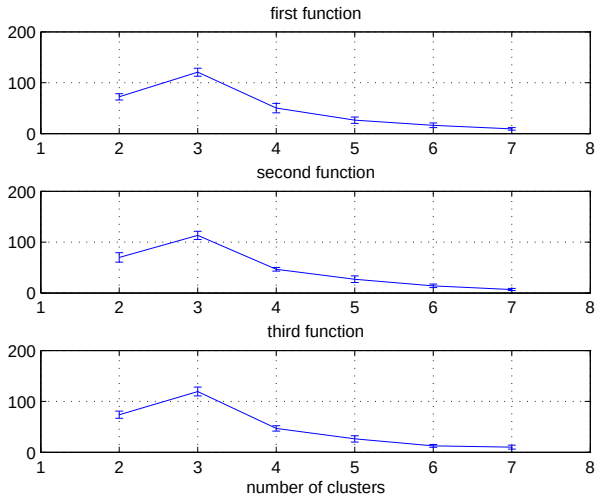


Fig. 2. Error-bar plots of $N_{k,g}$ the three text collection dataset.

5 Conclusion

We offer a new approach for detecting the true number of clusters based on a novel point of view where a physical magnitude of samples mixing within clusters characterizes the partition quality. Pairs of samples are drawn in order to construct an empirical distribution of the inner cluster potential defined by the worst partitions clusters. The most concentrated at the origin distribution indicates the true number of clusters. Numerical experiments, provided by means of the proposed methodology, demonstrate high ability of the approach.

References

1. Ben-Hur, A., Elisseeff, A. and Guyon, I., "A stability based method for discovering structure in clustered data", *Pacific Symposium on Biocomputing*, 6-17 (2002)
2. Ben-Hur, A. and Guyon, I., "Methods in Molecular Biology", 159–182, M.J. Brownstein and A. Khodursky, Humana press (2003)
3. Calinski, R. and Harabasz, J., "A dendrite method for cluster analysis", *Communications in Statistics*, 3, 1–27 (1974)
4. Chakravarthy, S.V. and Ghosh, J., "Scale-Based Clustering Using the Radial Basis Function Network", *IEEE Transactions on Neural Networks*, 7(5), 1250–1261 (1996)
5. Cheng, R. and Milligan, G.W., "Measuring the influence of individual data points in a cluster analysi", *Journal of Classification*, 13, 315–335 (1996)
6. Dunn, J.C., "Well Separated Clusters and Optimal Fuzzy Partitions", *Journal Cybern.*, 4, 95–104 (1974)
7. Gordon, A.D., "Identifying genuine clusters in a classification", *Computational Statistics and Data Analysis*, 18, 561–581 (1994)
8. Gordon, A.D., "Classification", *Chapman and Hall*, CRC, Boca Raton, FL (1999)
9. Hartigan, J.A., "Statistical theory in clustering", *J. Classification*, 2, 63–76 (1985)
10. Hubert, L. and Schultz, J., "Quadratic assignment as a general data-analysis strategy", *Br. J. Math. Statist. Psychol.*, 76, 190–241 (1976)
11. Jain, A. and Dubes, R., "Algorithms for Clustering Data", *Englewood Cliffs*, Prentice-Hall, New Jersey (1988)
12. Krzanowski, W. and Lai, Y., "A criterion for determining the number of groups in a dataset using sum of squares clustering", *Biometrics*, 44, 23–34 (1985)
13. Kuhn, K., "The hungarian method for the assignment problem", *Naval Research Logistics Quarterly*, 2, 83–97 (1955)
14. Lange, T., Roth, V., Braun, M. and Buhmann, J.M., "Stability-based validation of clustering solutions", *Neural Computation*, 15(6), 1299–1323 (2004)
15. Levine, E. and Domany, E., "Resampling Method for Unsupervised Estimation of Cluster Validity", *Neural Computation*, 13, 2573–2593 (2001)
16. Milligan, G. and Cooper, M., "An examination of procedures for determining the number of clusters in a data set", *Psychometrika*, 50, 159–179 (1985)
17. Sugar, C. and James, G., "Finding the Number of Clusters in a Data Set : An Information Theoretic Approach", *J. of the American Statistical Association*, 98, 750–763 (2003)
18. Tibshirani, R. , Walther, G. and Hastie, T., "Estimating the number of clusters via the gap statistic", *J. Royal Statist. Soc. B*, 63(2), 411–423 (2001)
19. Volkovich, Z., Barzily, Z. and Morozensky, L., "A statistical model of cluster stability", *Pattern Recognition*, 41(7), 2174–2188 (2008)
20. Zech, G. and Aslan, B., "New test for the multivariate two-sample problem based on the concept of minimum energy", *The Journal of Statistical Computation and Simulation*, 75(2), 109–119 (2005)

Non-Standard Behavior of Density Estimators for Functions of Independent Observations

U. U. Müller[1⋆], A. Schick[2⋆⋆], and W. Wefelmeyer[3]

[1] Department of Statistics, Texas A&M University
College Station, TX 77843-3143, USA
(email: uschi@stat.tamu.edu)
[2] Department of Mathematical Sciences, Binghamton University
Binghamton, NY 13902-6000, USA
(email: anton@math.binghamton.edu)
[3] Mathematical Institute, University of Cologne
50931 Cologne, Germany
(email: wefelm@math.uni-koeln.de)

Abstract. Densities of functions of two or more independent random variables can be estimated by local U-statistics. Frees (1994) gives conditions under which they converge pointwise at the parametric root-n rate. Giné and Mason (2007) give conditions under which this rate also holds in L_p-norms. We present several natural applications in which the parametric rate fails to hold in L_p or even pointwise.

1. The density estimator of a sum of squares of independent observations typically slows down by a logarithmic factor. For exponents greater than two, the estimator behaves like a classical density estimator.

2. The density estimator of a product of two independent observations typically has the root-n rate pointwise, but not in L_p-norms. An application is given to semi-Markov processes and estimation of an inter-arrival density that depends multiplicatively on the jump size.

3. The stationary density of a nonlinear or nonparametric autoregressive time series driven by independent innovations can be estimated by a local U-statistic (now based on dependent observations and involving additional parameters), but the root-n rate can fail if the derivative of the autoregression function vanishes at some point.

Keywords: Density estimator, Local U-statistic, Local von Mises statistic, Convergence rate, Autoregressive time series, Semi-Markov process.

1 Introduction

It is often of interest to estimate densities of known or unknown functions of independent observations. Consider for example a regression model $Y = r(X) + \varepsilon$ with independent error ε and covariate X. If we have independent observations (X_i, Y_i), $i = 1, \ldots, n$, then the density of the response Y could be estimated by a kernel estimator based on $Y_1, \ldots, Y_n$. However, a much

⋆ Supported by NSF Grant DMS 0907014.
⋆⋆ Supported by NSF Grant DMS 0906551.

better estimator is obtained if we exploit the independence of ε and X and write Y as a sum $r(X)+\varepsilon$ of independent random variables. Then the density p of Y can be estimated by a local von Mises statistic

$$\hat{p}(z) = \frac{1}{n^2} \sum_{i,j=1}^{n} k_b(z - \hat{r}(X_i) - \hat{\varepsilon}_j).$$

Here $k_b(z) = k(z/b)/b$ with kernel k and bandwidth b, $\hat{r}$ is some estimator of the regression function r, and $\hat{\varepsilon}_j = Y_j - \hat{r}_j(X_j)$ are the corresponding residuals. Under appropriate conditions, the estimator $\hat{p}(z)$ converges at the parametric rate $n^{1/2}$; see Støve and Tjøstheim, 2010 [19], Escanciano and Jacho-Chávez, 2010 [1], and, for nonlinear regression and with responses missing at random, Müller, 2010 [5]. It is the purpose of this review to indicate why such rates are possible, and to illustrate when they fail.

The most straightforward version of the problem is the following. Let $X_1, \ldots, X_n$ be independent real-valued observations with density f. We want to estimate the density p of some transformation $T(X_1, \ldots, X_m)$ of m of these observations, with m at least 2. Frees, 1994 [2] proposed as an estimator of $p(z)$ the local U-statistic

$$\hat{p}(z) = \frac{1}{\binom{n}{m}} \sum_{1 \le i_1 < \cdots < i_m \le n} k_b(z - T(X_{i_1}, \ldots, X_{i_m}))$$

with $k_b(x) = k(x/b)/b$ for a kernel k and a bandwidth b. He showed that this estimator can be pointwise $n^{1/2}$-consistent under some assumptions on f and T. Saavedra and Cao, 2000 [9] consider the function $T(X_1, X_2) = X_1 + \varphi X_2$. It is even possible to obtain $n^{1/2}$-consistency in various norms, together with functional central limit theorems in the corresponding spaces. Schick and Wefelmeyer, 2004 [11], 2007 [13] prove such results for transformations of the form $T(X_1, \ldots, X_m) = T_1(X_1) + \cdots + T_m(X_m)$ and $T(X_1, X_2) = X_1 + X_2$ in the sup-norm and in L_1-norms. Giné and Mason, 2007 [3] consider general transformations $T(X_1, \ldots, X_m)$ and obtain such results in the L_p-norms. Their results hold locally uniformly in the bandwidth. More general results applicable here are in Nickl, 2007 [6] and Nickl, 2009 [7].

These results are less generally valid than appears at first sight. In Section 2 we restrict attention to $m = 2$ and to transformations of the special form $T(X_1, X_2) = T_1(X_1) + T_2(X_2)$ and explain under which conditions the local U-statistic $\hat{p}(z)$ is asymptotically linear, $n^{1/2}$-consistent, and asymptotically normal. The rate is typically slower when, say, $T_1(y) = T_1(x) + c(y - x)^\nu + o(|y - x|^\nu)$ for y to the left or right of some point x, with $\nu \ge 2$. Then the density of $T_1(X)$ has a strong peak. Specifically, we consider $T_1(x) = T_2(x) = x^\nu$ and describe the rates of the local U-statistic. Then we discuss the two-sample case and applications to regression, to time series driven by independent innovations, and to renewal processes with multiplicative waiting times.

2 Results and Applications

Let $X_1, \ldots, X_n$ be independent real-valued observations with density f. An estimator for the density p of a transformation of the form $T(X_1, X_2) = T_1(X_1) + T_2(X_2)$ is the local U-statistic

$$\hat{p}(z) = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} k_b(z - T_1(X_i) - T_2(X_j)),$$

where $k_b(z) = k(z/b)/b$ for a kernel k and a bandwidth b. Suppose that $T_1(X)$ and $T_2(X)$ have densities g_1 and g_2. The estimator $\hat{p}(z)$ has the Hoeffding decomposition

$$\hat{p}(z) = p * k_b(z) + \frac{1}{n} \sum_{i=1}^{n} \big(g_1 * k_b(z - T_2(X_i)) - p * k_b(z) + g_2 * k_b(z - T_1(X_i)) - p * k_b(z) \big) + U(z),$$

where

$$U(z) = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} \big(k_b(z - T_1(X_i) - T_2(X_i)) - g_1 * k_b(z - T_2(X_i)) - g_2 * k_b(z - T_1(X_i)) + p * k_b(z) \big)$$

is a degenerate local U-statistic. We have

$$n(n-1)E[U^2(z)] \le 2E[k_b^2(z - T_1(X_1) - T_2(X_2))] = 2p * k_b^2(z)$$

and

$$p * k_b^2(z) = \frac{1}{b} \int p(z - bu) k^2(u)\, du \le \frac{\|p\|_\infty}{b} \int k^2(u)\, du.$$

If p is bounded and $\int k^2(u)\, du$ is finite, we obtain $U(z) = O_P(1/(nb^{1/2}))$, which is of order $o_P(n^{-1/2})$ if $nb \to \infty$. The Hoeffding decomposition then says that the centered local U-statistic $\hat{p}(z) - p * k_b(z)$ is approximated by a sum of two centered and smoothed empirical "estimators" of $p(z)$ (that involve the unknown densities g_1 and g_2). Under mild assumptions one can remove the smoothing; see e.g. Schick and Wefelmeyer, 2004 [11]. If p is Hölder with exponent α, then the bias $p * k_b(z) - p(z)$ is of order $o(n^{-1/2})$ if $nb^{2\alpha} \to 0$. This implies that $\hat{p}(z)$ is asymptotically linear,

$$\hat{p}(z) = p(z) + \frac{1}{n} \sum_{i=1}^{n} \big(g_1(z - T_2(X_i)) + g_2(z - T_1(X_i)) - 2p(z) \big) + o_P(n^{-1/2}). \quad (1)$$

If $E[g_1^2(z - T_2(X_2))]$ and $E[g_2^2(z - T_1(X_1))]$ are finite, then $\hat{p}(z)$ is $n^{1/2}$-consistent and asymptotically normal.

Remark 1. (Convolution of density estimators.) The density p has the convolution representation

$$p(z) = \int g_2(z-y)g_1(y)\,dy.$$

Therefore, it can also be estimated by a convolution of density estimators

$$\hat{g}_{conv}(z) = \int \hat{g}_2(z-y)\hat{g}_1(y)\,dy$$

with kernel estimator for $g_1(y)$ based on $T_1(X_1), \ldots, T_1(X_n)$,

$$\hat{g}_1(y) = \frac{1}{n}\sum_{i=1}^{n} k_b(y - T_1(X_i)),$$

and, correspondingly,

$$\hat{g}_2(y) = \frac{1}{n}\sum_{i=1}^{n} k_b(y - T_2(X_i)).$$

The estimator $\hat{g}_{conv}$ is asymptotically equivalent to $\hat{g}$. □

Remark 2. (Transform density estimator or transform observations.) Suppose that T_1, say, is strictly increasing and differentiable. Then the density of $T_1(X)$ at y is

$$g_1(y) = \frac{f(T_1^{-1}(y))}{T_1'(T_1^{-1}(y))}.$$

We obtain an alternative estimator of $g_1(y)$ by plugging in a kernel estimator for f,

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} k_b(x - X_i).$$

We expect that it depends on T_1 whether $\hat{g}_1(y)$ is better than

$$\tilde{g}_1(y) = \frac{\hat{f}(T_1^{-1}(y))}{T_1'(T_1^{-1}(y))}.$$

In the convolution representation $p(z) = \int g_2(z-y)g_1(y)\,dy$ we can use $\hat{g}_1$ or $\tilde{g}_1$. If T_2 is also strictly increasing and differentiable, we can combine $\hat{g}_1$ or $\tilde{g}_1$ with $\hat{g}_2$ or $\tilde{g}_2$. □

We now discuss cases in which $\hat{p}(z)$ is not $n^{1/2}$-consistent.

Remark 3. (Piecewise constant transformations.) The distribution of $T_1(X)$ does not always have a density. Suppose that T_1 is piecewise constant,

$$T_1(X) = \sum_{s=1}^{t} c_s \mathbf{1}[X \in I_s],$$

with $c_s \in \mathbb{R}$, and I_s, $s = 1, \ldots, t$, a partition of $\mathbb{R}$. If $T_2(X)$ has a density g_2, then $T_1(X_1) + T_2(X_2)$ has a density p that is a finite mixture of shifts of g_2,

$$p(z) = \sum_{s=1}^{m} a_s g_2(z - c_s)$$

with weights $a_s = P(X \in I_s)$. As soon as each interval I_s contains at least one observation, the constants c_s can be observed, and $p(z)$ can be estimated by

$$\hat{p}(z) = \sum_{s=1}^{t} \hat{a}_s \hat{g}_2(z - c_s),$$

where $\hat{a}_s = (1/n) \sum_{i=1}^{n} \mathbf{1}[X_i = c_s]$. The rate of $\hat{p}(z)$ equals the pointwise rate of $\hat{g}_2$. □

Even if T_1 and T_2 are not constant on any interval, $\hat{p}(z)$ can fail to be $n^{1/2}$-consistent. In the following we describe a situation in which $T_1(X)$ and $T_2(X)$ have densities, but $g_1(z - T_2(X))$ does not necessarily have finite variance. For notational simplicity, assume that f is supported on $(0, \infty)$, and set $T_1(x) = T_2(x) = x^\nu$ for some $\nu > 0$. Then $g_1 = g_2 = g$ with

$$g(y) = \frac{1}{\nu} y^{1/\nu - 1} f(y^{1/\nu}),$$

and the stochastic expansion (1) of $\hat{p}(z)$ specializes to

$$\hat{p}(z) = p(z) + \frac{2}{n} \sum_{i=1}^{n} g(z - X_i^\nu) + o_P(n^{-1/2}). \tag{2}$$

In the theorems below, we take the kernel k to be continuously differentiable with support $[-1, 1]$. We also assume that f is bounded. First let $\nu < 2$. Then g is square-integrable, and $g(z - X^\nu)$ has finite variance. By the arguments of Schick and Wefelmeyer, 2004 [11] and 2009 [17] or Giné and Mason, 2007 [3], we have the following result.

Theorem 1. *Let $\nu < 2$. Suppose the density f is of bounded variation and $f(0+)$ is positive. Let $b \sim (\log n)^{1/2}/n$. Then $\hat{p}(z)$ has the stochastic expansion (2), and*

$$n^{1/2}\big(\hat{p}(z) - p(z)\big) \Rightarrow N\big(0, 4 \operatorname{Var}(g(z - X^\nu))\big).$$

For $\nu = 2$, square-integrability of g fails just barely, resulting in a rate for $\hat{p}(z)$ that is only slightly worse than $n^{-1/2}$. More precisely, Schick and Wefelmeyer, 2009 [16] prove the following result.

Theorem 2. *Let $\nu = 2$. Suppose f is of bounded variation, and $f(0+)$ and $g(z-)$ are positive. Let $b \sim (\log n)^{1/2}/n$. Then*

$$\Big(\frac{n}{\log n}\Big)^{1/2} \big(\hat{p}(z) - p(z)\big) \Rightarrow N(0, f^2(0+) g(z-)).$$

For $\nu > 2$, the rate of $\hat{p}(z)$ is of order $n^{-1/\nu}$ if f is of bounded variation and $f(0+)$ and $g(z-)$ are positive. Faster rates are possible under additional smoothness assumptions on p at z.

Theorem 3. *Let $\nu > 2$. Suppose f is of bounded variation, and $f(0+)$ and $g(z-)$ are positive. Let $b \sim 1/n$. Then*

$$\hat{p}_b(z) - p(z) = O_P(n^{-\beta}).$$

Even in the case $\nu \geq 2$, the estimator $\hat{p}(z)$ can be $n^{1/2}$-consistent if $g(z-) = 0$ since this works against the peak of g at 0 in the representation $p(z) = g * g(z)$. For details we refer to Schick and Wefelmeyer, 2009 [16] and 2009 [17].

We will now briefly discuss possible applications of the above results.

Remark 4. (Several samples.) The above results carry over to m-sample cases. We restrict attention to $m = 2$. Suppose $X_1, \ldots, X_n$ and $Z_1, \ldots, Z_n$ are real-valued and independent with densities f_1 and f_2, respectively. An estimator for the density p of a transformation $T_1(X) + T_2(Z)$ is the local von Mises statistic

$$\hat{p}(z) = \frac{1}{n^2} \sum_{i,j=1}^{n} k_b(z - T_1(X_i) - T_2(Z_j)).$$

Let g_1 and g_2 denote the densities of $T_1(X)$ and $T_2(Z)$. As in the one-sample case (1) we obtain a stochastic expansion

$$\hat{p}(z) = p(z) + \frac{1}{n} \sum_{i=1}^{n} \big(g_1(z - T_2(Z_i)) + g_2(z - T_1(X_i)) - 2p(z)\big) + o_P(n^{-1/2}).$$

Appropriate versions of Theorems 1–3 continue to hold. □

Remark 5. (Regression.) Two-sample results can be applied to regression models $Y = r(X) + \varepsilon$ with ε independent of X. If we have independent observations (X_i, Y_i), $i = 1, \ldots, n$, then the density p of Y can be estimated by the local von Mises statistic

$$\hat{p}(z) = \frac{1}{n^2} \sum_{i,j=1}^{n} k_b(z - \hat{r}(X_i) - \hat{\varepsilon}_j)$$

based on some estimator $\hat{r}$ of the regression function r, and on residuals $\hat{\varepsilon}_j = Y_j - \hat{r}(X_j)$. Note that the "pseudo-observations" $\hat{r}(X_i)$ and $\hat{\varepsilon}_j$ are only approximately independent, so we are close to the two-sample case with $Z = \varepsilon$, $T_1(X) = r(X)$, and $T_2(\varepsilon) = \varepsilon$. As seen above, we can expect a rate $n^{-1/2}$ for $\hat{p}(z)$ if r has a derivative that is bounded away from 0.

Suppose r is only piecewise monotone and continuously differentiable, and there are points x with

$$r(y) = c(y-x)^\nu + o(|y-x|^\nu)$$

for y to the left or right of x. Then the convergence rate of $\hat{p}(z)$ will be determined by the largest such ν. □

Remark 6. (Time series.) Results for regression carry over to time series driven by independent innovations. Consider a first-order moving average process $X_i = \varepsilon_i + \varphi\varepsilon_{i-1}$, with independent innovations ε_i that have mean 0, finite variance, and density f. If $\varphi \neq 0$, the stationary density p can be estimated by a local von Mises statistic

$$\hat{p}(z) = \frac{1}{n^2}\sum_{i,j=1}^{n} k_b(z - \hat{\varepsilon}_i - \hat{\varphi}\hat{\varepsilon}_j)$$

with $\hat{\varphi}$ an estimator of φ. Saavedra and Cao, 1999 [8] obtain $n^{1/2}$-consistency; see also Schick and Wefelmeyer, 2004 [10]. Functional results for higher-order moving average processes and general linear processes are obtained in Schick and Wefelmeyer, 2004 [12], 2007 [14], 2008 [15] and 2009 [18]. Nonlinear and nonparametric time series can also be treated. □

Remark 7. (Renewal processes.) Here is a two-sample case where $T(X,Z)$ is a product rather than a sum of functions $T_1(X)$ and $T_2(Z)$. Let (X_i, T_i), $i = 0, \ldots, n$ be observations of a Markov renewal process with real state space. Assume that the embedded Markov chain is stationary. We make the structural assumption that the waiting times depend multiplicatively on some power of the distance between the previous and the present state of the embedded Markov chain,

$$T_i - T_{i-1} = |X_i - X_{i-1}|^\nu W_i,$$

where $\nu > 0$ and the W_i are independent with density g and independent of the embedded Markov chain. Note that W_i is observable as a function of the observations (X_{i-1}, T_{i-1}) and (X_i, T_i). We can estimate the waiting time density p of $T_i - T_{i-1}$ by the local von Mises statistic

$$\hat{p}(z) = \frac{1}{n^2}\sum_{i,j=1}^{n} k_b(z - |X_i - X_{i-1}|^\nu W_j).$$

Greenwood et al., 2009 [4] give conditions under which $\hat{p}(z)$ has rate $n^{-1/2}$ and is asymptotically linear and asymptotically normal. □

References

1. Escanciano, J. C. and Jacho-Chávez, D. T. "$\sqrt{n}$-Uniformly Consistent Density Estimation in Nonparametric Regression Models". Submitted (2010).
2. Frees, E. W. "Estimating Densities of Functions of Observations". *J. Amer. Statist. Assoc.* 89:517–525 (1994).
3. Giné, E. and Mason, D. "On Local U-Statistic Processes and the Estimation of Densities of Functions of Several Sample Variables". *Ann. Statist.* 35:1105–1145 (2007).
4. Greenwood, P. E., Schick, A. and Wefelmeyer, W. "Estimating the Inter-Arrival Time Density of Semi-Markov Processes under Structural Assumptions on the Transition Distribution". Submitted (2009).
5. Müller, U. U. "Estimating the Density of a Possibly Missing Response Variable in Nonlinear Regression". Submitted (2010).
6. Nickl, R. "Donsker-Type Theorems for Nonparametric Maximum Likelihood Estimators". *Probab. Theory Related Fields* 138:411–449 (2007). Erratum 141:331–332 (2008).
7. Nickl, R. "On Convergence and Convolutions of Random Signed Measures". *J. Theoret. Probab.* 22:38–56 (2009).
8. Saavedra, A. and Cao, R. "Rate of Convergence of a Convolution-Type Estimator of the Marginal Density of an MA(1) process". *Stochastic Process. Appl.* 80:129–155 (1999).
9. Saavedra, A. and Cao, R. "On the Estimation of the Marginal Density of a Moving Average Process". *Canad. J. Statist.* 28:799–815 (2000).
10. Schick, A. and Wefelmeyer, W. "Root n Consistent and Optimal Density Estimators for Moving Average Processes". *Scand. J. Statist.* 31:63–78 (2004).
11. Schick, A. and Wefelmeyer, W. "Root n Consistent Density Estimators for Sums of Independent Random Variables". *J. Nonparametr. Statist.* 16:925–935 (2004).
12. Schick, A. and Wefelmeyer, W. "Functional Convergence and Optimality of Plug-in Estimators for Stationary Densities of Moving Average Processes". *Bernoulli*, **10**, 889–917 (2004).
13. Schick, A. and Wefelmeyer, W. "Root-n Consistent Density Estimators of Convolutions in Weighted L_1-Norms". *J. Statist. Plann. Inference* 37:1765–1774 (2007).
14. Schick, A. and Wefelmeyer, W. "Uniformly Root-n Consistent Density Estimators for Weakly Dependent Invertible Linear Processes". *Ann. Statist.* 35:815–843 (2007).
15. Schick, A. and Wefelmeyer, W. "Root-n Consistency in Weighted L_1-Spaces for Density Estimators of Invertible Linear Processes". *Stat. Inference Stoch. Process.*, 11:281–310 (2008).
16. Schick, A. and Wefelmeyer, W. "Convergence Rates of Density Estimators for Sums of Powers of Observations". *Metrika* 69:249–264 (2009).
17. Schick, A. and Wefelmeyer, W. "Non-Standard Behavior of Density Estimators for Sums of Squared Observations". *Statist. Decisions* 27:55-73 (2009).
18. Schick, A. and Wefelmeyer, W. "Improved Density Estimators for Invertible Linear Processes". *Comm. Statist. Theory Methods* 38:3123-3147 (2009).
19. Støve, B. and Tjøstheim, D. "A Convolution Estimator for the Density of Nonlinear Regression Observations". Submitted (2010).

Branching Walks in Inhomogeneous Random Environments

Elena Yarovaya[1]

Department of Probability Theory
Faculty of Mechanics and Mathematics
Moscow State University
Leninskie Gory
119992 Moscow, Russia
(e-mail: yarovaya@mech.math.msu.su)

Abstract. We consider a branching random walk on $\mathbf{Z}^d$ ($d \geq 1$) with one source of branching at the origin. The birth-and-death processes are defined by a random potential V. The evolution of the first moment of the number of particles at a point $x \in \mathbf{Z}^d$ or the total particle population is described by the Cauchy problem with a random potential V for the (t-differential x-pseudo-difference) equation $\partial_t m_1 = Am_1 + V\delta_0(x)m_1$ on $\mathbf{Z}^d$ ($d \geq 1$) with the initial conditions $m_1(0,\cdot) = \delta_0(\cdot)$ or $m_1(0,\cdot) \equiv 1$, where the (x-pseudo-difference) operator A is a generator of symmetric random walk on $\mathbf{Z}^d$. The long-time behavior of the moments $\langle m_n^p \rangle$ ($p \geq 1$) for the potentials V with "heavy" upper tails is obtained.
Keywords: Branching random walks, Inhomogeneous environment, Random environment, Kolmogorov backward equations, Feynman–Kac representation, Random moments.

1 Introduction

The present paper is devoted to the study of continuous-time symmetric branching random walks (BRW) under the assumption that the intensities of birth and death of particles at the source are random. As well known (see, Gärtner and Molchanov, 1990 [6], 1998 [7], Molchanov, 1994 [8], 1996 [9] and bibliography therein) the concept of "strong centers" is used for the interpretation of the intermittency phenomenon in the theory of random media. Much attention in the theory of random media, in particular in connection with the localization problem (see, e.g., Carmona and Lacroix, 1990 [3]), has been devoted to the study of spectral properties of the Anderson operator $\kappa\Delta + V(x)$ with $\kappa > 0$, where Δ is the discrete Laplacian on $\mathbf{Z}^d$ acting in variable x as $\Delta\psi(x) = \frac{1}{2d}\sum_{|x'-x|=1}\psi(x') - \psi(x)$, and the potential $V(x) = V(x,\omega)$, $x \in \mathbf{Z}^d$, $d \geq 1$, is a random function determined by the random branching medium. The evolution of the first order moments for the local and total numbers of particles in continuous-time BRW for a spatially homogeneous random branching environment is described by the operator $\kappa\Delta + V(x)$. For example, the expected total number of particles (i.e., the

moment of the first order) satisfies the Cauchy problem for the Anderson operator with a random potential:

$$\partial_t m_1(t,x) = \kappa \Delta m_1(t,x) + V(x) m_1(t,x), \quad m_1(0,x) \equiv 1. \tag{1}$$

It has been discovered that the evolution of the field $m_1(t,x)$ leads to the formation of highly irregular time-space structures, characterized by the generation of rare high peaks accumulating the bulk "mass" of the evolving field (see, Zeldovich et al 1988 [11], Gärtner and Molchanov, 1990 [6], 1998 [7], and Molchanov, 1994 [8], 1996 [9]). Such a phenomenon has received the name "intermittency". Studying of intermittency in the papers Gärtner and Molchanov, 1990 [6], 1998 [7] and Molchanov, 1994 [8], 1996 [9] is based on the asymptotic analysis of the moments $\langle m_1 \rangle$ obtained by averaging the random moment m_1 over medium's realizations. The angular brackets $\langle \cdot \rangle$ indicate the expectation with respect to a random environment. In particular, these works have shown that intermittency manifests itself as an anomalous, progressive growth, as $t \to \infty$, of the moments $\langle m_1^p \rangle$ with respect to their order p. For instance, the second moments grow much faster than the squared first moments: $\langle m_1^2 \rangle \gg \langle m_1 \rangle^2$, $\langle m_1^4 \rangle \gg \langle m_1^2 \rangle^2$, and so on.

Equations for the higher-order moments are similar to equation (1) (see, Yarovaya, 1990 [10] and Albeverio et al 2000 [1]), but they satisfy inhomogeneous Cauchy problems. In addition, the right-hand parts of these equations for the higher-order moments contain terms which depend on the branching environment and on the moments of lower orders. In the paper Albeverio et al 2000 [1] the case where the potential $V(x)$ has a Weibull type upper tail was studied and the Lyapunov exponents for the moments of all orders were calculated.

In the present paper, the evolution of the mean number of particles m_1 in an inhomogeneous random environment is determined by the operator $A + V\Delta\!\!\!/_0$, where the random walk generator A is a bounded self-adjoint operator in $l^2(\mathbf{Z}^d)$ defined by the following x-pseudo-difference formula $(Au)(x) = \sum_{x' \in \mathbf{Z}^d} a(x,x')u(x')$, and $\Delta\!\!\!/_0 = \delta_0 \delta_0^T$. Unlike (1), V is a random variable depending only on birth-and-death intensities at the source. Here as usual $\delta_0 = \delta_0(\cdot)$ denotes the column-vector on $\mathbf{Z}^d$ taking the value 1 at the origin and 0 at other points. To avoid confusion with the standard notation of the Laplace operator, the above operator is denoted as "slanted delta" $\Delta\!\!\!/$. The aim of the paper is to demonstrate that the condition

$$\lim_{t \to \infty} \frac{t}{\ln \langle e^{Vt} \rangle} = 0 \tag{2}$$

implies growth of the moments $\langle m_n^p \rangle$, $n = 1, 2, \ldots$, with respect to p, as $t \to \infty$, typical for the phenomenon of intermittency.

2 BRW in an Inhomogeneous Random Environment

In the model of a symmetric BRW the random walk of particles is given by the infinitesimal transition matrix $A = \|a(x,y)\|_{x,y\in\mathbf{Z}^d}$. It is assumed to be symmetric: $a(x,y) = a(y,x)$, homogeneous: $a(x,y) = a(0,y-x) = a(y-x)$, irreducible, regular: $\sum_{x\in\mathbf{Z}^d} a(x) = 0$ with $a(x) \geq 0$, $x \neq 0$, $a(0) < 0$, and having a finite variance of jumps: $\sum_{x\in\mathbf{Z}^d} x^2 a(x) < \infty$. In virtue of symmetry and homogeneity of the random walk, the conditions $\sum_{y\in\mathbf{Z}} a(x,y) = 0$ and $\sum_{x\in\mathbf{Z}} a(x,y) = 0$ are satisfied for the matrix A. In particular, this class includes the simple symmetric random walk defined by $a(x,y) = -a(0)/2d$ for $|y-x| = 1$, $a(x,x) = a(0)$ and $a(x,y) = 0$ otherwise. If $\kappa = a(0)/2d$, then we get the simple symmetric random walk defined by the operator $\kappa\Delta$ in the Cauchy problem (1).

Suppose now that a branching process at the origin is defined by intensities of death ξ^- and binary splitting ξ^+ of particles. In this case a branching random environment is formed by only one pair of non-negative random variables, $\xi := (\xi^-, \xi^+)$ defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$. It is assumed that $\Omega = \mathsf{R}^2_+$.

Therefore, if at time t there are $\mu_t(0) > 0$ particles at the origin, then each particle in the time interval $[t, t+h)$ independently of others can either jump with the probability $p(h,0,y) = a(y)h + o(h)$ to the point $y \neq 0$, or produce two particles including itself with the probability $\xi^+ h + o(h)$, or die with the probability $\xi^- h + o(h)$, or survive (no changes) with the probability $1 - \sum_{y\neq 0} a(y)h - (\xi^- + \xi^+)h + o(h)$. The standard method used in the case of the fixed medium realization ω can be applied to prove that the sojourn time of a particle at the point x is distributed exponentially with the parameter $-a(0) + \xi^+(\omega) + \xi^-(\omega)$. As above, we suggest that newly born particles evolve by the same rule, independently of the other particles and the past history. We also specify the initial conditions by assuming that at time $t = 0$ there is a single particle at a point $x \in \mathbf{Z}^d$. The evolution of the system of particles on $\mathbf{Z}^d$ is described by the number $\mu_t(y)$ of particles at time t at each point $y \in \mathbf{Z}^d$ and the total particle population size $\mu_t := \sum_{y\in\mathbf{Z}^d} \mu_t(y)$.

Hence the generating functions associated with the random variables $\mu_t(y)$ and μ_t are defined for $z > 0$ by

$$F(z;t,x,y) := \mathbf{E}_x^{(\omega)}[e^{-z\mu_t(y)}], \quad F(z;t,x) := \mathbf{E}_x^{(\omega)}[e^{-z\mu_t}],$$

where $\mathbf{E}_x^{(\omega)}$ is the corresponding expectation. The label ω is referring to the fixed realization of the branching medium ξ, and the subscript x indicates the initial position of the single original particle. Similar to Albaverio et al 2000 [1], it can be obtained that the functions $F(z;t,x,y)$ and $F(z;t,x)$ satisfy the (t-differential x-pseudo-difference) equation

$$\partial_t F = AF - [\xi^+ + \xi^-]\delta_0(x)F + \xi^+\delta_0(x)F^2 + \xi^-\delta_0(x) \tag{3}$$

with the initial conditions $F(z;0,x,y) = e^{-z\delta_y(x)}$ and $F(z;0,x) = e^{-z}$. Let $V := \xi^+ - \xi^-$. In this case the following equations for the moment func-

tions $m_n(t,x,y)$, $m_n(t,x)$ are obtained formally from the equation for the generating functions (3) and satisfy the chain of linear differential equations

$$\partial_t m_1 = Am_1 + V\delta_0(x)m_1, \tag{4}$$

$$\partial_t m_n = Am_n + V\delta_0(x)m_n + \xi^+\delta_0(x)g_n[m_1,\dots,m_{n-1}], \quad n = 1,2,\dots, \tag{5}$$

with the initial conditions $m_n(0,\cdot,y) = \delta_y(\cdot)$ and $m_n(0,\cdot) \equiv 1$, where $g_1 \equiv 0$ and $g_n[m_1,\dots,m_{n-1}](\cdot)$, for $n \geq 2$, is the following function of the variable x:

$$g_n[m_1,\dots,m_{n-1}](x) := \sum_{i=1}^{n-1} \binom{n}{i} m_i(x)m_{n-i}(x).$$

3 Main Results

As well known, the Feynman-Kac representation is a generalization of Kolmogorov's backward equation. Let us point out that the main technical tool used in the papers of Gärtner et al 2007 [5], Albeverio et al 2000 [1], Gärtner and Molchanov, 1990 [6], 1998 [7], Yarovaya, 1990 [10] is the Feynman-Kac repersentation of the solution for the problem (1) with different initial conditions such as $m_1(0,\cdot) \equiv 1$ or $m_1(0,\cdot,y) = \delta_y(\cdot)$.

Below, we give the Feynman-Kac representations for the solutions of the Cauchy problems (4). In what follows let x_t be a continuous-time "jumping" trajectory of an auxiliary continuous-time symmetric random walk on $\mathbf{Z}^d$ with the generator A, and E_x be the expectation under the condition that the random walk has started at the point x.

Theorem 1 (Kolmogorov's backward equation). *Define* $p(t,x,y) = E_x\delta_y(x_t)$. *Then* $p(t,\cdot,y) \in l^2(\mathbf{Z}^d)$ *for each* $t > 0$ *and*

$$\partial p_t = Ap, \quad p(0,x,y) = \delta_y(x), \tag{6}$$

where the right hand side is interpreted as the linear operator A applied to the function $x \mapsto p(t,x,y)$ *by the formula:* $Ap(t,x,y) = \sum_{x'} a(x,x')p(t,x',y)$.

Moreover, if $p^*(t,x,y)$ *satisfies the Cauchy problem* (6), *then* $p^*(t,x,y) = p(t,x,y)$ *with* $p(t,x,y) = E_x\delta_y(x_t)$.

Under assumption that the Cauchy problem (4) **P**-a.s. has a unique non-negative solution the following generalizations of Theorem 1 also hold.

Theorem 2. *Let* V *be a positive random variable, and let the Cauchy problem* (4) **P***-a.s. have a unique non-negative solution. Put*

$$m_1(t,x,y) = E_x\left[\exp\left(V\int_0^t \delta_0(x_s)\,ds\right)\delta_y(x_t)\right],$$

$$m_1(t,x) = E_x\left[\exp\left(V\int_0^t \delta_0(x_s)\,ds\right)\right].$$

Then $m_1(t,x,y)$ $\mathbf{P}$-a.s. satisfies the Cauchy problem (4) *with the initial condition $m_1(0,\cdot,y)=\delta_y(\cdot)$ while $m_1(t,x)$ $\mathbf{P}$-a.s. satisfies the Cauchy problem* (4) *with the initial condition $m_1(0,\cdot)\equiv 1$.*

Theorem 3. *Let V be a positive random variable, let the Cauchy problem* (4) $\mathbf{P}$*-a.s. have a unique non-negative solution, and let*

$$\lim_{t\to\infty}\frac{t}{\ln\langle e^{Vt}\rangle}=0. \tag{7}$$

Then for all integer moments $\langle m_n^p\rangle$, $n\geq 1$, where m_n are the solutions of Cauchy problems (4), (5) *with the initial conditions $m_n(0,\cdot,y)=\delta_y(\cdot)$ and $m_n(0,\cdot)\equiv 1$, respectively, we get*

$$\lim_{t\to\infty}\frac{\ln\langle m_n^p\rangle}{\ln\langle e^{pnVt}\rangle}=1. \tag{8}$$

The proof is based on the following upper and lower estimates for the moments $\langle m_n^p\rangle$. These estimates are obtained by using the Feymann-Kac representations for the solutions of the studied Cauchy problems. The estimates for $n\geq 2$ are cumbersome and the volume of the paper does not allow for their presentation, so we give below the estimates only for $n=1$.

Upper estimate for $\langle m_1^p\rangle$. Applying Lyapunov's and Jensen's inequalities to the Feymann-Kac formula (Theorem 2) and using Fubini's theorem, we obtain

$$\begin{aligned}\langle m_1^p(t,x,y)\rangle &= \left\langle\left(E_x e^{V\int_0^t\delta_0(x_s)\,ds}\delta_y(x_t)\right)^p\right\rangle \leq \left\langle E_x e^{pV\int_0^t\delta_0(x_s)\,ds}\delta_y^p(x_t)\right\rangle\\ &= E_x\left\langle e^{pV\int_0^t\delta_0(x_s)\,ds}\right\rangle\delta_y(0)\leq\frac{1}{t}\int_0^t E_x\left\langle e^{ptV}\right\rangle ds\,\delta_y(0)=\left\langle e^{ptV}\right\rangle\delta_y(0),\end{aligned}$$

and similarly

$$\begin{aligned}&\langle m_1^p(t,x)\rangle = \left\langle\left(E_x e^{V\int_0^t\delta_0(x_s)\,ds}\right)^p\right\rangle \leq \left\langle E_x e^{pV\int_0^t\delta_0(x_s)\,ds}\right\rangle\\ &= E_x\left\langle e^{\frac{1}{t}\int_0^t ptV\delta_0(x_s)\,ds}\right\rangle \leq E_x\left\langle\frac{1}{t}\int_0^t e^{ptV\delta_0(x_s))}\,ds\right\rangle=\frac{1}{t}\int_0^t E_x\left\langle e^{ptV}\right\rangle ds.\end{aligned}$$

Here the expectation in the last integrand does not depend on s and is equal to $\langle e^{ptV}\rangle$. Hence we get

$$\langle m_1^p(t,x)\rangle\leq\left\langle e^{ptV}\right\rangle.$$

Lower estimate for $\langle m_1^p\rangle$. If $y=x=0$, then the required estimate can be obtained by taking into account in the Feynman–Kac formula from Theorem 2 only those paths of the random walk x_s which stay at the initial point $x=0$ during the time interval $[0,t]$. Indeed, let τ denote the time spent by the

random walk at the initial state until exit. Since τ is exponentially distributed with the parameter $-a(0)$, we have

$$m_1(t,0,0) \geq E_0\left[I\{\tau > t\}\, e^{V\int_0^t \delta_0(x_s)\,ds}\delta_0(x_t)\right] = \mathrm{P}\{\tau > t\}\cdot e^{Vt} = e^{a(0)t}\cdot e^{Vt},$$
$$m_1(t,0) = E_0\left[e^{\int_0^t V(x_s)\,ds}\right] \geq e^{a(0)t}\cdot e^{Vt}$$

and

$$\langle m_1^p(t,0)\rangle \geq \langle m_1^p(t,0,0)\rangle \geq e^{pa(0)t}\left\langle e^{ptV}\right\rangle.$$

It is not difficult to show that

$$\langle m_1^p(t,x)\rangle \geq \langle m_1^p(t,x,y)\rangle \geq f(t)e^{pa(0)t}\left\langle e^{ptV}\right\rangle,$$

where $\ln f(t) \approx t$, as $t \to \infty$. Therefore using the obtained upper and lower estimates we get that condition (7) implies (8).

4 The Potential with Gumbel Type Upper Tail

Here we construct an example in which the distribution of the random potential V satisfy (7). In the next theorem the upper tail of the distribution of the potential V has the following form:

$$\ln \mathrm{P}\{V > \theta\} \sim -\exp\left(\frac{\theta}{c}\right), \quad \theta \to \infty, \quad c > 0. \tag{9}$$

Theorem 4. *Under the assumption* (9) *for every* $p \geq 1$ *the following relation holds*

$$\ln\left\langle e^{pVt}\right\rangle \sim cpt\ln t, \quad t \to \infty,$$

and conditions (7) *are valid.*

Proof. Let $P(\theta) = \mathrm{P}\{V > \theta\}$. Then by the definition of $\langle\cdot\rangle$ we have

$$\left\langle e^{pVt}\right\rangle = \int_{-\infty}^{\infty} e^{ptu}\, d(1-P(u)).$$

Hence

$$\int_0^{\infty} e^{ptu}\, d(1-P(u)) \leq \left\langle e^{pVt}\right\rangle \leq 1 + \int_0^{\infty} e^{ptu}\, d(1-P(u)).$$

Using the representation

$$\int_0^{\infty} e^{ptu}\, d(1-P(u)) = P(0) + pt\int_0^{\infty} e^{ptu}P(u)\, du,$$

the proof may be reduced to finding the logarithmic asymptotic behavior, as $t \to \infty$, of the following integral

$$w(t) := pt \int_0^\infty e^{ptu} P(u)\, du.$$

Put $Q(u) := \ln P(u)$. Condition (9) implies the existence of a function $z(u)$ such that $z(u) \to 0$, as $u \to \infty$, and $Q(u) = -\exp(u/c)(1 + z(u))$. So, for any $0 < \varepsilon < 1$ we can select $t(\varepsilon) \geq 0$ such that $Q_\varepsilon^- \leq Q(u) \leq Q_\varepsilon^+$ for $u \geq t(\varepsilon)$ where $Q_\varepsilon^- := -(1+\varepsilon)\exp(u/c)$ and $Q_\varepsilon^+ := -(1-\varepsilon)\exp(u/c)$. Hence

$$w_\varepsilon^-(t) + W_\varepsilon^-(t) \leq w(t) \leq w_\varepsilon^+(t) + W_\varepsilon^+(t), \tag{10}$$

where

$$w_\varepsilon^-(t) := pt \int_0^{t(\varepsilon)} e^{ptu} P(u)\, du - pt \int_0^{t(\varepsilon)} e^{ptu + Q_\varepsilon^-(u)}\, du,$$

$$w_\varepsilon^+(t) := pt \int_0^{t(\varepsilon)} e^{ptu} P(u)\, du - pt \int_0^{t(\varepsilon)} e^{ptu + Q_\varepsilon^+(u)}\, du,$$

$$W_\varepsilon^-(t) := pt \int_0^{\infty} e^{ptu + Q_\varepsilon^-(u)}\, du,$$

$$W_\varepsilon^+(t) := pt \int_0^{\infty} e^{ptu + Q_\varepsilon^+(u)}\, du.$$

Then clearly

$$|w_\varepsilon^-(t)|, |w_\varepsilon^+(t)| \leq C^* e^{ptt(\varepsilon)}, \tag{11}$$

where C^* is a constant. Each of the functions $W_\varepsilon^-(t)$, $W_\varepsilon^+(t)$ may be represented as an integral, the asymptotic behavior of which can be found by the saddle point method (see, e.g., de Bruijn, 1958 [2] and Fedoryuk, 1987 [4]). Thus we get

$$W_\varepsilon^-(t) \sim S_\varepsilon^-(t) \exp\left(pct \ln pct - (1+\varepsilon)pct\right),$$

$$W_\varepsilon^+(t) \sim S_\varepsilon^+(t) \exp\left(pct \ln pct - (1-\varepsilon)pct\right),$$

where the functions $S_\varepsilon^-(t)$ and $S_\varepsilon^+(t)$ grow, as $t \to \infty$, no faster than the power functions. Hence by (10), (11) and by the form of the asymptotic behavior of $W_\varepsilon^-(t)$ and $W_\varepsilon^+(t)$ we obtain the following asymptotic upper and lower estimations

$$pct(\ln pct - (1+\varepsilon)) \lesssim \ln w(t) \lesssim pct(\ln pct - (1-\varepsilon)),$$

from which due to arbitrariness of ε, we get $\ln w(t) \sim pct \ln t$, as $t \to \infty$. Now it is easy to verify that (9) implies (7).

5 Conclusion

It is not difficult to show that if a distribution of the potential V satisfies $\ln \mathrm{P}\{V > \theta\} \sim -c\theta$, as $\theta \to \infty$, with some $c > 0$, then $\ln \left\langle e^{pVt} \right\rangle \sim cpt$, as

$t \to \infty$, and (7) is not valid. The validity of (7) means that the tail of a distribution of the potential V is "heavier" than the exponential one.

Moreover, note that (7) implies validity of relation (8) for the models of BRW in spatially homogeneous random environments too.

Remark at last that principal assumption in Theorems 2 and 3 is the **P**-a.s. uniquiness of a non-negative solution for the Cauchy problem (4). We conjecture that, for a positive random variable V, such a uniquiness is guaranteed by the condition

$$\langle (\xi^+ (\ln_+ V)^{-1})^d \rangle < \infty,$$

where $\ln_+ V := \ln \max\{V, e\}$.

Acknowledgment

This work is supported by the RFBR grant 10-01-00266.

References

1. Albeverio, S., Bogachev, L., Molchanov, S., and Yarovaya, E., "Annealed Moment Lyapunov Exponents for a Branching Random Walk in a Homogeneous Random Branching Environment" ", *Markov Processes and Related Fields* 6, 473–516 (2000).
2. Bruijn, N. de., *Asymptotic methods in analysis*, North-Holland Publishing Co.-Amsterdam P. Noordhoff Ltd., Groningen (1958).
3. Carmona, R.,and Lacroix J., "Spectral Theory of Random Schroedinger Operators", *Probability and Its Applications* Birkhäuser Verlag, Basel (1990).
4. Fedoryuk, M.V., *Asymptotics: Integrals and Series*, Nauka, Moscow (1987).
5. Gärtner, J., König, W., and Molchanov S., "Geometric characterization of intermittency in the parabolic Anderson model", *The Annals of Probability*, 35, 339–499 (2007).
6. Gärtner, J., and Molchanov S., "Parabolic problems for the Anderson model. I: Intermittency and related topics", *Commun. Math. Phys.*, 132, 613–655 (1990).
7. Gärtner, J., and Molchanov S., "Parabolic problems for the Anderson model. II: Second-order asymptotics and structure of high picks", *Probab. Theory Relat. Fields*, 111, 17–55 (1998).
8. Molchanov, S., "Lectures on random media", *Lect. Notes Math.* 1581, 242–411 (1994).
9. Molchanov, S., "Reaction-diffusion equations in the random media: Localization and intermittency", in *Nonlinear Stochastic PDEs. Hydrodynamic Limit and Burgers' Turbulence, Springer-Verlag, Berlin–Heidelberg–New York. IMA, Math. Appl.* 77, 81–109 (1996).
10. Yarovaya, E., "Intermittency of leading moments in the model of branching process with diffusion in a random medium", *Mosc. Univ. Math. Bull.*, 45, 49–51 (1990).
11. Zeldovich, Ya., Molchanov, S., Ruzmaikin, A., Sokoloff, D., "Intermittency, diffusion and generation in a nonstationary random medium" *Sov. Sci. Rev., Sect. C: Math. Phys. Rev.* 7, 1–110.