Modelling Point’s Position Time Series in the Light of Cointegration

Tomáš BACIGÁL and Magda KOMORNÍKOVÁ, Slovakia

Key words: GPS, time series, cointegration, common trend, prediction.

SUMMARY

Advancements and automation on the field of measuring instruments provide us with quantum of more and more precise data to be further processed in an appropriate way. In our paper we deal with such a set of data arranged in time – time series – taken from continuous GPS observations on permanent station Borowiec which takes part in EUREF network for monitoring Earth's crust kinematics. The outcome is a set of point's coordinates in local horizontal system (*n*,*e*,*v*) of which we took those two making horizontal plane, i.e. (*n*,*v*), and compare three methods of statistical processing. First, we model two univariate time series as if they are independent, another way is to accept the interrelationship and model it as one bivariate time series. The third one incorporates geometrical nature of both variables and makes use of a common trend presence. As a criterion of model's suitability we used mean square error and mean percentage error of predicted values.

Whole one subsection dwells on testing for the presence of stochastic trend and subsequently for a cointegration, which is essential when investigating the series for common stochastic trend. This is represented by augmented Dickey-Fuller test and Johansen’s test, respectively. Having affirmed cointegration, we transform the (*n*,*e*) system to obtain a new one, (*y*,*x*), oriented according to the common deterministic (linear) trend. This is processed on as usually and transformed back, finally. The usual procedure consists of trend and seasonality decomposition and applying the autoregressive models to still correlated residuals.

Mean square and mean percentage errors computed for 5 predicted values per variable speak very clearly for the model supporting cointegration.

Modelling Point's Position Time Series in the Light of Cointegration

Tomáš BACIGÁL and Magda KOMORNÍKOVÁ, Slovakia

1. INTRODUCTION

Many technical disciplines involved in civil engineering, such as geology, geodesy, statics of structures and others deal with position of particular points in time-space to figure out processes that influence our environment (both original and man-made). Supported by advancements and automation on the field of measuring instruments, monitoring becomes robust and effective, yet demanding more appropriate methods of processing. In this paper we'll focus on modelling time-series arisen from observations by NAVSTAR Global positioning system (GPS), which is satellite based navigational system developed and provided by the American Department of Defence. Observations had been performed daily in years 2001-2002 on GPS permanent station Borowiec (BOR1, Poland) which takes part in EUREF Permanent Network representing a regional densification of global IGS net in Europe which is used, among other purposes, for regular monitoring of recent kinematics of the Earth's crust (see Hefty and Husár 2003). The standard outcome, being in the form of three coordinates (X,Y,Z) in geocentric coord.system, was transformed into local topocentric horizontal coordinate system (n,e,v - north, east, vertical component) - with the origin in the mean position of the two year period - to be further processed.

Because of significantly lower precision and negligible linear trend in vertical direction, we only deal here with the two time series n and e each containing 730 data points. Figure 1 shows two dimensional representation of point variation on Earth's surface and Figure 2 time plot for each coordinate.

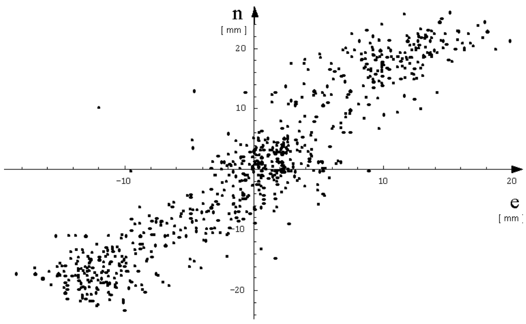

Figure 1: Daily record of point's position in a ground plane.

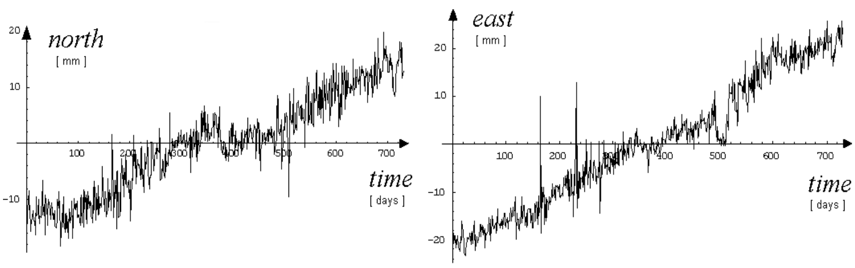

Figure 2: Time plot of point's position variation.

There's easily seen the data following linear trend with a high level of fit. It's a consequence of the long-term drift of the Eurasian tectonic plate, anyway, this overt drift is pretty suitable for applying several approaches of data processing mostly used in mathematical statistics and for showing a plus of the proposed key procedure.

2. DATA PROCESSING

Basically, we may treat our data as (a) two independent time series or (b) use the fact that both series just reject the same systematic and random disturbing effects, in other words, they are significantly interconnected.

The first approach has been and still is the most preferred way of time series modelling in general, which provides solid results in fitting. However, there is a slightly higher danger here of modelling spurious processes and, consequently, coming to misleading interpretations. The standard procedure includes modelling polynomial (linear) and periodic (seasonal) trends, and then applying Box-Jenkins methodology to cover some residual autocorrelations. This is well described in Reference and user's guide to Time series pack for Mathematica (1995) and we'll reenter it later in more details.

As for the (b)-group, it's a reasonable tendency of evolution in data processing to look for further relations and to develop more effective techniques (gratefully using computers), such as turning from single equations to vector representation of mathematical relations, etc. Vector regression analysis gives additional information about modelled processes and the way they are linked together (in the form of cross-correlation matrices, basically). We chose this modern approach as the second alternative to be compared in conclusion.

Still staying in the last group, we should introduce a theory largely elaborated by econometricians and given a name "cointegration". For brief explanation, two non-stationary I(1) time series (means integrated of the order 1, having the first differences stationary) are cointegrated, if one of their linear combinations is I(0) and hence stationary. There are several tests for cointegration, for details and references see Bognár (2005) and Franses(1998). The most used ones was employed for proving our series to be cointegrated, the procedures are briefly described in section 3.1.

3. COINTEGRATION

In this section, we perform some tests at first to find out what kind of trend is present in the data and to prove cointegration relation between our two time series. This is essential for applying common trend methodology in the later subsection.

If speaking about trend, it must be understood there are deterministic and stochastic trend being dealt with in time series theory and are often defined in the context of autoregressive models. Time series generated by deterministic trend (DT) model display mean or trend–reverting behaviour, while those generated by stochastic trend (ST) model lack the reverting forces. An illustrative example of DT model can be $X_t = \delta t + \varepsilon_t$ and ST model

$X_t = \delta + X_{t-1} + \varepsilon_t = X_0 + \delta t + \sum_{i=1}^{t} \varepsilon_i$, where ε_t is N(0, σ^2) random process and the model of ST is called random walk (i.e., AR parameter by X_{t-1} equals unity) with drift δ. The term $\sum_{i=1}^{t} \varepsilon_i$ is now called the stochastic trend, but we can see it may be accompanied by (any) deterministic trend component. The key difference is that ST series can deviate from this trend for lengthy period of time.

In our geometrical application, we have good reason to believe the trend in *n*, *e* has deterministic nature, however many times it is not the case.

3.1 Testing

For *identifying the stochastic trend*, there are two groups of methods

- tests for unit roots
- stationarity tests

An example of the second set of methods may be KPSS test or LM test. However, here we focus on widely used *augmented Dickey-Fuller* (ADF) tests from the first group (for details see Franses (1998, p.80). Procedure starts with choosing the order *p* of AR(*p*) model standardly from AIC, BIC information criteria, and continues by performing auxiliary regression

$$\Delta X_t = \mu + \delta t + \rho X_{t-1} + \phi_1 \Delta X_{t-1} + \ldots + \phi_{p-1} \Delta X_{t-p+1} + \varepsilon_t \qquad (1)$$

where X_t represents a variable, *t* is time, μ, δ, ρ, ϕ are parameters and ε residuals. ρ is the parameter of interest, for which we need to compute test statistic $t(\hat{\rho}) = \hat{\rho}/\mathrm{SE}(\hat{\rho})$, SE denotes standard error. The relevant null hypothesis is that $\rho = 0$ against alternative $\rho < 0$ (one-sided test), that means if $t(\hat{\rho}) > t_{critical}$ then we do not reject H_0 of unit root and hence, series contain stochastic trend. Otherwise there is no ST and we may solely think of eventual deterministic trend. Note, that test statistic does not follow standard asymptotic distribution, some critical values are provided, for example, in Franses (1998, p.82). The procedure was executed three times, firstly omitting both deterministic components (constant μ and trend δt), then including only constant and finally both of them. Table 1 shows the results, that speak clearly for the primacy of deterministic trend in both time series.

Table 1: Augmented Dickey-Fuller test

time series	deterministic component	$t(\rho)$	t_{crit} (α=0.05)	consequence
n	none	-1.70	-1.95	$\rho = 0$ indicates ST
	constant	-1.69	-2.86	$\rho = 0$ indicates ST
	constant & trend	-7.59	-3.41	$\rho < 0$, DT accepted
e	none	-0.95	-1.95	$\rho = 0$ indicates ST
	constant	-0.96	-2.86	$\rho = 0$ indicates ST
	constant & trend	-7.19	-3.41	$\rho < 0$, DT accepted

Having found trending behaviour of both our time series (within the framework of AR(*p*) model), it is natural to investigate whether these I(1) processes are "commonly integrated", i.e., there exists a common trend pattern.

There has been devised several methods of testing for cointegration. The first, *Engle-Granger two steps method* comes from single-equation model of two variables $X_{1,t}$,$X_{1,t}$ and works as follows. Residuals u_t from static regression

$$X_{2,t} = \beta_0 + \beta_1 X_{1,t} + u_t \quad (2)$$

are used in auxiliary regression

$$\Delta\hat{u}_t = \gamma_0 + \rho\hat{u}_{t-1} + \gamma_1\Delta\hat{u}_{t-1} + \ldots + \gamma_p\Delta\hat{u}_{t-p} + \varepsilon_t , \quad (3)$$

and the t-test for the significance of ρ is evaluated. When $\rho = 0$ (that is H_0, $t(\hat{\rho}) > t_{crit}$), u_t has a unit root and thus (2) does not reflect a stationary cointegration relationship. Otherwise, when $\rho < 0$, that is, when $t(\hat{\rho})$ is significantly negative, $X_{1,t}$ and $X_{2,t}$ are cointegrated. Some critical values are given in Franses (1998, p.217), test results in Table 2 shows indisputable presence of cointegration. By the way, R^2 (index of determination) by the regression of n_t on e_t is slightly higher, therefore this regression is to be more preferred here.

Engle-Granger is useful when we analyze two time series, but it may become less useful for increasing number of time series. This occurs, e.g. if we decide to include the third coordinate observations. Hence, multivariate methods appear to be more helpful.

Table 2: Engle-Granger testing for cointegration

regression of	deterministic component	$t(\hat{\rho})$	t_{crit} (α=0.05)	conclusion
e_t on n_t	constant constant & trend	-8.69 -9.56	-3.37 -3.80	cointegration
n_t on e_t	constant constant & trend	-9.71 -9.90	-3.37 -3.80	cointegration

To better understand cointegration and all associate terms, let's describe two time series by following VAR(1) model

$$\begin{bmatrix} 1 & \delta \\ 1 & \eta \end{bmatrix}\begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} = \begin{bmatrix} \mu_1^* \\ \mu_2^* \end{bmatrix} + \begin{bmatrix} \rho_1 & \delta\rho_1 \\ \rho_2 & \mu\rho_2 \end{bmatrix}\begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t}^* \\ \varepsilon_{2,t}^* \end{bmatrix}, \quad (4)$$

where $\delta \neq \eta$, μ_1^*, μ_2^* are intercept terms and $\varepsilon_{1,t}^*$, $\varepsilon_{2,t}^*$ are assumed to be mutually independent white noise error processes. Multiplying both sides with the inverse of the left-hand side matrix and subtracting the one period lagged X_{t-1} from both sides gives

$$\Delta X_t = \mu + \Pi X_{t-1} + \varepsilon_t , \quad (5)$$

where μ and ε_t are functions of μ^* and ε^* , respectively. Interesting matrix is Π, because when $0 \leq \rho_i < 1$ for $i = 1,2$, Π has full rank 2, on the other hand, when $\rho_1 = \rho_2 = 1$, the rank of Π is equal 0. Now interesting is the cointegration case, when (for example) $\rho_1 = 1$ and $0 \leq \rho_2 < 1$, the matrix Π can be written as

$$\Pi = \alpha\beta^{\mathrm{T}} \quad (6)$$

where $\beta = [1 \ \ \eta]^{\mathrm{T}}$ is the cointegration parameters vector. βX_t is an equilibrium (or long-run) relation between $X_{1,t}$, $X_{2,t}$, and the parameter matrix α reflects the speed of adjustment toward equilibrium. Equation (5) incorporating (6) is called a vector error correction model.

Multivariate method of testing for cointegration, proposed above, comes by considering again the VAR(p) model, more convenient if written in error correction format

$$\Delta X_t = \mu + \Gamma_1 \Delta X_{t-1} + \ldots + \Gamma_{p-1} X_{t-p+1} + \Pi X_{t-p} + \varepsilon_t , \qquad (7)$$

where Π contains the information on possible cointegrating relations between the m (in our case m = 2) elements of X_t. If Π is close to rank deficiency, there may be cointegration. The *Johansen's method* is such a statistical method to investigate the rank of Π (which, essentially, amounts to a multivariate extension of the univariate ADF method). The procedure goes like this. First, we perform regressions

$$\begin{aligned} \Delta X_t &= a_0 + a_1 \Delta X_{t-1} + \ldots + a_{p-1} \Delta X_{t-p+1} + r_{0t} \\ X_{t-p} &= b_0 + b_1 \Delta X_{t-1} + \cdots + b_{p-1} \Delta X_{t-p+1} + r_{1t} \end{aligned} \qquad (8)$$

then construct the matrices S_{00}, S_{10}, S_{11}, S_{01} of ($m \times m$) from

$$S_{ij} = \frac{1}{n} \sum_{t=1}^{n} r_{it} r_{jt}^{\mathrm{T}} , \qquad \text{for } i, j = 0{,}1 . \qquad (9)$$

The next step is to solve the eigenvalue problem $\left| \lambda S_{11} - S_{10} S_{00}^{-1} S_{01} \right| = 0$ which gives the eigenvalues $\hat{\lambda}_1 \geq \ldots \geq \hat{\lambda}_m$ and the corresponding eigenvectors $\hat{\beta}_1$ through $\hat{\beta}_m$. Now, a test for the rank of Π can be performed by testing how many eigenvalues λ_t equal zero. The first test statistic (which is a likelihood ratio test) $\lambda_{trace} = -n \sum_{i=r+1}^{m} \ln(1 - \hat{\lambda}_i)$ tests the null hypothesis of at most r cointegration relations against the alternative there are more of them, while the second test statistic $\lambda_{max} = -n \ln(1 - \hat{\lambda}_r)$ can be used to test the null of r – 1 against r cointegration relations (vectors). H_0 shall not be rejected if test statistic is smaller than critical value. Table of critical values can be found, e.g. in Franses (1998, p.224), our case evaluation is summarized in Table 3.

Table 3: Johansen's tests for cointegration (m = 2)

r	test statistic		$\lambda_{crit}(m\text{-}r)$ (α = 0.05)	conclusion
0	λ_{trace}	127.36	17.95	H_0 rejected
1	λ_{trace}	1.08	8.18	H_1 rejected, r =1 cointegr.vector
1	λ_{max}	126.28	14.90	H_0 rejected
2	λ_{max}	1.08	8.18	H_1 rejected, r =1 cointegr.vector

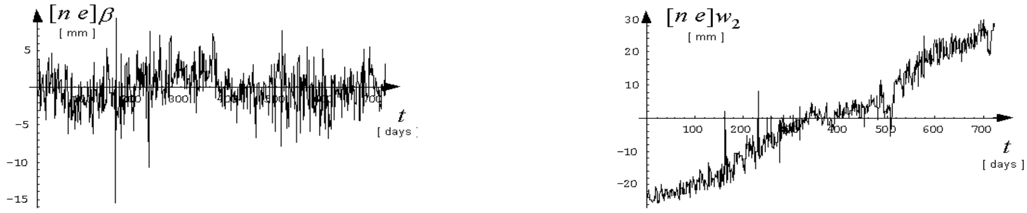

Figure 3: a) cointegration relation between n, e and b) common stochastic trend.

Having found 1 cointegration relation $[n_t \; e_t] \hat{\beta}_1$, it's not a bad idea to plot it (see Figure 3a) for later comparison. Anyway, when there are r cointegration relations among m variables, there has to be (m - r) independent common stochastic trends in the system. Gonzalo and Granger proposed a method to estimate the stochastic trends, procedure that use the Johansen's but differs in eigenvalue problem $\left|\lambda S_{00} - S_{01}S_{11}^{-1}S_{10}\right| = 0$, solution of which has the same eigenvalues $\hat{\lambda}_i$ but different eigenvectors $\hat{w}_1, \ldots, \hat{w}_m$. Because in our case r = 1, only one common stochastic trend variable can be constructed (using eigenvector $\hat{w}_{r+1}$), that is [n_t e_t] $\hat{w}_2$, plotted in Figure 3b.

3.2 Common Trend

Once having found cointegration, it's naturally leading us to investigate a common trend (see Komorníková a Komorník 2002). We look for linear combination

$$\begin{aligned} y &= \gamma_1 n + \delta_1 e \\ x &= \gamma_2 n + \delta_2 e \end{aligned} \qquad (10)$$

such that y represents a common trend direction and x is a stationary trend-free variable, orthogonal to y. In the light of our geometrical application, it's easy to rewrite a general common trend problem into familiar transformation (in 2D cartesian system)

$$\begin{aligned} y &= n\cos\alpha + e\sin\alpha \\ x &= -n\sin\alpha + e\cos\alpha \end{aligned} \qquad (11)$$

as shown in Figure 4.

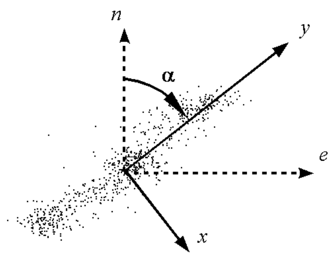

Figure 4: Transformation into common trend direction

The angle α can be determined either from analysis of stochastic trend

$$e_t = a_0 + b_0 n, \qquad \tan\alpha = b_0 \qquad (12)$$

or analysis of deterministic trend starting at linear regression

$$n_t = a_1 + b_1 t + \varepsilon_{1,t}, \qquad e_t = a_2 + b_2 t + \varepsilon_{2,t} \qquad (13)$$

where t denotes time and a, b regression parameters. If we place (13) into (11) and focus on series x, which is supposed to be trend-free, then

$$\begin{aligned} x_t &= -(a_1 + b_1 t + \varepsilon_{1,t})\sin\alpha + (a_2 + b_2 t + \varepsilon_{2,t})\cos\alpha, \\ x_t &= (a_2\cos\alpha - a_1\sin\alpha) + \underbrace{(b_2\cos\alpha - b_1\sin\alpha)}_{0}t + (\varepsilon_{2,t}\cos\alpha - \varepsilon_{1,t}\sin\alpha) \end{aligned} \qquad (14)$$

(linear trend term in x is eliminated), so

$$\tan\alpha = b_2 / b_1. \qquad (15)$$

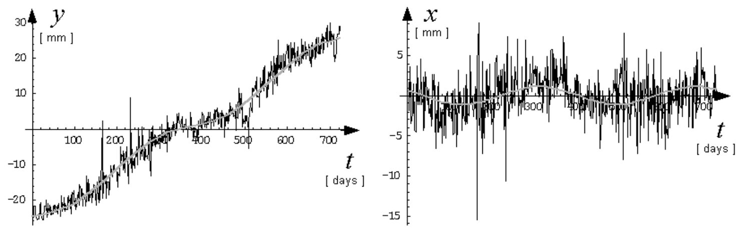

Figure 5: New series *y* and *x* fitted by linear and/or cyclical trend.

All right, we have got a new couple of time series *y*, *x*. At this point it is more than interesting to realize that Figures 3 and 5 show the same variables.

The next step is to model it the same way as two univariate series ((a) approach), at first by subtracting linear trend and seasonal component, then by testing it for residual auto-correlations and applying Box-Jenkins methodology. For comparing purposes we decided to include only annual seasonality and exclude any cyclical component. Figure 5 shows both series fitted by corresponding deterministic model. Correlogram of residuals confirmed the presence of significant correlations. This small residual dependencies may further be modelled by ARMA, ARCH, GARCH or some kind of TAR models, however here we simply employ the more standard autoregressive model of order *p*, which is chosen either from plot of residuals' variances (watching the relative steepness) or information criteria (finding a minimum), where Akaike's AIC and Schwarz's BIC are most used.

So the model of *y, x* is ready, schematically y_m, x_m = *trend* + *seasonality* + *AR*(*p*), however, this is not a final point we are supposed to come to. The new, model series must be transformed back to (*n*,*e*) system. If (11) is written in matrix notation, transformation matrix $M_{n,e\to y,x}$ is clearly orthogonal and therefore a backward transformation can easily be performed

$$\begin{bmatrix} n_m \\ e_m \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} y_m \\ x_m \end{bmatrix} \tag{16}$$

(because $M_{y,x\to n,e} = M^{-1}_{n,e\to y,x} = M^{\mathrm{T}}_{n,e\to y,x}$). For visual review Figure 7 joins original data with the *model*.

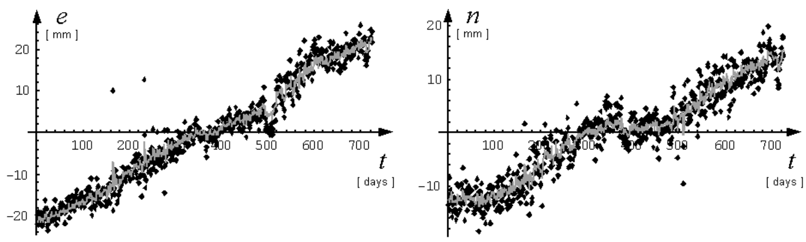

Figure 6: Original time series (black) and model (grey).

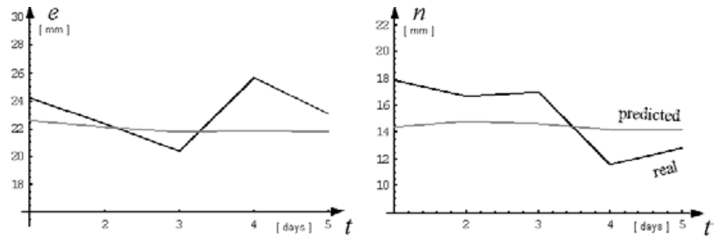

Figure 7: Prediction.

One of the two cardinal purposes of data processing (that's: to understand and be able to forecast) is the next values prediction (Fig. 8).

It can be utilized well for comparing the methods. We did it. Having computed model values for next 5 days and got the corresponding GPS measurements, we decided to quantify prediction efficiency by these measures:

- mean square error $$mse = \frac{1}{k}\sum_{t=1}^{k}(real_t - model_t)^2 ,\qquad (17)$$

- mean percentage error $$mpe = \frac{1}{k}\sum_{t=1}^{k}\frac{real_t - model_t}{real_t}100\% ,\qquad (18)$$

where k is a number of predicted time points.

4. RESULTS

First to mention are the parameters of deterministic model, i.e trend and seasonality, shown in Table 4. These results are approximately the same for all three methods (excepting those relating to *y*, *x*, of course), and serve for data description. There's pretty seen the quantity of Eurasian tectonic plate long-term drift (25.2mm per year) and the effect of seasonal forces in particular direction, too.

Table 4: Deterministic model parameters.

variable	trend [mm/year]	seasonality amplitude [mm]	seasonality period [days]
n	13.2	2.2	365
e	21.5	1.6	
y	25.2	2.5	
x	0	1.1	

What is more interesting is certainly in Table 5, which contains results from each method in separate line, namely mean square and mean percentage error of predicted values per variable. This is accompanied by the order of autoregressive model, properly chosen according to information criteria. *Mse* and *mpe* speak positively for the method that respects the presence of common trend. However, if outliers are removed using criterion of triple standard deviation (1% confidence level), better accuracy is attained (Table 6).

Table 5: Mean square and mean percentage error of predicted values

method	variable	order p	*mse* [mm^2]	*mpe* [%]
1.) independent univariate time series	*n*	1	7.40	5.08
	e	4	3.70	2.88
2.) multivariate time series	*n*	2	8.13	5.44
	e	2	5.06	5.13
3.) respecting common trend	*n*	2 (*y*)	**5.90**	**4.04**
	e	4 (*x*)	**4.10**	**2.49**

Table 6: *mse* and *mpe* of predicted values after removing outliers

method	variable	order p	*mse* [mm^2]	*mpe* [%]
1.) independent univariate time series	*n*	1	7.37	4.94
	e	4	3.34	1.85
2.) multivariate time series	*n*	4	7.52	5.01
	e	4	4.23	4.20
3.) respecting common trend	*n*	4 (*y*)	**5.59**	**0.50**
	e	4 (*x*)	**3.33**	**1.97**

REFERENCES

Bognár, T., 2005, 'Cointegration and coordinate transformation applied for geodetical time series', forthcoming *Journal of Electrical Engineering*.

Franses, P.H., 1998, '*Time Series Models for Business and Economic Forecasting*', Cambridge, Cambridge University Press.

Hefty, J. and Husár, L., 2003, '*Družicová geodézia/Globálny polohový systém*', Bratislava, Slovenská technická univerzita v Bratislave.

Komorníková, M. and Komorník, J., 2002, 'Time series models for Earth's crust kinematics', *Kybernetika*, 38, 383-387.

Reference and user's guide to Time series pack for Mathematica, 1995, Wolfram Research.

ACKNOWLEDGEMENTS

The authors acknowledge financial support from grants TARSKI COST 274, VEGA 1/0273/03 and VEGA 1/0033/04.

BIOGRAPHICAL NOTES

Tomáš Bacigál, Ing., graduated 2003 in the field Geodesy and Cartography, nowadays an PhD student of Applied Mathematics at the Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology in Bratislava

Magda Komorníková, Prof. PhD., is a full professor at the Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology in Bratislava

CONTACTS

Ing. Tomáš Bacigál
Faculty of Civil Engineering at Slovak University of Technology
Radlinskeho 11, 813 68
Bratislava
SLOVAKIA
Fax + 421 2 52961 477
Email: bacigal@math.sk,
Web site: www.math.sk/bacigal

Prof. Magda Komorníková PhD
Faculty of Civil Engineering at Slovak University of Technology
Radlinskeho 11, 813 68
Bratislava
SLOVAKIA
Tel. + 421 2 59274 428
Fax + 421 2 52961 477
Email: magda@math.sk