TESTING FOR COMMON DETERMINISTIC TRENDS IN GEODETIC DATA

Tomáš Bacigál *

Due to a long-term drift of the Eurasian tectonic plate, monitored points (GPS permanent stations) move toward north-east direction. If analyzed through time series of horizontal coordinates, a linear trend with high level of fit is found. In our paper, we test the time series of five points for common deterministic trend slopes using multivariate analysis proposed by Vogelsang and Franses (2001). Getting grouped according to direction, most of the time series clearly follow the same common trend and similar results were obtained after transforming the north-east system into common trend direction. However, some points reflects regional effects on position change, which are subject of further study.

K e y w o r d s: time series, common deterministic trend, multivariate analysis

2000 *Mathematics Subject Classification*: 37M10 Time series analysis

1 INTRODUCTION

In our work we focus on modelling time-series arisen from observations by NAVSTAR Global positioning system (GPS), performed daily in years 2001-2002 on GPS permanent stations used, among other purposes, for regular monitoring of recent kinematics of the Earth's crust (see [5]). The input data are being in the form of local topocentric horizontal coordinates (`n,e,v` - north, east and vertical) and come from 5 permanent stations denoted in an international framework as BOR1 (Poland), GOPE (Bohemia), POTS (Germany), HFLK (Austria), PENC (Hungary). However, because of significant trending behaviour only north and east component is counted in processing. The at-first-sight-visible north-east linear movement of the points ascribed to long-term drift of Eurasian tectonic plate gives us an question, if all the trends are statistically the same, i.e., if all the time series possess common linear trend, and thus may be modelled with far less parameters. Time series of BOR1 monitoring are plotted on Figure 1.

2 THEORY

In [1], we test our time series separately for the presence of stochastic trend, which might contain constant and linear deterministic component (details in [2]). After modelling both of the components, tests shows no random-walk behaviour in each data set. Here it is of interest to examine if two or more of such a trend-stationary time series have the same slope. More concretely, do all the five concerned points (realized by permanent stations) move to the north with the same upward-moving trend? The question includes also the west-east direction and the velocity in the resultant, which we get from common trend transformation theory, dealt with in [1], too.

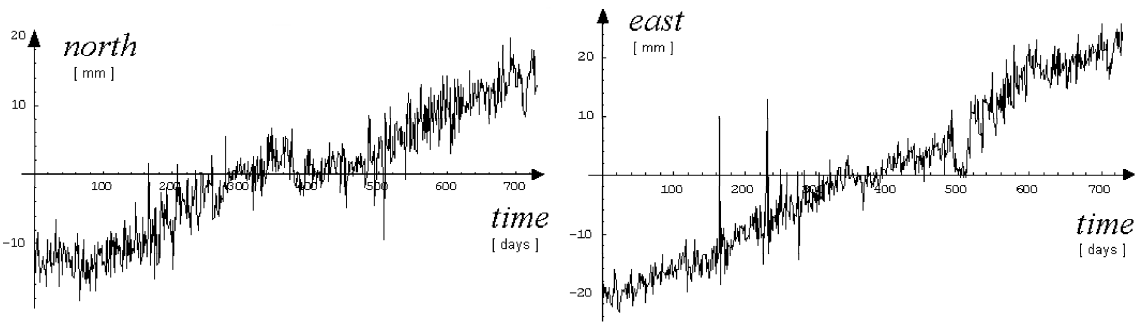

Fig. 1. Time series with linear trending. North and east coordinate.

* Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, 813 68 Bratislava, Slovakia, E-mail: bacigal@math.sk

Research supported by VEGA 1/1033/04, VEGA 1/0273/03, APVT - 20 - 003204.

ISSN 1335-3632

Such a hypothesis can be written as linear restrictions on the slope parameters across the series and we can apply the multivariate linear trend tests [3]. Consider the multivariate trend model

$$\begin{aligned} z_{1,t} &= \mu_1 + \beta_1 t + u_{1,t} \\ z_{2,t} &= \mu_2 + \beta_2 t + u_{2,t} \\ &\vdots \\ z_{k,t} &= \mu_k + \beta_k t + u_{k,t} \end{aligned} \tag{1}$$

that can be compactly written as $\boldsymbol{z}_t = \boldsymbol{\mu} + \boldsymbol{\beta} t + \boldsymbol{u}_t$, where μ and β are classical constant and linear trend parameters, u denotes residuals and k is the number of time series, in our case $k = 5$. We are interested in testing hypotheses of the form

$$\mathrm{H}_0 : \boldsymbol{R\beta} = \boldsymbol{r}\,, \qquad \mathrm{H}_1 : \boldsymbol{R\beta} \neq \boldsymbol{r}\,, \tag{2}$$

where $\boldsymbol{R}$ is $q \times k$ matrix and $\boldsymbol{r}$ is a $q \times 1$ vector of known constants. The linear hypotheses of (2) are quite general, they include linear hypotheses on slopes within given trend equations ($q = k - 1$) as well as joint trend hypotheses across equations ($q = k$). According to [3] we apply two F-tests, both test statistics are functions of the following HAC (heteroskedasticity autocorrelation) variance covariance matrix estimator. Let $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\beta}}$ denote the stacked single equation OLS estimates and $\hat{\boldsymbol{u}}_t = \boldsymbol{z}_t - \hat{\boldsymbol{\mu}} - \hat{\boldsymbol{\beta}} t$ be the residuals. Define

$$\hat{\boldsymbol{\Omega}}_{HAC} = \hat{\boldsymbol{\Gamma}}_0 + \sum_{j=1}^{n-1} (1 - \frac{j}{L})(\hat{\boldsymbol{\Gamma}}_j + \hat{\boldsymbol{\Gamma}}_j^\top)\,, \tag{3}$$

which is the Bartlett kernel estimator, where $\hat{\boldsymbol{\Gamma}}_j = \frac{1}{n}\sum_{t=j+1}^{n} \hat{\boldsymbol{u}}_t \hat{\boldsymbol{u}}_{t-j}^\top$ and L is the truncation lag or bandwidth. Usually a consistent $\hat{\boldsymbol{\Omega}}_{HAC}$ is needed, yet [3] offers an alternative, where $L = n$. Although it does not result in consistent estimator, valid testing is still possible because of asymptotic proportionality and moreover it has certain advantage coming from the choice of bandwidth. It holds that

$$\hat{\boldsymbol{\Omega}}_{L=n} = \frac{2}{n^2} \sum_{t=1}^{n} \hat{\boldsymbol{S}}_t \hat{\boldsymbol{S}}_t^\top\,, \tag{4}$$

where $\hat{\boldsymbol{S}}_t = \sum_{j=1}^{t} \hat{\boldsymbol{u}}_j$. It is also convenient to express an element of $\hat{\boldsymbol{\beta}}$ as

$$\hat{\beta}_i = \left(\sum_{t=1}^{n} \tilde{t}^2\right)^{-1} \left(\sum_{t=1}^{n} \tilde{t} z_{i,t}\right) \qquad \text{for } i = 1, 2, \ldots, k\,, \tag{5}$$

where $\bar{t} = \frac{1}{n}\sum_{t=1}^{n} t$ and $\tilde{t} = t - \bar{t}$. Now the first of test statistics can be defined

$$F_1 = (\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})^\top \left[\boldsymbol{R}(\sum_{t=1}^{n} \tilde{t}^2)^{-1} \hat{\boldsymbol{\Omega}}_{L=n} \boldsymbol{R}^\top\right]^{-1} (\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})/q\,. \tag{6}$$

Following [3] we also consider an alternative to $\hat{\boldsymbol{\Omega}}_{L=n}$ which is constructed using $\tilde{t}\hat{\boldsymbol{u}}_t$ instead of $\hat{\boldsymbol{u}}_t$. Because $\tilde{t}\hat{\boldsymbol{u}}_t$ is not a vector of stationary time series, establishing consistency of HAC estimator would be difficult if even feasible, yet again if we use $L = n$, the asymptotic behaviour of the HAC estimator can be derived. We can write

$$\tilde{\boldsymbol{\Omega}}_{L=n} = \frac{2}{n^2} \sum_{t=1}^{n} \tilde{\boldsymbol{S}}_t \tilde{\boldsymbol{S}}_t^\top\,, \tag{7}$$

where $\tilde{\boldsymbol{S}}_t = \sum_{j=1}^{t} (j - \bar{t})\tilde{\boldsymbol{u}}_j$, and then the second test statistic is

$$\begin{aligned} F_2 = n(&\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})^\top \\ &\left(\boldsymbol{R}(\frac{1}{n}\sum_{t=1}^{n} \tilde{t}^2)^{-1} \tilde{\boldsymbol{\Omega}}_{L=n} (\frac{1}{n}\sum_{t=1}^{n} \tilde{t}^2)^{-1} \boldsymbol{R}^\top\right)^{-1} \\ &(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})/q\,. \end{aligned} \tag{8}$$

The null hypothesis in (2) is rejected if test statistic F_1 (F_2) exceeds critical value given for q restrictions in [3, Tab.3 (Tab.2)]. It is worth noting that due to practical reasons indices of the F-statistics has been swapped in our work.

The asymptotic distribution theory for these F statistics is nonstandard and was developed for the case where the errors are covariance stationary. Simulation evidence reported by [3] suggests that the F-tests suffers much less from over-rejection problem caused by strong positive serial correlation than the compared standard alternative, whereas the power of F-s is slightly lower. Finite sample simulation evidence in [3] also suggested that the performance of the tests are improved when $\hat{\boldsymbol{\Omega}}$ estimator is computed using VAR(1) prewhitening. However, this we do not do here.

The standard alternative to F_1 and F_2 is a Wald test based on consistent $\hat{\Omega}_{HAC}$ estimator, which uses the same Bartlett kernel. For $\hat{\Omega}_{HAC}$ to be consistent, the bandwidth L must increase as the sample increases but at the slower rate. As referred in [3], the rate $\sqrt[3]{n}$ minimizes the approximate mean square error for $\hat{\Omega}$ and considering this in (3), the Wald test is defined as

$$W = (\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})^\top \left[\boldsymbol{R}(\sum_{t=1}^{n} \tilde{t}^2)^{-1} \hat{\boldsymbol{\Omega}}_{HAC} \boldsymbol{R}^\top\right]^{-1} (\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})\,.$$

Asymptotic distribution of the Wald test is χ^2 with q degrees of freedom.

For reference we recommend to see an interesting application of this theory in [4] too.

3 APPLICATION

So we have got 5 time series from permanent stations denoted BOR1, GOPE, POTS, HFLK, PENC and they make vector $\boldsymbol{z}_t$ in respective order. Table 1 contains OLS estimates of parameter β in units mm/year for the north and east coordinate time series n_t, e_t and even for common trend direction time series y_t (see [1]).

Table 1. Trend parameter estimates [mm/year]

point:	BOR1	GOPE	POTS	HFLK	PENC
$\hat{\beta}_{n_t}$	14.2	14.7	14.4	13.4	12.2
$\hat{\beta}_{e_t}$	22.3	23.1	21.1	21.5	23.9
$\hat{\beta}_{y_t}$	26.4	27.4	25.5	25.3	26.9

By defining the $q \times k$ matrix $\boldsymbol{R}$ and $q \times 1$ vector $\boldsymbol{r}$

$$\boldsymbol{R} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \boldsymbol{r} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_1 \\ \hat{\beta}_1 \\ \hat{\beta}_1 \end{pmatrix},$$

the null hypothesis (2) says that the first (reference) time series has the same slope as the rest of time series. Moving the zero column in $\boldsymbol{R}$ and filling $\boldsymbol{r}$ with corresponding parameter, slopes of other series are tested.

We also consider a joint test, when $q = k$, $\boldsymbol{R}$ is k-dimensional identity matrix and $\boldsymbol{r}$ contains trend parameter of an average time series, $\bar{n}_t = \frac{1}{k}\sum_{j=1}^{k}(\boldsymbol{n}_t)_j$ for instance. Test results are collected in Table 2 and 3, critical values for F-tests are given in a separate line bellow. Every Wald test statistic is provided with p-value, which (given null hypothesis is valid) represents a probability that we get the particular value or even the more contradictory to tested hypothesis, in other words it expresses a probability the trend slopes are statistically the same.

4 CONCLUSIONS

First to conclude, all the tests confirmed common deterministic trend in north direction coordinate for all the series except for PENC. Although the joint test for $q = 5$ did not reject the null hypothesis due to quite a large set of variables, from the five tests within equations it is obvious that this last slope does not correspond to the others. As for other two variables, e_t and y_t, no significant deterministic relation was found. Therefore next we took foursomes into account (Table 3) removing seemingly most troublesome series, and observed how it changed the test statistics. Removing PENC accentuated the common trend in n_t, and pointed slightly at promises in y_t. It is useful now to repeat that slope of y_t represents a velocity

Table 2. F_1, F_2 and Wald test for all 5 time series and selected triads.

station of reference trend	F_1			F_2			W					
	n_t	e_t	y_t	n_t	e_t	y_t	n_t	p-value	e_t	p-value	y_t	p-value
BOR1	17.2	96.6	501.4	16.7	57.7	281.0	0.82	0.935	2.85	0.583	13.88	0.008
GOPE	34.5	68.6	75.4	29.0	95.7	64.6	1.43	0.839	4.73	0.216	3.19	0.526
POTS	22.6	215.1	314.1	19.0	209.3	171.4	0.94	0.919	10.34	0.035	8.46	0.075
HFLK	32.3	159.6	415.7	31.3	188.6	215.8	1.55	0.818	9.31	0.054	10.66	0.031
PENC	143.2	338.8	526.3	198.0	170.0	290.4	9.78	0.044	8.39	0.078	14.34	0.006
Joint	21.5	218.7	423.0	24.5	127.2	218.9	1.50	0.913	7.85	0.164	13.51	0.019
Critical value for q=4 and α=5%: 46.8 (F_1), 43.8 (F_2) and for joint test(q=5): 70.1 (F_1), 78.3 (F_2)												
BOR1	0.6	1.4	1.5	1.0	3.2	3.8	0.03	0.988	0.08	0.962	0.09	0.954
POTS	2.4	4.8	3.0	3.8	6.5	2.9	0.10	0.953	0.16	0.923	0.07	0.964
HFLK	7.9	14.9	11.2	7.9	9.0	7.0	0.19	0.907	0.22	0.895	0.17	0.917
Joint	6.0	33.1	16.5	4.6	33.2	12.8	0.17	0.982	1.23	0.746	0.48	0.924
BOR1	16.4	27.8	4.8	27.9	22.6	5.0	0.69	0.709	0.56	0.756	0.12	0.940
GOPE	9.5	82.6	4.4	24.9	78.3	3.1	0.61	0.735	1.93	0.381	0.08	0.963
PENC	124.2	18.7	4.6	137.7	14.7	4.4	3.40	0.183	0.36	0.834	0.11	0.948
Joint	30.8	117.9	46.4	33.8	129.3	26.4	1.25	0.740	4.79	0.188	0.98	0.807
Critical value for q=2 and α=5%: 40.7 (F_1), 43.8 (F_2) and for joint test(q=3): 68.7 (F_1), 73.4 (F_2)												

Table 3. Joint test for common trend slope of 4-time-series set.

Station removed from set	F_1			F_2			W					
	n	e	y	n	e	y	n	p-value	e	p-value	y	p-value
POTS	25.1	295.0	149.4	36.7	131.4	94.6	1.81	0.771	6.49	0.166	4.67	0.323
HFLK	24.5	120.6	252.7	23.4	119.6	157.3	1.15	0.886	5.91	0.206	7.77	0.100
PENC	11.5	119.3	504.4	12.0	48.4	266.9	0.59	0.964	2.39	0.665	13.18	0.010
Critical value for joint test (q=4 and α=5%): 69.3 (F_1), 76.7 (F_2)												

in the direction a particular point moves, and does not contain information about direction.

At last, we tested triads, two of which appear quite interesting and are summarized in Table 3. First (BOR1, POTS, HFLK) was found to have statistically the same linear trend significant for every direction, though the p-value of Wald test leaves some space for questions. In short and simple, these three points move in the same direction and at the same speed. Second (BOR1, GOPE, PENC) shows no significant harmony in direction, yet it indicates a common velocity.

Honestly, we don't know the cause of this minor effects in tested parameters, there may be some speculations uttered about local instability of given points in particular direction or some residual systematic components in time series that weren't taken into account in model specification and which might cause spurious estimates. Nevertheless, although the data visibly show significant linear trend behaviour caused by tectonic plate drift, the tests rejected common deterministic trend for two of the observed points. It is subject to study, why this happened.

Acknowledgement

The author gratefully acknowledges many helpful suggestions of Professor Magda Komorníková.

References

[1] BACIGÁL, T.—KOMORNÍKOVÁ, M.: Modelling points position time series in the light of cointegration, Proceedings of INGEO 2004 (CD-ROM edition), Bratislava, 2004.

[2] FRANSES, P.H.: Time Series Models for Business and Economic Forecasting, Cambridge University Press, Cambridge, 1998.

[3] FRANSES, P.H.—VOGELSANG, T.J.: Testing for Common Deterministic Trend Slopes, Center for Analytic Economics Working Paper, Cornell University, 2002, pp. 1–15.

[4] FOMBY, T.B.—VOGELSANG, T.J.: Tests of Common Deterministic Trend Slopes Applied to Quarterly Temperature Data, Advances in Econometrics **17** (2003), 29–43, chapter 2, Elsevier Science.

[5] HEFTY, J.: The permanent Modra-Piesok GPS station and its long-term and short-term stability, Slovak Journal of Civil Engineering (2001), 31-37.

Received ?? June 2005

Tomáš Bacigál (Ing) graduated in geodesy and cartography at the Faculty of Civil Engineering of the Slovak University of Technology, Bratislava, at present he is a PhD student in applied mathematics at the same faculty. His supervisor is Professor Magda Komorníková.