Testing for Common Deterministic Trends in Geodetic Data

PhD. student: Ing. Tomáš Bacigál
Supervisor: Prof. RNDr. Magda Komorníková, CSc.

Department of Mathematics and Descriptive Geometry
Faculty of Civil Engineering
Slovak University of Technology Bratislava

ISCAM 2005

Topics:

- Point on the Earth's surface
  - dynamics and concerns
- Time series analysis
  - a way to learn
- Testing for common trend
- Results
  - information or confusion?

Point on the Earth

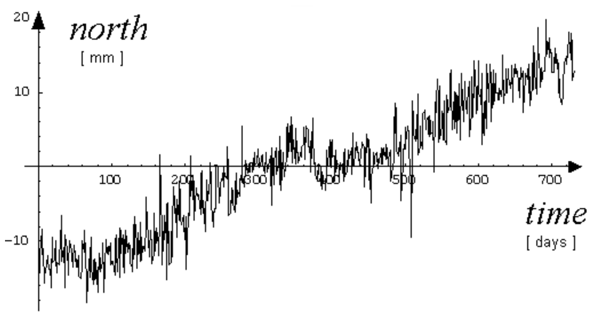
- geometrical and physical quantities related to point
- variation being temporarily or permanently monitored by means of geodesy
- satellite navigational systems support the research activities in geodesy and geophysics
- NAVSTAR GPS
- permanent stations for precise determination of position
- monitoring Earth's crust kinematics, hydrosphere, atmosphere and ionosphere variation, rotation of Earth, disturbing forces in the gravity field...
- results serve in civil sector for practical purposes

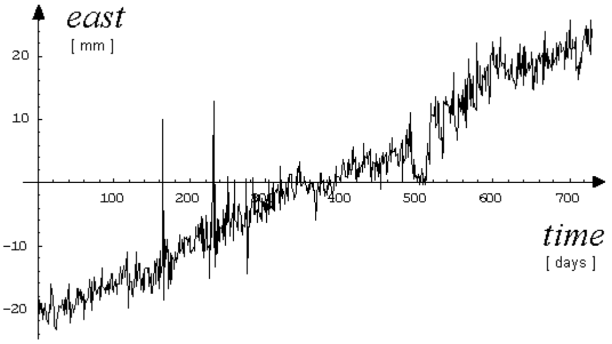
Time series

- time ordered data taken from observations of some phenomenon
- purpose:

  to understand the underlying mechanism and

  to forecast
- daily GPS observations in 2 years period = 730 time points
- local topocentric horizontal coordinate system (n,e,v) → north and east component time series (significant trending due to Eurasian tectonic plate drift)





Testing

Question:
Do all five concerned points (realized by permanent stations) move to the north (east) at the same speed?

Then, it is of our interest to examine if $k = 5$ given trend-stationary time series have the same deterministic trend slope.
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.

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 ($q = k$).

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 a heteroskedasticity autocorrelation (HAC) variance covariance matrix estimator

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

which in this particular case use the Bartlett kernel, 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 Franses and Vogelsang (2002) 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.

Now, let's define two **test statistics** coming from this theory:

1.

$$F_1 = (\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})^\top \left[\boldsymbol{R}(\sum_{t=1}^{n} \tilde{t}^2)^{-1} \hat{\boldsymbol{\Omega}}_{L=n} \boldsymbol{R}^\top \right]^{-1} (\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})/q \,. \tag{4}$$

where
$\hat{\boldsymbol{\Omega}}_{L=n} = \frac{2}{n^2}\sum_{t=1}^{n} \hat{\boldsymbol{S}}_t \hat{\boldsymbol{S}}_t^\top$,
$\hat{\boldsymbol{S}}_t = \sum_{j=1}^{t} \hat{\boldsymbol{u}}_j$,
$\bar{t} = \frac{1}{n}\sum_{t=1}^{n} t$ and $\tilde{t} = t - \bar{t}$

2.

$$F_2 = n(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})^\top \left[\boldsymbol{R}(\frac{1}{n}\sum_{t=1}^{n} \tilde{t}^2)^{-1} \tilde{\boldsymbol{\Omega}}_{L=n} (\frac{1}{n}\sum_{t=1}^{n} \tilde{t}^2)^{-1} \boldsymbol{R}^\top \right]^{-1} (\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})/q \,, \tag{5}$$

where $\tilde{\boldsymbol{\Omega}}_{L=n} = \frac{2}{n^2}\sum_{t=1}^{n} \tilde{\boldsymbol{S}}_t \tilde{\boldsymbol{S}}_t^\top$ and $\tilde{\boldsymbol{S}}_t = \sum_{j=1}^{t} (j - \bar{t})\tilde{\boldsymbol{u}}_j$.

The null hypothesis in (2) is rejected if test statistics F_1, F_2) exceed critical value given for q restrictions. The asymptotic distribution theory for these statistics is nonstandard and was developed for the case where the errors are covariance stationary. Simulation evidence reported by Franses and Vogelsang (2002) 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.

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. 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

3.

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

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

For our data, $q \times k$ matrix $\boldsymbol{R}$ and $q \times 1$ vector are defined

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

in the case we want the null hypothesis (2) to say that the first (reference) time series has the same slope as the rest of time series. **Otherwise** a joint test is considered, 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.

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	n_t	e_t	n_t	p-value	e_t	p-value
BOR1	17.2	96.6	16.7	57.7	0.82	0.935	2.85	0.583
GOPE	34.5	68.6	29.0	95.7	1.43	0.839	4.73	0.216
POTS	22.6	215.1	19.0	209.3	0.94	0.919	10.34	0.035
HFLK	32.3	159.6	31.3	188.6	1.55	0.818	9.31	0.054
PENC	143.2	338.8	198.0	170.0	9.78	0.044	8.39	0.078
Joint	**21.5**	218.7	**24.5**	127.2	1.50	**0.913**	7.85	0.164
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	1.0	3.2	0.03	0.988	0.08	0.962
POTS	2.4	4.8	3.8	6.5	0.10	0.953	0.16	0.923
HFLK	7.9	14.9	7.9	9.0	0.19	0.907	0.22	0.895
Joint	**6.0**	**33.1**	**4.6**	**33.2**	0.17	**0.982**	1.23	**0.746**
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)								

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

Conclusion

- For the set of 5 series

  All joint tests confirmed common deterministic trend in north direction for all series. However, by individual testing, PENC was found to be an exception.

  As for east direction, no significant deterministic relation was found.

- Only one triad (BOR1, POTS, HFLK) seems to have statistically significant common linear trend in both direction.

- Honestly, we don't know the cause of this minor effects in tested parameters. We may speculate 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.

- Anyway, 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.