

Time variations and anomalies in the air pressure admittance of superconducting tidal gravity data

J. Arnos¹⁾, B. Ducarme²⁾, A.P. Venedikov³⁾ and R. Vieira¹⁾

- 1) Instituto de Astronomía y Geodesia (CSIC-UCM). Universidad Complutense de Madrid. Facultad de Matemáticas. 28040 Madrid, Spain. arnos@iagmat1.mat.ucm.es, vieira@iagmat1.mat.ucm.es
- 2) Royal Observatory of Belgium (Belgian National Fund for Scientific Research), Av.Circulaire 3, B-1180 Brussels. ducarme@oma.be
- 3) Geophysical Institute, Bulgarian Academy of Sciences, Acad. Bonchev street block 3, Sofia 1113, vened@geophys.bas.bg

Abstract

The new VAV tidal analysis program is designed to deal simultaneously with non-tidal as well as tidal signals. The main goal of this paper is to study a well-known non-tidal signal: the air pressure influence, through the time variations of its admittance in some data sets of superconducting gravimeters from the Global Geodynamics Program (GGP). Unhappily, except in Boulder, where seasonal variations of the air pressure influence are obvious, the detected anomalies are due to gross errors in the data which are reducing strongly the signal to noise ratio. Thus the first conclusion is a trivial one: it is impossible or too difficult to find useful non-tidal signals, when the data still contain gross errors, i.e. non-useful non-tidal signals. Another conclusion of our investigation is that non-careful interpolations can generate very bad, non-stationary noise in data series of very good quality. It is thus important to reduce to a minimum any “repair” of the data prior to the analysis. The amount of “repair” depends strongly from the analysis method. With its ability to analyze non-equidistant data with any sampling rate, VAV can simply suppress perturbed data and ignore gaps.

1. Introduction

The Earth tide data processing has to deal with the general model:

$$\text{Tidal data} = \text{Tidal signal} + \text{Non-tidal signals} + \text{noise}$$

The tidal signal is the useful signal for the tidal domain, e.g. in connection with the Earth’s structure and properties. Nevertheless, the model above clearly shows that the lack of attention paid to the non-tidal signals may badly affect just the useful signal.

For other domains, in particular the so-called “Earth deformations”, the non-tidal signals are useful signals, e.g. for the search of earthquake and volcano precursors. In this case the tidal signal is only a noise. However, just in the same way as above, a careless treatment of the tidal signal will spoil the detection of the useful non-tidal signals. Examples of such inadequate treatments are the eliminations of the tidal signal through an elementary filtration.

A conclusion from these simple considerations is that the tidal data processing should deal, for whatever purpose, most carefully with both tidal and non-tidal signals.

A general problem of the model is how to distinguish between noise and non-tidal signals.

All methods for Earth tide data analysis accept implicitly or explicitly that the noise of the data is a random stationary phenomenon. Hence, all components of the data, which are not tidal waves and, in the same time, which are non-stationary phenomena, should be considered as non-tidal signals. Typical examples are instrumental drift, jumps, too big residuals, time variations of the tidal parameters, etc.

A problem for the use of the non-tidal signals is that

$$\text{non - tidal signals may be } \begin{cases} \text{useful signals of geophysical origin or} \\ \text{anomalies or perturbations of instrumental and human origin} \end{cases}$$

The problem is to distinguish between these two phenomena. One thing is clear. If we are sure that the anomalies are rare, it is easy to find the useful signals and, vice versa, if there are too many anomalies, it may be too difficult or impossible to distinguish the useful signals. A bad news is that some of the anomalies may be generated by improper manipulation and processing of the data. The good news is that we have the possibility to avoid all improper manipulations and apply a correct data processing.

This present paper is an attempt to study a well known non-tidal signal, the time variations of the air-pressure admittance, i.e. the variations of the cross-regression coefficient, say b , of the observed tidal gravity with the air-pressure. We usually suppose that b is a constant. The deviations from this hypothesis may indicate that either we need an improvement of the model of the admittance or we have a kind of anomalous signal.

The results presented in this paper have been obtained through special options of the new computer program VAV (Venedikov et al., 2001). In the next section 2 we shall describe the algorithm of the option.

In the following sections we shall give some examples of application. They have been obtained through the application of VAV on 3 series of superconducting gravity data (Table 1) of the Global Geodynamic Project (GGP) (Crossley, 2000), collected in the International Centre for Earth Tides (Ducarme et al., 2000).

Table 1. List of series of superconducting gravimeter data.

Station, Country	Instrument or sensor	Latitude, Longitude	Time interval
Boulder, USA	GWR CO24	40.13°, 254.77°	12.04.1995-01.08.1998
Cantley, Canada	GWR T012	45.58°, 284.19°	07.11.1989-31.12.1998
Strasbourg, France	GWR T005, GWR C026	48.62°, 7.68°	11.07.1987-25.06.1996 01.03.1997-30.04.1998

In the examples we have used, in parallel to the time variations of the b coefficients, residuals of the filtered numbers (e.g. Figures 2 and 3 in section 3), provided by VAV. These residuals are a convenient tool to find non-tidal signals because we get them at once for a whole time window, e.g. every 48 hours. The filtered numbers, as well as the residuals are complex numbers, but we use only the real modulus. A threshold level (the horizontal straight line in the graphics) is computed with very high confidential probability. In such a way every value, which exceeds this level, can be considered as a non-tidal signal with a high confidence.

2. The main algorithm

The trivial way to study the time variations of some parameters is to partition the data into segments and process the data, separately in every segment. Here we propose to apply a global analysis on the whole series of the data, but to accept that in every segment we have a different or individual regression coefficient. In such a way we get global estimates of the tidal parameters, by using a highest possible separation in tidal groups, accompanied by a set of regression coefficients, related with every segment.

Generally, the multi channel analysis is based on model equations like

$$\mathbf{U} = \mathbf{Ax} + \mathbf{Vb} + \mathbf{e}, \quad (1)$$

where \mathbf{U} is a vector (column vector) of filtered tidal data, after the elimination of the drift, \mathbf{x} is a vector of the tidal unknowns $\xi = \delta \cos \kappa$ and $\eta = -\delta \sin \kappa$ for a set of tidal groups, \mathbf{A} is a known matrix, appropriately created by using the theoretical amplitudes and phases, \mathbf{b} is a vector of unknown cross-regression coefficients, representing one or several admittance functions, \mathbf{V} is a known matrix created by using, in one or another way, the observed values of the 2nd, 3rd, ... channels, e.g. air-pressure, temperature etc. and \mathbf{e} is the noise.

The estimation of the time variations of \mathbf{b} is always related with a partition of the data into N time segments: $S(t_1), S(t_2), \dots, S(t_N)$ related with the epochs t_1, t_2, \dots, t_N , as shown in Figure 1

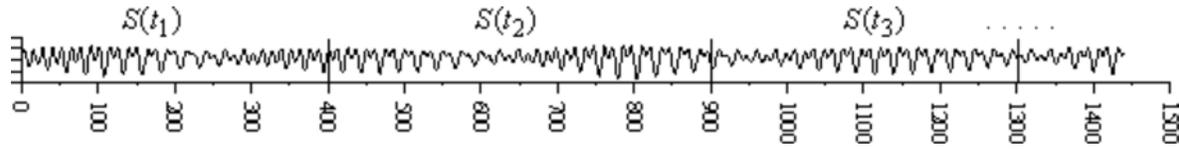


Figure 1. Partition of the data into segments.

Corresponding to the partitioning we have

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \\ \vdots \\ \mathbf{U}_k \\ \vdots \\ \mathbf{U}_N \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_N \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \\ \vdots \\ \mathbf{V}_N \end{pmatrix} \quad \text{and} \quad \mathbf{e} = \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_N \end{pmatrix} \quad (2)$$

where $\mathbf{U}_j, \mathbf{A}_j, \mathbf{V}_j$ and \mathbf{e}_j are matrix components of $\mathbf{U}, \mathbf{A}, \mathbf{V}$ and \mathbf{e} respectively, related with the segment $S(t_j), j = 1, \dots, N$.

One possibility to obtain the time variations is to apply the general equation (1) separately for every $S(t_j)$, i.e. to deal with the equations

$$\mathbf{U}_j = \mathbf{A}_j \mathbf{x}_j + \mathbf{V}_j \mathbf{b}_j + \mathbf{e}_j \quad \text{separately for every } S(t_j), j = 1, \dots, N \quad (3)$$

From the separate solution of these N systems we shall get the estimates

$$\tilde{\mathbf{x}}_j = \tilde{\mathbf{x}}(t_j) \quad \text{and} \quad \tilde{\mathbf{b}}_j = \tilde{\mathbf{b}}(t_j), j = 1, \dots, N \quad (4)$$

i.e. the tidal unknowns and the regression coefficients as discrete functions of the time.

A weak point is that $\tilde{\mathbf{x}}_j$ so obtained are not estimates of the global tidal unknowns \mathbf{x} in (1). E.g., for data larger than 1 year, \mathbf{x} in (1) may involve unknowns for each of the tidal groups: CHI1, PI1, P1, S1, K1, PSI1, PHI1 and TET1, all of them having different amplitude factors. If the segments are as short as a few months, all these groups should be included in one and the same group K1. Hence the elements of $\tilde{\mathbf{x}}_j$ are not the same as those of \mathbf{x} and they are estimated with a lower precision. Something more, the association of groups with different amplitude factors in one group will produce time variations that may affect the estimated regression coefficients.

In this relation VAV works with a single system of equations.

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \\ \vdots \\ \mathbf{U}_N \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_N \end{pmatrix} \mathbf{x} + \begin{pmatrix} \mathbf{V}_1 & & \mathbf{O} \\ & \mathbf{V}_2 & \\ & & \ddots \\ \mathbf{O} & & & \mathbf{V}_N \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_N \end{pmatrix} \quad (5)$$

Thus the terms \mathbf{U} and \mathbf{A} from (1) remain the same, but the matrix \mathbf{V} and the vector \mathbf{b} are considerably transformed.

From the solution of (5) by the method of the least squares we get global estimates $\tilde{\mathbf{x}}$ of the same tidal unknowns \mathbf{x} in (1) but a set of estimates $\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2, \dots, \tilde{\mathbf{b}}_N$ of the regression coefficients $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_N$. Since $\tilde{\mathbf{b}}_j$ is related with the epoch t_j we can consider $\tilde{\mathbf{b}}_j = \tilde{\mathbf{b}}(t_j)$ as a discrete function of the time that describes the time variations of the regression coefficients.

If the number N of the segments is high, the number of the unknowns may become too high which can embarrass the computations. Due to this it is convenient to use the classical algorithm for the separation of the unknowns. In our case it consists in the following way to obtain the estimates.

We transform the matrices \mathbf{A}_j into \mathbf{C}_j through

$$\mathbf{C}_j = \mathbf{A}_j - \mathbf{V}_j(\mathbf{V}_j^T \mathbf{V})^{-1} \mathbf{V}_j^T \mathbf{A}_j \text{ and create the matrix } \mathbf{C} = \begin{pmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \vdots \\ \mathbf{C}_N \end{pmatrix} \quad (6)$$

Then the estimates of the tidal unknowns are directly obtained through

$$\tilde{\mathbf{x}} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{U} \quad (7)$$

where

$$\mathbf{C}^T \mathbf{C} = \sum_{j=1}^N \left(\mathbf{A}_j^T \mathbf{A}_j - \mathbf{A}_j^T \mathbf{V}_j (\mathbf{V}_j^T \mathbf{V}_j)^{-1} \mathbf{V}_j^T \mathbf{A}_j \right) \quad (8)$$

$$\mathbf{C}^T \mathbf{U} = \sum_{j=1}^N \left(\mathbf{A}_j^T \mathbf{U}_j - \mathbf{A}_j^T \mathbf{V}_j (\mathbf{V}_j^T \mathbf{V}_j)^{-1} \mathbf{V}_j^T \mathbf{U}_j \right) \quad (9)$$

Afterwards, we can compute the estimates $\tilde{\mathbf{b}}_j$ through

$$\tilde{\mathbf{b}}_j = (\mathbf{V}_j^T \mathbf{V}_j)^{-1} \mathbf{V}_j^T (\mathbf{U}_j - \mathbf{A}_j \tilde{\mathbf{x}}) \quad (10)$$

The results in the following sections are obtained for a single auxiliary channel – the air-pressure - and a single regression coefficient for every segment. In this case the vectors $\tilde{\mathbf{b}}_j$ become a scalar b , which is the cross-regression coefficient and the expressions above are considerably simplified.

3. Station Cantley

Figure 3, compared to Figure 2, shows the strong effect of the application of a usual multi-channel analysis, to take into account the effect of the air-pressure, with a constant in time cross-regression coefficient b . In addition to the evident decrease of the level of the residuals, there is a huge decrease of the

AIC value

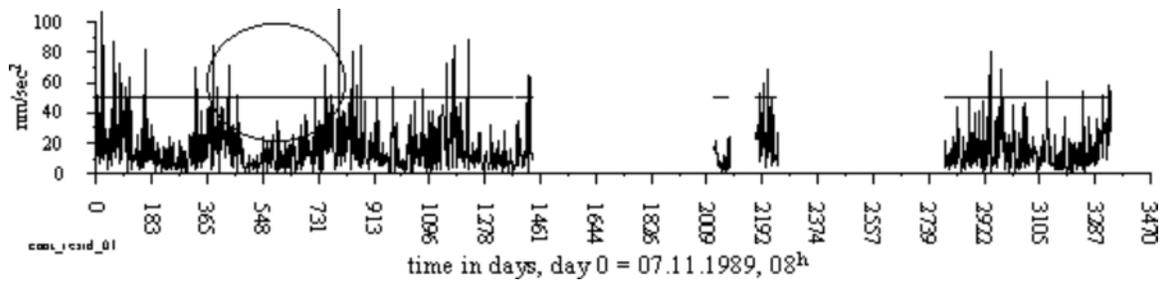


Figure 2. Cantley, modulus of residuals in the D-frequency domain; the effect of the air-pressure is ignored, i.e. no cross-regression. Akaike criterion for the analysis:

AIC = 88,495.

A curious phenomenon is manifested at the place, indicated by a circle. In Figure 2 the residuals are relatively low. Surprisingly, in Figure 3, at the same place, they become relatively high. It looks like a noise, namely anomaly or non-tidal signal, generated by the reduction of the air-pressure effect.

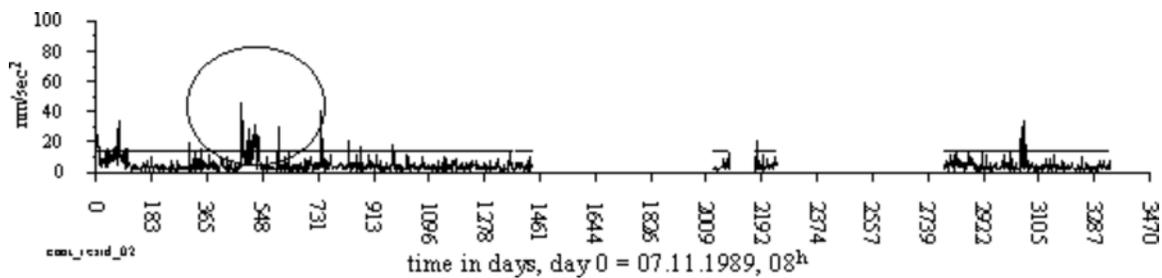


Figure 3. Cantley, modulus of residuals in the D-frequency domain; the effect of the air-pressure is taken into account, i.e. a cross-regression model is applied with a constant coefficient b ; AIC= 60,201.

Figure 4 shows an attempt to check eventual time variations of b . There are 3 segments, defined by two large gaps. In parallel with the horizontal lines, giving the value of b in every segment, the 95% confidential intervals are given. These confidence intervals shows clearly that we have a significantly lower b value in the first segment.

It should be noticed that the m.s.d. (mean square deviations) of b are usually determined, e.g. by the ETERNA program (Wenzel,1996), on the basis of a white noise assumption. VAV computes the m.s.d. of b on the assumption of a colored noise, i.e. we get frequency dependent m.s.d. In the case when a single b coefficient is determined, the m.s.d. is determined by using the m.s.d. of the data at the lowest frequency, i.e. the highest possible m.s.d. Due to this VAV provides higher m.s.d. of b than other programs, e.g. than ETERNA. This gives us more confidence when a significant difference is observed. The significance of the time variation is also supported by a further decrease of AIC.

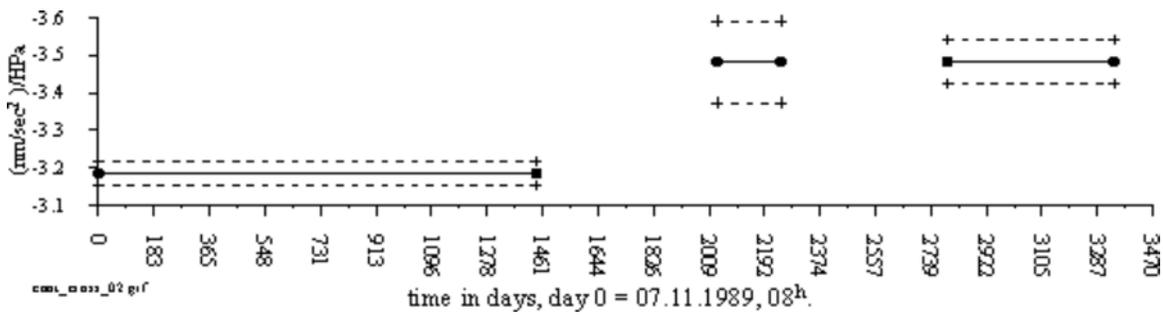


Figure 4. Cantley, time variations of the regression coefficient b . The data are partitioned into 3 segments, defined by two important gaps; AIC = 60,020.

The next Figure 5 shows a more detailed study of the time variations of b . There is an obviously significant decrease of b , around $t = 548$ days. A further decrease of AIC indicates the reality of this deviation.

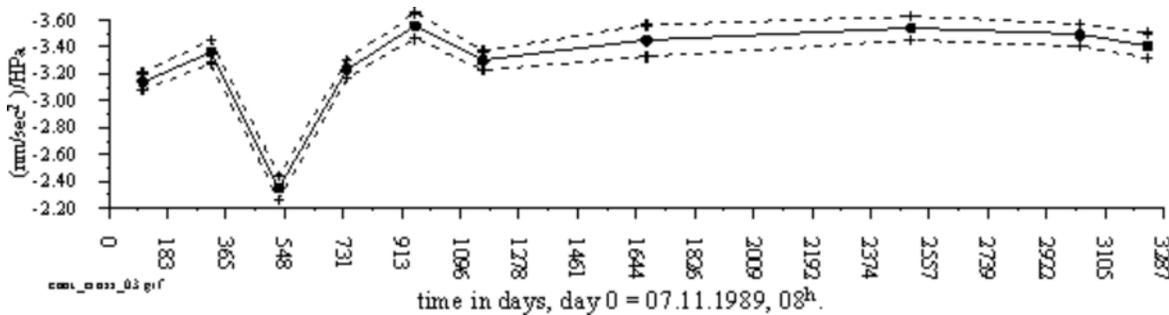


Figure 5. Cantley - time variations of the coefficient b . The data are partitioned into 10 segments, each segment including one and the same number of days; AIC = 59,151.

After several experiments we succeeded to isolate a segment in the interval of t (480, 540 days) (Figure 6) for which we have an extremely low coefficient $b \approx 0$. For this case we have got AIC considerably lower than in the other cases, which shows the reality of this strange value of b .

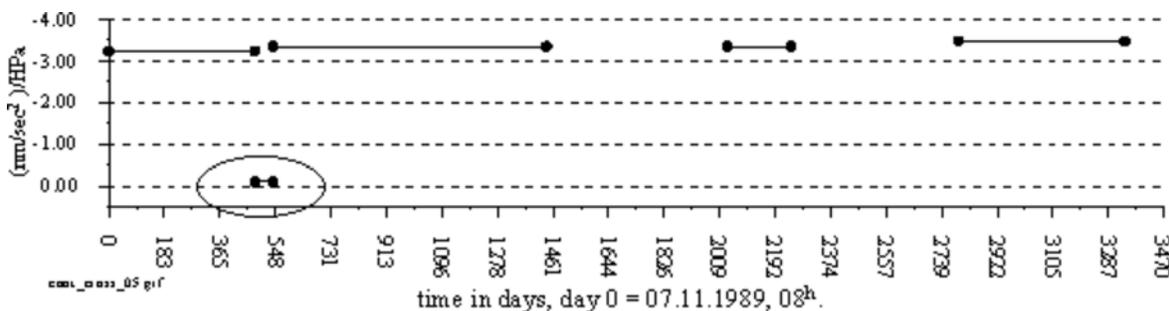


Figure 6. Cantley - time variations of the coefficient b ; the segments are defined by two large gaps and the interval of (480, 540 days) of length 60 days; the regression coefficient in the interval is $b \approx 0$; AIC = 56,881.

As shown in Figure 7, when we accept the model of time variation of b shown in Figure 6, the anomalous residuals in Figure 3 have disappeared.

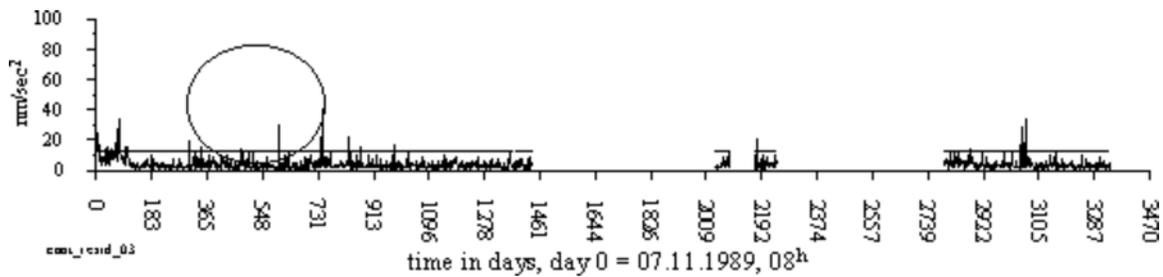


Figure 7. The D-residuals corresponding to the time variations of the b – coefficient given in Figure 6.

Unfortunately, this strange phenomenon cannot be any useful non-tidal signal. It is an example of what we have defined in section 1 as an anomaly and a serious one. Our supposition is that the data in the controversial interval have been repaired in one or another way, most likely through interpolation, but without restoring the atmospheric pressure effect. Thus, instead of finding an useful signal, we found an example showing how a careless intervention in the data, may be an unnecessary interpolation, can generate noise.

Table 2. Results of the analysis of a series of data 01.07.1990–1.11.1991 in variants: whole series without any gap and the series with a gap of 60 days, at the place of the controversial interval 02.03.1991 – 01.05.1991 (480-540 days).

Diurnal tides						
Data used	d(Q1)	m.s.d.	d(O1)	m.s.d.	d(K1)	m.s.d.
Whole series	1.16568	±.00141	1.16648	±.00028	1.14836	±.00021
With gap 60 days	1.16474	±.00085	1.16657	±.00017	1.14823	±.00013
Semidiurnal tides						
Data used	d(N2)	m.s.d.	d(M2)	m.s.d.	d(S2)	m.s.d.
Whole series	1.20961	±.00064	1.20429	±.00013	1.18396	±.00027
With gap 60 days	1.20964	±.00041	1.20429	±.00008	1.18369	±.00018
Cross-regression coefficient (admittance)						
Data used	b	m.s.d.				
Whole series	-2.8264	±0.0568				
With gap 60 days	-3.3201	±0.0358				

The results in Table 2 show that when such data are introduced, we get worse results instead of an improvement or, which is the same, that the exclusion of such data actually improves the results.

4. Station Boulder

Figures 8 and 9 are similar to Figures 2 and 3 in the previous section. They also demonstrate the strong effect of the cross-regression, with an important reduction of the AIC value.

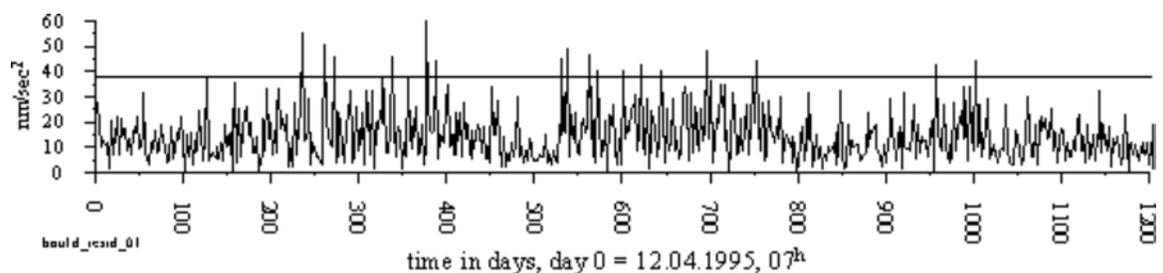


Figure 8. Boulder, residuals (modulus) in the D-frequency domain; the effect of the air-pressure is ignored, i.e. no cross-regression model; AIC = 46,504

Unlike the case of Figure 3, in Figure 9 we see very few residuals over the threshold level. The indicated case is interesting because it does not exist in Figure 8. It turned out, that this anomaly is due to some interpolated air-pressure data. It is again an example of anomaly, generated by an unnecessary data manipulation.

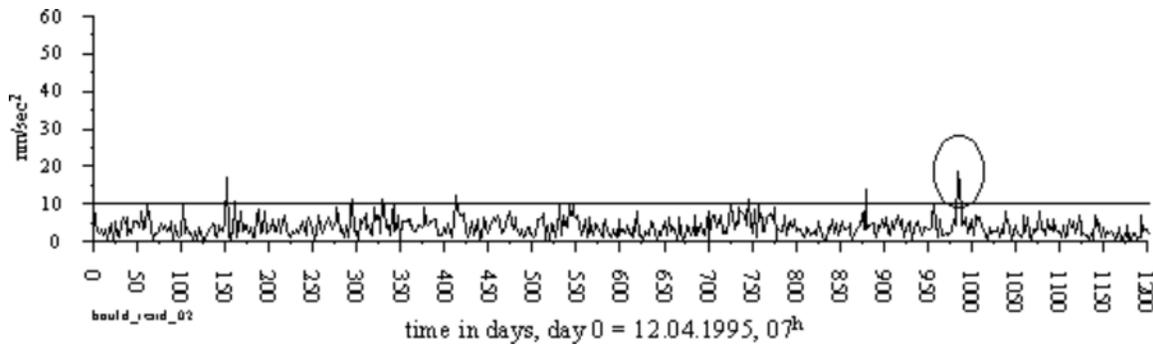


Figure 9. Boulder, residuals (modulus) in the D-frequency domain; the effect of the air-pressure is taken into account. A cross-regression model is applied with a constant coefficient b ; AIC= 29,121.

Otherwise, Figure 9 shows data with very good general behavior that are certainly interesting to study for time variations of the air-pressure admittance.

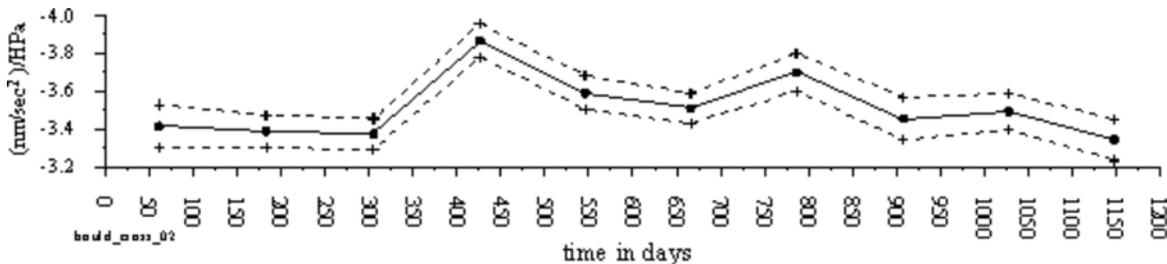


Figure 10. Time variations of the b coefficients; the data are partitioned into 10 equal segments; AIC = 28,913.

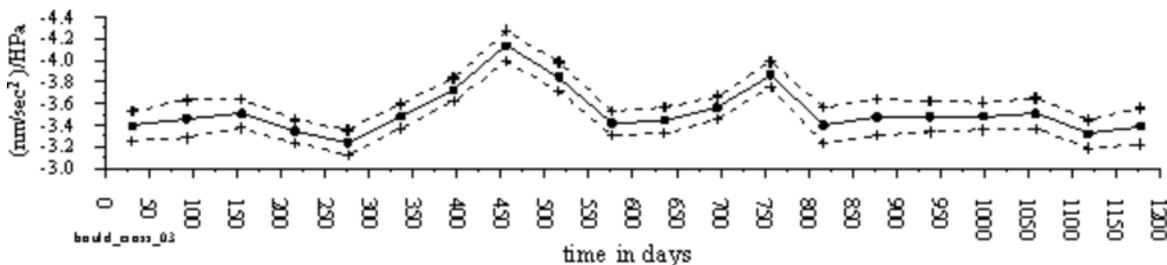


Figure 11. Time variations of the b coefficients; the data are partitioned into 20 equal segments; AIC = 28,809.

Figures 10 and 11 represent the time variations of the b coefficients, when the data are partitioned into 10 and 20 equal segments respectively. We have a reduction of the AIC value, compared to Figure 9. This means that we have really some variations of b . The two peaks are significant deviations from the general behavior of the observed curve. The distance between the peaks is close to one year. This is an indication

about a possible yearly period, i.e. for some seasonal variations. We shall return to this point in section 6.

The attempts to find more details in the time variations of b were not successful. Moreover, the introduction of a time variable b has not seriously improved the analysis results.

5. Station Strasbourg

In Strasbourg two instruments were successively installed: T005 from 1987 to 1996 and CT26 later on (Table 1). Figures 12 and 13 are similar to the Figures 2 and 3 or Figures 8 and 9 in the previous sections 3 and 4. In figure 12 the noise level is very similar for both instruments and is mainly due to the atmospheric pressure effects. In Figure 13 we have again a large reduction of the level of the residuals, especially for CT26, and a corresponding considerable decrease of the AIC value. Nevertheless, in Figure 13 a considerable number of values are exceeding the threshold level for T005. It is obvious that the records of the new CT instrument are of much better quality.

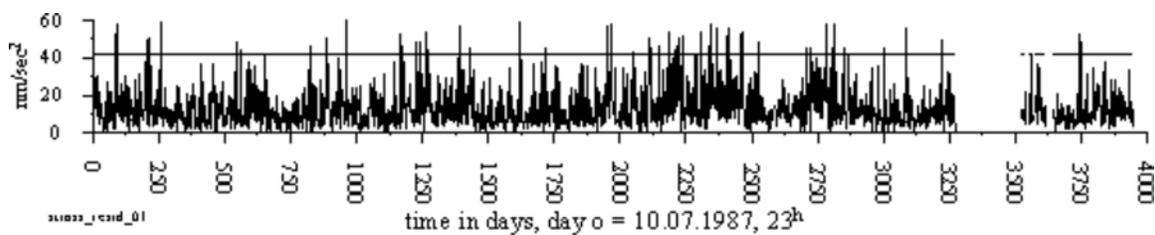


Figure 12. Strasbourg, residuals (modulus) in the D-frequency domain; the effect of the air-pressure is ignored, i.e. no cross-regression model; AIC = 145,247.

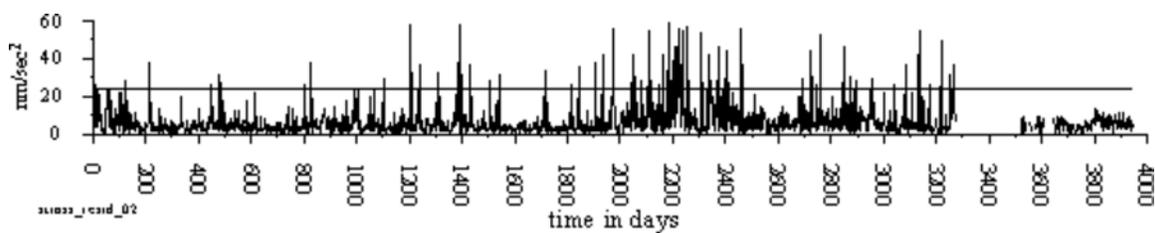


Figure 13. Strasbourg, residuals (modulus) in the D-frequency domain; the effect of the air-pressure is taken into account; a cross-regression model is applied with a constant coefficient b ; AIC= 124,076.

The first attempt to study the time variations of b is shown in Figure 14. The data have been partitioned in two segments corresponding to the different instruments. The results are confirmed in Ducarme & al., 2002 (Table 1), where the authors got using ETERNA software (Wenzel, 1996):

For T005 (3,272days) $b = -3.128 \pm 0.010 \text{ nm.s}^{-2}/\text{hPa}$

For CT26 (817days) $b = -3.394 \pm 0.007 \text{ nm.s}^{-2}/\text{hPa}$

A premature conclusion is that we have different admittances for the two instruments. Such a conclusion could lead us to state that an important part of the air-pressure effect is instrumental.

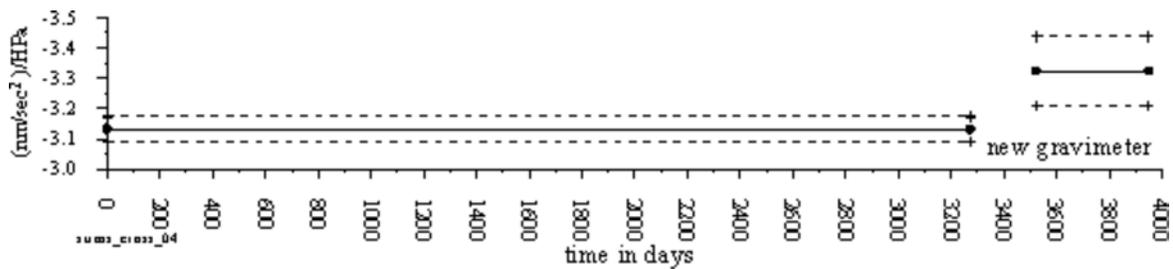


Figure 14. Strasbourg, time variations of b through partition of the data in two segments defined by a gap coinciding with the replacement of the gravimeters; AIC = 124 065.

Figure 15 shows an attempt to study in more details the time variations of b . According to the AIC value this case is more reliable and the premature conclusion made above is no more so convincing. Now we can see several relatively low values, in particular at the very beginning, as well as at two points in the interval (1950, 2250 days).

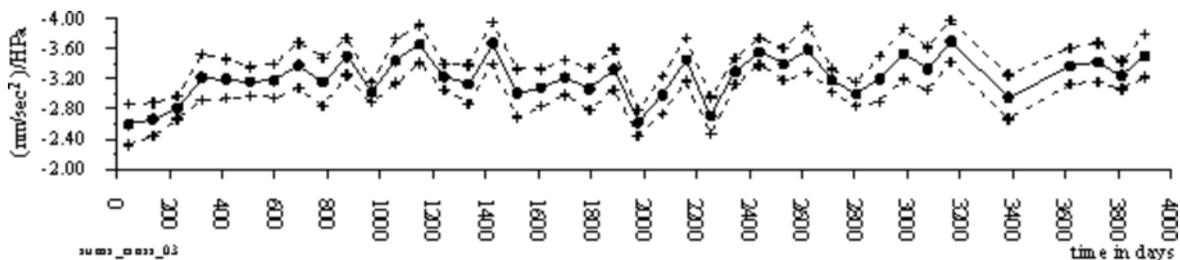


Figure 15. Strasbourg, time variations of b through partition of the data in 30 equal segments; AIC = 123,872.

The next Figures 16 and 17 are samples showing more details in these areas. We have namely four segments of length between 10 and 24 days in which we have practically $b = 0$. The confidential intervals of b are not shown, but actually in all these cases of low b they cover the zero. The further decrease of AIC is in support of this result. In both Figures the AIC has one and the same value because the graphics are obtained through one and the same analysis.

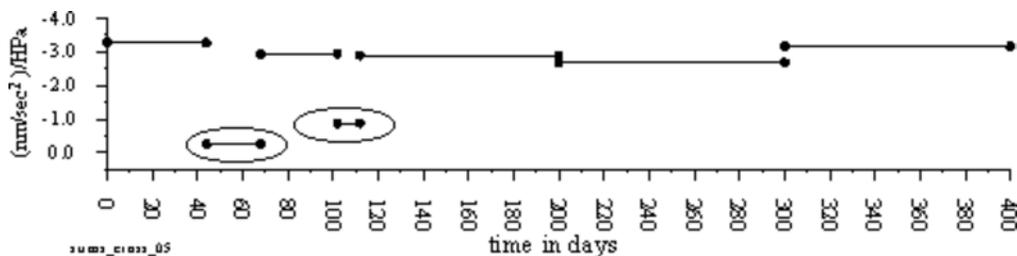


Figure 16. Strasbourg, location of two segments, (44, 68 days) and (102, 112 days), with values of b , which do not differ significantly from the zero; AIC = 123,345.

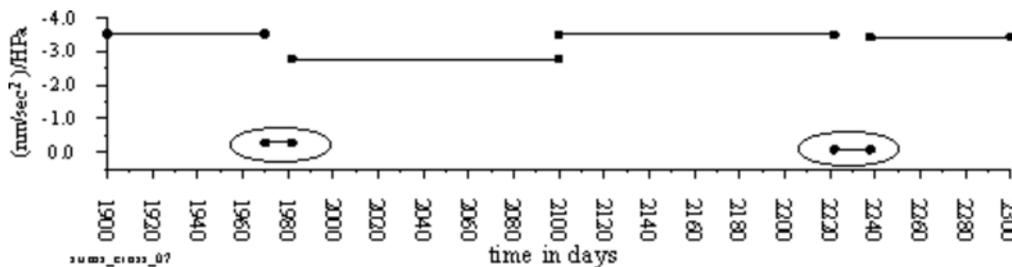


Figure 17. Strasbourg, location of two segments, (1970, 1982 days) and (2222, 2238 days), with values of b which do not differ significantly from the zero; AIC = 123,345.

In (Hinderer et al., 2002) the authors are in favor of the use of interpolated data, as well as of “repaired data”. Due to this we can suppose that the segments with zero admittance are actually interpolated data. Normally the air pressure effect should be artificially introduced in the interpolated sections, using its nominal admittance. Perhaps it was not done properly. We can thus suppose that there are other parts in this series with interpolated data without introduction of the air pressure correction. The result is the introduction of additional noise in the residues of Figure 13, which exceed the threshold level.

It is too difficult to find all places with interpolations and repairs. We have thus simply applied the option of VAV for automatic elimination of the doubtful data, which are of course not only interpolated data. Table 3 shows the results of this procedure, applied in 4 iterations. Although this option resulted in a massive elimination of data, till 23.6% of all data as well as it introduced a huge number of gaps, we have got a fair diminution of the associated RMS errors.

The cases of ETERNA and VAV without any elimination, i.e. in the 0-iteration, are practically identical. However, when we apply VAV possibility to check and eliminate the doubtful data, which is also based on the VAV capacity to deal with a great number of gaps, we get considerable deviations from ETERNA. The precision for the D tides is raised more than twice and, for the SD tides, nearly twice.

Table 3. Station Strasbourg (T005 and CT26): analysis by ETERNA on the whole series and by the VAV program in an iteration procedure, eliminating data with too big residuals.

Software	Nr of iteration	Elim. Data	Amplitude d factor (diurnal tides)		
			Q1 m.s.d.	O1 m.s.d.	K1 m.s.d.
ETERNA		0%	1.14598 ±.00073	1.14733 ±.00014	1.13540 ±.00010
VAV	0	0%	1.14571 ±.00069	1.14726 ±.00013	1.13541 ±.00009
VAV	1	7.8%	1.14577 ±.00041	1.14748 ±.00008	1.13554 ±.00005
VAV	2	15.0%	1.14572 ±.00032	1.14757 ±.00006	1.13569 ±.00004
VAV	3	20.3%	1.14578 ±.00030	1.14756 ±.00006	1.13569 ±.00004
VAV	4	23.6%	1.14586 ±.00029	1.14756 ±.00005	1.13573 ±.00004

Software	Nr of iteration	Elim. Data	Amplitude d factor (semidiurnal tides)		
			N2 m.s.d.	M2 m.s.d.	S2 m.s.d.
ETERNA		0%	1.17173 ±.00040	1.18520 ±.00008	1.18784 ±.00017
VAV	0	0%	1.17173 ±.00039	1.18518 ±.00007	1.18784 ±.00015

VAV	1	7.8%	1.17231 ±.00026	1.18537 ±.00005	1.18795 ±.00010
VAV	2	15.0%	1.17211 ±.00023	1.18547 ±.00004	1.18787 ±.00009
VAV	3	20.3%	1.17214 ±.00022	1.18547 ±.00004	1.18782 ±.00009
VAV	4	23.6%	1.17220 ±.00021	1.18548 ±.00004	1.18786 ±.00009

It is interesting to see how the iterations affect the picture on Figure 14 and the possible conclusion that the air-pressure admittance is significantly different for the different instrument.

In Figure 18 the iteration 0 is a reproduction of Figure 14. After iteration 1, with a moderate quantity of eliminated data, we have still a difference between the two segments, i.e. the two gravimeters. Nevertheless, the difference is smaller and the confidence intervals are overlapping, although they are narrower than in Figure 14. Hence the difference becomes statistically not significant. We can give a sigh of relief, because the instruments are not responsible for the difference and they cannot be accused to have a different admittance. The other iterations in Figure 18 follow the tendency to decrease the difference, so that after iteration 4 we get practically identical admittance for the two segments.

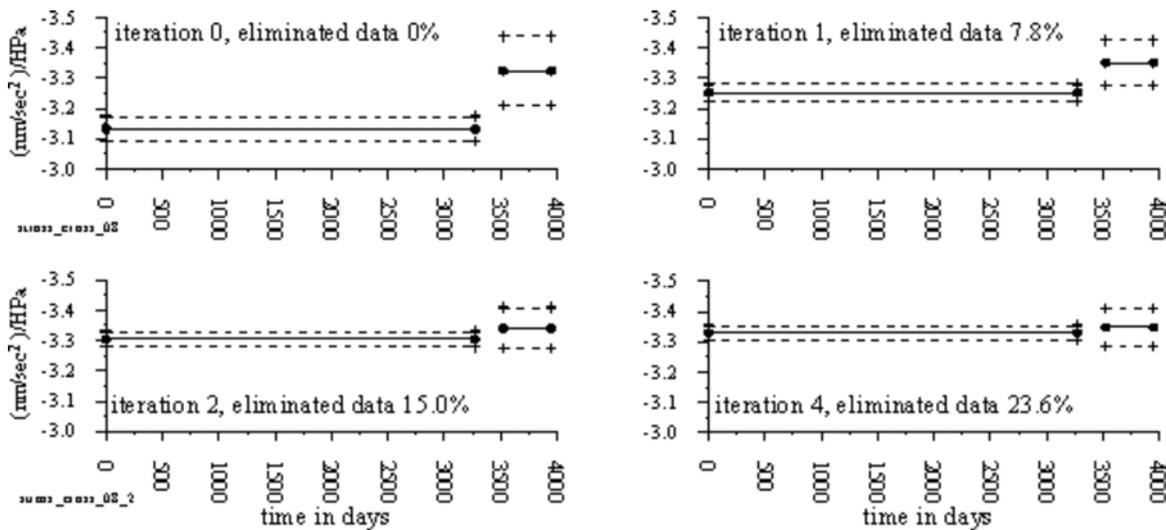


Figure 18. Strasbourg, time variations of b through partition of the data in two segments with elimination of data in an iteration procedure.

6. Seasonal variations of the air-pressure admittance

As said in section 4, in the series of Boulder, there are some indications of a seasonal variation of the air pressure admittance b with a yearly period. As far as such a supposition is correct, the time variations we can see in Figures 10 and 11 may be considered as a useful non-tidal signal. To check this hypothesis we have used a specific way. The data have been split into segments, which are defined as sets of data, instead of segments, which are data intervals. More concretely, in Figure 19, segment 1 includes all data in all years during the month of January, segment 2 – all data during February, and so on. Then the global analysis provides us the b coefficients for every segment. It is possible to say that we partition the data into seasons and get b estimated for every season – in Figure 19 for every month.

The curves in Figure 19, in particular the confidence intervals, indicate the reality of a seasonal dependence, characterized by an increase of the b values during the warmer seasons.

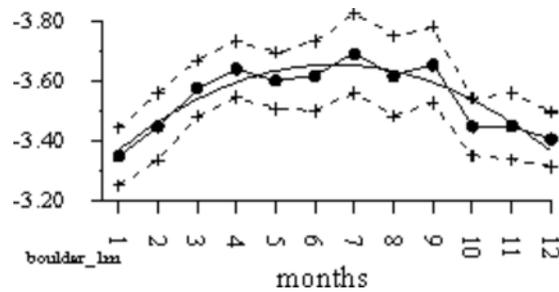


Figure 19. Seasonal variations of the cross-regression coefficient b with confidential interval, a season being chosen as a month.

Other variants of choosing the seasons are shown in Figure 20. Both cases **A** and **B** of this Figure confirm the inference for higher b during the warmer seasons.

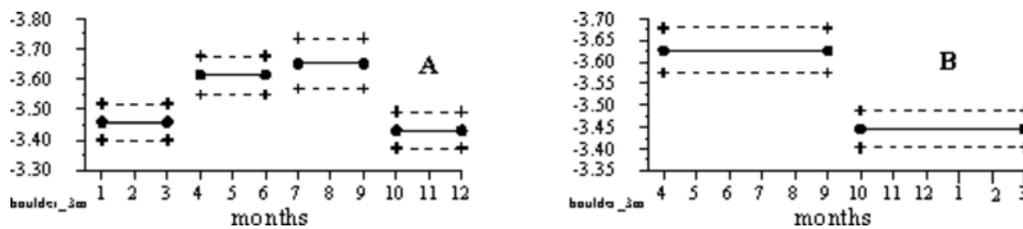


Figure 20. Seasonal variations of the cross-regression coefficient b ; in case **A** we have 4 seasons, every season involving 3 months and in case **B** we have only two seasons, every one of them involving 6 months.

We confirm here the results obtained by T.van Dam & O.Francis (1998) who gave a meteorological explanation of this phenomenon.

In other series of data we have observed only very slight similarity to the picture in Figure 19. Obviously, the problem is very sophisticated as it depends essentially of the regional meteorological conditions. It should be investigated in more details. One of the necessary conditions for such investigation, is a careful check of the data and the elimination of all artificially introduced anomalies.

7. Conclusions

The initial purpose of this paper was to check the existing conception about the air pressure admittance and, eventually, to encourage the research of more sophisticated models. In this sense, the only result, which may be considered as moderately successful, is the observation of time variations in Boulder.

In the other examples, considered here, our computing technique turned out to be helpful only for finding some anomalies in the data. Thus the first conclusion is a trivial one: it is impossible or too difficult to find useful non-tidal signals, when the data still contain gross errors, i.e. non-useful non-tidal signals. It is traditionally the goal of the preprocessing of the tidal data to "repair" the data by eliminating the spikes and jumps and to fill small gaps in the series. Ideally abnormal data should always be eliminated, but the different tidal analysis methods had and still have different requirements on the preprocessed data, even if the term of "repaired data" may sound to a specialist on data processing no better than "false money" to a cashier. In any case another conclusion of our investigation is that non-careful interpolations can generate very bad, non-stationary noise in data series of very good quality.

The fear of gaps is linked to the classical Fourier methods as they have been applied in the early times, the epoch without computers or with very slow computers. When such methods are applied on evenly data without gaps we get fine results and the computational work is very simple. If there are gaps, the computations become considerably more complicated. If the gaps are simply ignored and replaced by zeros, we get spikes and all possible deformations of the results.

The situation is completely different when the method of the least squares (MLSQ) is used. MLSQ is a more general method than the Fourier methods. There is not any condition of evenly spaced data. We have only to correctly create the observation equations about the data at the time moments at which they exist. The equations take into account all interrelations or interactions between the unknown parameters and the functions in which the parameters take part. Then the solution of the equations provides the estimates of the parameters without any spikes.

Of course, it is better when the data are without gaps, by the simple reason that higher quantity of data provides a higher precision. Nevertheless, it is a bad illusion to believe that, if a gap is fulfilled by artificially created data, this operation will improve the accuracy. In the best case, it is possible to get an apparent reduction of the m.s.d. Moreover, if the model used for interpolation is different from the real tidal factors of the series, interpolation will bias the results. Actually, the interpolated and repaired data are often source of additional noise. It is much more useful to work with longer series including gaps, to improve the resolution, than with shorter series without gaps. It is better to use a series of 12 month with a two months gap inside, than a series of 10 months without gaps.

Actually, we think that the interpolation is imposed by purely practical reasons. The methods of analysis like ETERNA (Wenzel, 1996), based on the general scheme of Chojnicki (1972), use high-pass filters like the filter of Pertsev. Such filters have a low signal-to-noise ratio, of the order of 1. Due to this they need a moving filtration, which replaces one original data by one filtered data (nearly). In the same time the moving filtration cannot support a great number of gaps, because they may produce losses of a great number of data. To obtain filters with better signal to noise and cut-off characteristics it is necessary to increase the length of application and thus to increase the loss of data at each gap. One cannot avoid these sharp filters when decimating the data from the original sampling rate to minutes or hourly data. Moreover the direct analysis without filtering, which is used for the determination of the long period tides including the so called "pole" tides, requires to model the drift in each data section. A large number of gaps or jumps will ends up with an unrealistic number of unknowns.

The methods like VAV use narrow band-pass filters, which have considerably higher signal-to-noise ratio. For example, if the time window is 48 hours, these filters provide, for each tidal family, a pair of filtered numbers, everyone with a signal-to-noise ratio very close to $48/2=24$. Due to this we can apply the filters without overlapping, with all gaps remaining between the filtered intervals, with a very small quantity of lost data. Some people consider such a filtration as a decimation of the data with a step equal to the time window. Actually, this operation is a transformation of the data from the time domain in a time/frequency domain, which keeps the whole useful information.

The problem of spikes and tares is more delicate. If non-harmonic perturbations are left in the original sampling of the data, they will produce biased data after decimation to lower sampling rates, minutes or hours. In tidal analysis methods such as ETERNA, which can be applied to any sampling rate without decimation, spikes and tares have nevertheless to be "repaired" as their simple elimination will create a large numbers of small gaps. Other classical analysis methods such as VEN66 (Venedikov, 1966) or NSV98 (Venedikov & al., 1997) still require uninterrupted data sets of let's say 48h and it was the usual practice to interpolate missing data up to a few hours to complete blocks and to smooth out small perturbations. Only VAV accepts unevenly spaced data and will no more require interpolations or data smoothing, as abnormal data can be directly eliminated in the original sampling.

References

Chojnicki, T., 1972. Détermination des paramètres de marée par la compensation des observations au moyen de la méthode des moindres carrés, *Institute of Geophysics, Polish Academy of Sciences, Marées Terrestres*, **55**, 48-80.

Crossley, D., 2000. Report on the status of GGP, *Cahiers Centre Européen Geodyn. Seism.*, **17**, 1-7.

Ducarme, B., Vandercoilden, L., 2000. First results of the GGP data Bank at ICET, *Cahiers Centre Européen Geodyn. Seism.*, **17**, 117-124; *Bull. d'Inform. Marées Terrestres*, **132**, 10291-10298.

Ducarme B., Sun, H.P., Xu, J.Q., 2002. New investigation of tidal gravity results from GGP network, *Third WORKSHOP of the Global Geodynamics Project (GGP) on Superconducting Gravimetry, Jena, March 11-14, 2002*, *Bull. Obs. Marées Terrestres*, **135**,...

Hinderer, J., Rosat, S., Crossley, D., 2002. Influence of different processing methods on the retrieval of gravity signals from GGP data, *Third WORKSHOP of the Global Geodynamics Project (GGP) on Superconducting Gravimetry, Jena, March 11-14, 2002*, *Bull. Obs. Marées Terrestres*, **135**,...

van Dam, T., Francis, O., 1998. Two years of continuous measurements of tidal and non-tidal variations of gravity in Boulder, Colorado, *Geophys. Research Letters*, **25**, 3, 393-396.

Venedikov, A.P., 1966, Une méthode d'analyse des marées terrestres à partir d'enregistrements de longueurs arbitraires, *Observatoire Royal de Belgique, Série Géophysique*, **71**, 463-485.

Venedikov, A.P., Vieira, R., de Toro, C., Arnosó, J., 1997. A new program developed in Madrid for tidal data processing, *Bull. Obs. Marées Terrestres*, **126**, 9969-9704.

Venedikov, A.P., Arnosó, J., Vieira, R., 2001, Program VAV/2000 for tidal analysis of unevenly spaced data with irregular drift and colored noise, *Journal Geodetic Society of Japan*, **47**, 1, 281-286.

Wenzel, H.G., 1996, The nanogal software: data processing package ETERNA 3.3, *Bull. Inf. Marées Terrestres*, **124**, 9425-9439.