IMPORTANCE OF AUTOCORRELATION FOR PARAMETER ESTIMATION IN REGRESSION MODELS

Heiner Kuhlmann

Institute for Application of Geodesy to Engineering, University of Stuttgart
Geschwister-Scholl-Str. 24D, 70174 Stuttgart, Germany
E-Mail: Heiner.Kuhlmann@uni-stuttgart.de

Abstract

In deformation analysis the functional relationship between the acting forces and the resulting deformations should be established. If time depending observations are given, a regression could be used as functional model. In case of stochastic model uncorrelated observations with identical variance are assumed. Due to the high sampling rate a small time difference arises between two observations. Thus the assumed stochastic model is not suitable. The calculation has to be effected by means of autocorrelated obervations.

In this paper the influences of autocorrelation on the estimated parameters, the variances and degree of freedom are shown. The autocorrelation can be described via Gauss-Markov-process. This method is demonstrated on the basis of a deformation monitoring of a jointless monolithic bridge.

1. Reason for the deformation measurement

In the year 1999 in Stuttgart, Germany a new bridge was build along a highway crossing a small valley „Nesenbachtal“. It is a jointless monolithic bridge. The abutments of the bridge are connectet directly with a tunnel on each side. Outlines of the bridge are shown in figure 1.

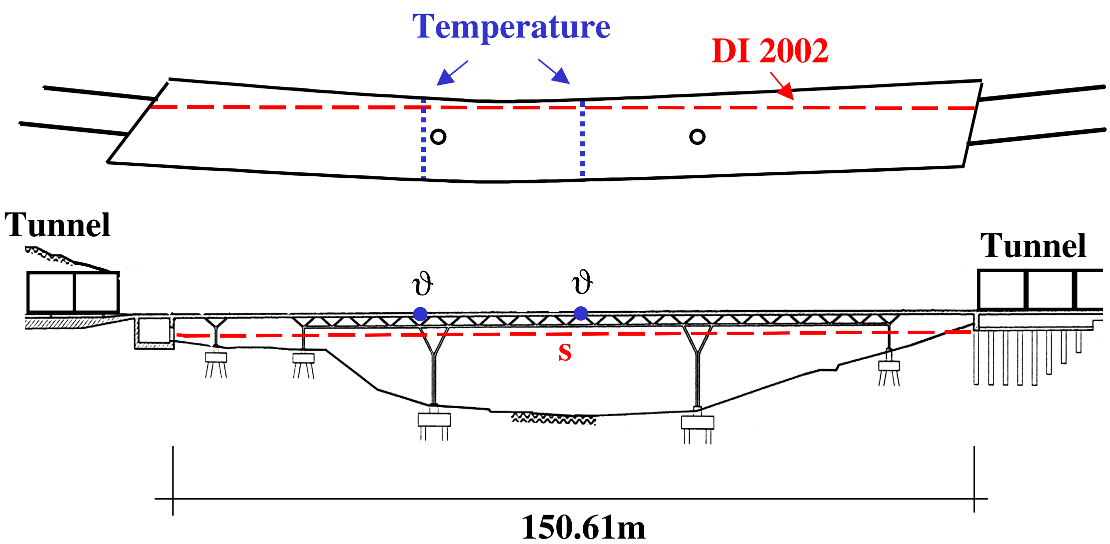

Fig. 1: Bridge crossing „Nesenbachtal“

Monolithic bridges have two basic advantages contrary to conventional bridges: For bridges with joints the maintenance for the construction joints are enormous due to corrosion and filling of the joint with sand, etc. Using monolthic bridges there are no articulations at the joints, thus a continuous beam lead to advantages for the static behaviour, and thus the cross-section of the superstructure gets smaller with effects to building costs (ENGELSMANN et. al. 1999).

A problem of this construction method is that changes in length occur due to the changes of the temperature, which cannot move into expansion joints. The results are higher tensions in material and other deformations. At this building no transmission of the deformations were allowed due to the temperature in the tunnels. The result of a static calculation of the bridge shows no movements of the abutments. Due to the circular shape of the bridge the expansion leads to a radial displacement.

Aim of the deformation measurement was to prove that no temperature deformation is transmitted in the tunnels. Thus the abutment must not move.

2. Measurements and deformation model

The measurements were carried out in May 1999 during a lecture course. The temperatures of the concrete ϑ were measured by the mean value of 10 sensors, mounted in two cross-sections inside the superstructure. The changes of the distance between the two abutments were measured by electronic distance meter Leica DI2002. The sampling rate was $\Delta t = 10$ min and the time of measurement approximately 2 days. That means $n = 268$ measured values. Besides this some points of the building were observed by means of a motorized self-tracking tacheometer. These measurements are not subject of this paper. The time series of the distance and the temperature are shown in figure 2.

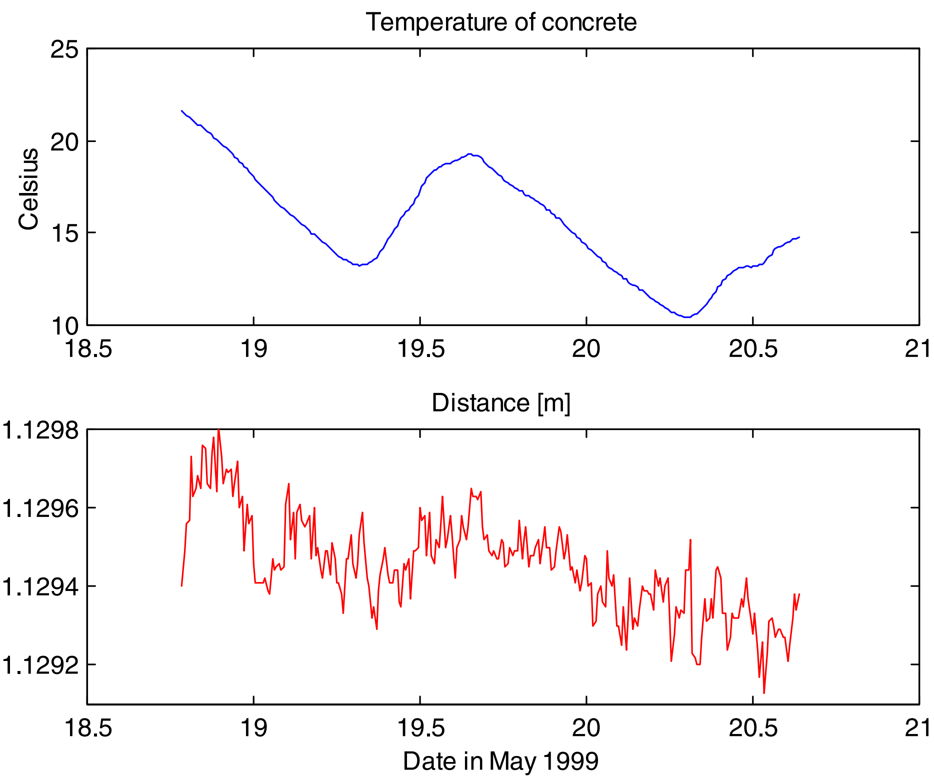

Fig. 2: Measurements of temperature and distance

In this distance to the target an accuracy of the electronic distance meter $\sigma_s = 1$ mm is given by the manufacturer. It is empirically known that the accuracy for distance differences ist about $\sigma_{\Delta s} = 0.2\text{-}0.3$ mm.

The investigation should serve the purpose if the changes in temperature have significant influence onto the changing of the distance. Thus the deformation model should fulfil the functional relationship between temperature of concrete and measured distances. In the model one has to take into consideration input signal and time depending deformations. That means a dynamic deformation model has to be chosen [WELSCH et. al. 2000]. In this case the dynamic model can be reduced to static deformation model because the concrete temperatue is acting immediately, without time delay on the deformations [WELSCH et. al. 2000]. This is proved by calculating a cross-covariance-function between the temperature and the distance. This function has a maximum at the time delay of $\tau = 0$ min.

The approached model is

$$\Delta s = b \cdot \Delta \vartheta \,. \qquad (1)$$

The parameter b includes the length of the bridge multiplied by the effective extension coefficient due to the temperature, which includes the stiffness of the construction.

3. Numerical solution of the deformation model

3.1. Regression with not-correlated observations

Aim of the investigation is calculating parameter b and its standard deviation. If b is significant to zero, the changes of temperature have influence on the movement of the abutment.

The calculation is done by a regression approach, directly derived from (1):

$$\boldsymbol{\Delta s} = \boldsymbol{\Delta \vartheta} \cdot b + \boldsymbol{\varepsilon} \text{ resp. } \boldsymbol{l} = \boldsymbol{A} \cdot \hat{\boldsymbol{x}} - \boldsymbol{v} \qquad (2)$$

In this model the vector of random errors $\boldsymbol{\varepsilon}$ includes the deviations of distance measurement. The temperature has in this model, this approach is often used in regression. In case of stochastic parts for the temperature, (2) a Gauß-Helmert-model will result. The essential message of this paper does not change due to this modification.

The observation vector $\boldsymbol{\Delta s}$ is calculated from the measured distances subtracted by the mean value. The same procedure is applied to the temperature, and leads to vector $\boldsymbol{\Delta \vartheta}$ (centering of the values)

For stochastic model

$$\boldsymbol{\Sigma}_{\Delta s \Delta s} = E\left\{\boldsymbol{\varepsilon} \cdot \boldsymbol{\varepsilon}^T\right\} \text{ with } \boldsymbol{Q}_{\Delta s \Delta s} = \boldsymbol{I} \qquad (3)$$

is used at first: all observation have the same accuracy and are not correlated, thus cofactor matrix is identity matrix $\boldsymbol{I}$. σ_0^2 complies the variance of the difference distance measurement.

The functional and stochastic model given in (2) and (3) lead to the well-known regression algorithm. It is identical to the solution of Gauß-Markov-model (see for example [PELZER, 1985] and [HÖPCKE, 1980]). The essential results are shown in tab. 1:

value	equation		numerical value
matrix of weight	$\boldsymbol{P} = \boldsymbol{Q}_{\Delta s \Delta s}^{-1}$	(4)	$\boldsymbol{P} = \boldsymbol{I}$
matrix of normal equations	$\boldsymbol{N} = \boldsymbol{A}^T \cdot \boldsymbol{P} \cdot \boldsymbol{A}$	(5)	$\boldsymbol{N} = 2216.2$ °C^2
constant term	$\boldsymbol{n} = \boldsymbol{A}^T \cdot \boldsymbol{P} \cdot \boldsymbol{l}$	(6)	$\boldsymbol{n} = 0.0709$ °C·m
cofaktor matrix of parameter	$\boldsymbol{Q}_{\hat{x}\hat{x}} = \boldsymbol{N}^{-1}$	(7)	$\boldsymbol{Q}_{\hat{b}\hat{b}} = 4.5123 \cdot 10^{-4}$ °C^{-2}
parameter	$\hat{\boldsymbol{x}} = \boldsymbol{Q}_{\hat{x}\hat{x}} \cdot \boldsymbol{n}$	(8)	$\hat{b} = 0.032$ mm/°C
degree of freedom	$f = n - u$	(9)	$f = 266$
residuals	$\boldsymbol{v} = \boldsymbol{A} \cdot \hat{\boldsymbol{x}} - \boldsymbol{l}$	(10)	
empiric variance of unit of weight	$s_0^2 = \frac{\boldsymbol{l}^T \cdot \boldsymbol{P} \cdot \boldsymbol{l} - \hat{\boldsymbol{x}} \cdot \boldsymbol{n}}{f} = \frac{\boldsymbol{v}^T \cdot \boldsymbol{P} \cdot \boldsymbol{v}}{f}$	(11)	$s_0 = 0.08$ mm
covariance matrix of parameter	$\boldsymbol{S}_{\hat{x}\hat{x}} = s_0^2 \cdot \boldsymbol{Q}_{\hat{x}\hat{x}}$	(12)	$s_{\hat{b}} = 0.0018$ mm/°C
multiple value of definiteness	$B = \frac{\boldsymbol{n} \cdot \hat{\boldsymbol{x}}}{\boldsymbol{l}^T \cdot \boldsymbol{P} \cdot \boldsymbol{l}} = \frac{\boldsymbol{l}^T \cdot \boldsymbol{P} \cdot \boldsymbol{l} - \boldsymbol{v}^T \cdot \boldsymbol{P} \cdot \boldsymbol{v}}{\boldsymbol{l}^T \cdot \boldsymbol{P} \cdot \boldsymbol{l}}$	(13)	$B = 0.55$

Tab. 1: Results of the regression with $\boldsymbol{Q}_{\Delta s \Delta s} = \boldsymbol{I}$

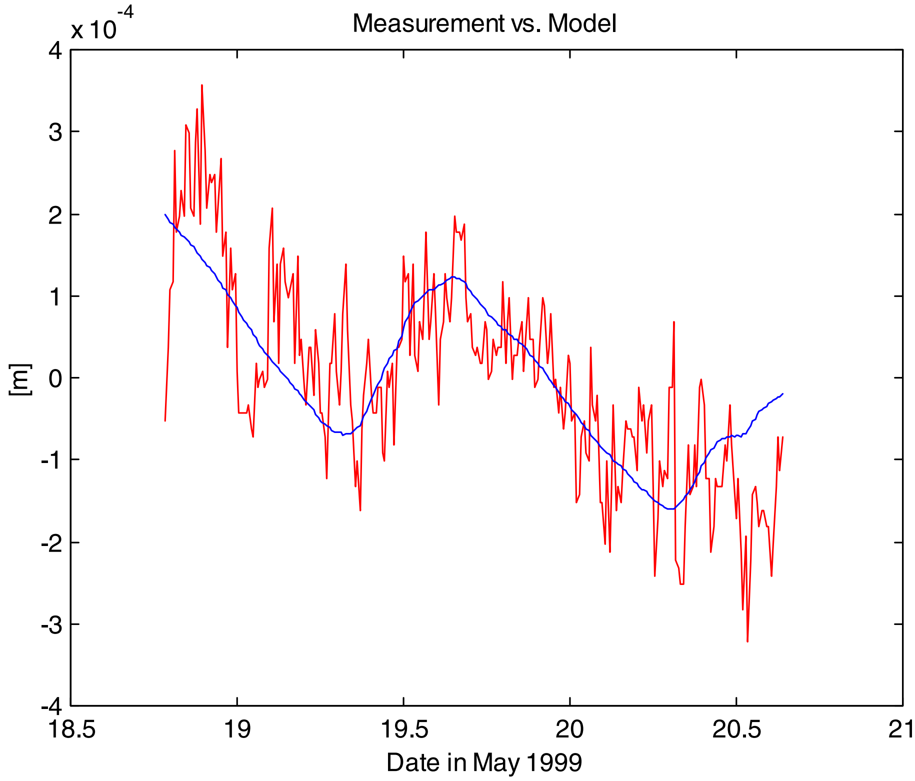

Fig. 3: measured (red) vs. calculated (blue) deformation with $\boldsymbol{Q}_{\Delta s \Delta s} = \boldsymbol{I}$

The multiple value of definiteness B is the relative decrease of the sum of the square of deviations. It describes the representation of the maesured distances by the temperature [HÖPCKE, 1980]. B and the parameter $\hat{b}$ can be checked to significane with statistical tests [WELSCH et. al., 2000]. An error propability of $\alpha = 5\%$ leads to the results shown in tab. 2.

value	test value	quantile	significant?
$\hat{b}$	$T_{\hat{b}} = \hat{b} / s_{\hat{b}} = 17.8$ (14)	$t_{f,1-\alpha/2} = 1.97$	yes
B	$T_B = f \cdot \frac{B}{1-B} = 325.1$ (15)	$F_{1,f,1-\alpha} = 3.841$	yes

Tab. 2: Test values with $\boldsymbol{Q}_{\Delta s \Delta s} = \boldsymbol{I}$

Regarding the graphic representation of the results shown in figure 3, one can't understand this clear test decision. The trend is represented by the model, but there are considerable deviations especially at the first few measurements. Due to this fact a narrow test decision is expected. Additionally the empiric standard deviation of unit of weight $s_0 = 0.08$ mm is very optimistic in comparison with the expected accuracy of the EDM $\sigma_0 = 0.2\text{-}0.3$ mm. The reason for this too optimistic measures of accuracy is a wrong stochastic model (3).

3.2. Description of auto-correlation with Gauß-Markov-process

The empric auto-correlation function $\hat{K}(k)$ can be calculated from the residuals $\boldsymbol{v}$ (10) by

$$\hat{C}(k) = \frac{1}{n-k-1} \sum_{j=1}^{n-k} v_j \cdot v_{j+k} \quad , k = 0,1,2,\ldots,n/10, \qquad \hat{K}(k) = \frac{\hat{C}(k)}{\hat{C}(0)} \tag{16}$$

[WELSCH et.el., 2000]. The result is shown in figure 4.

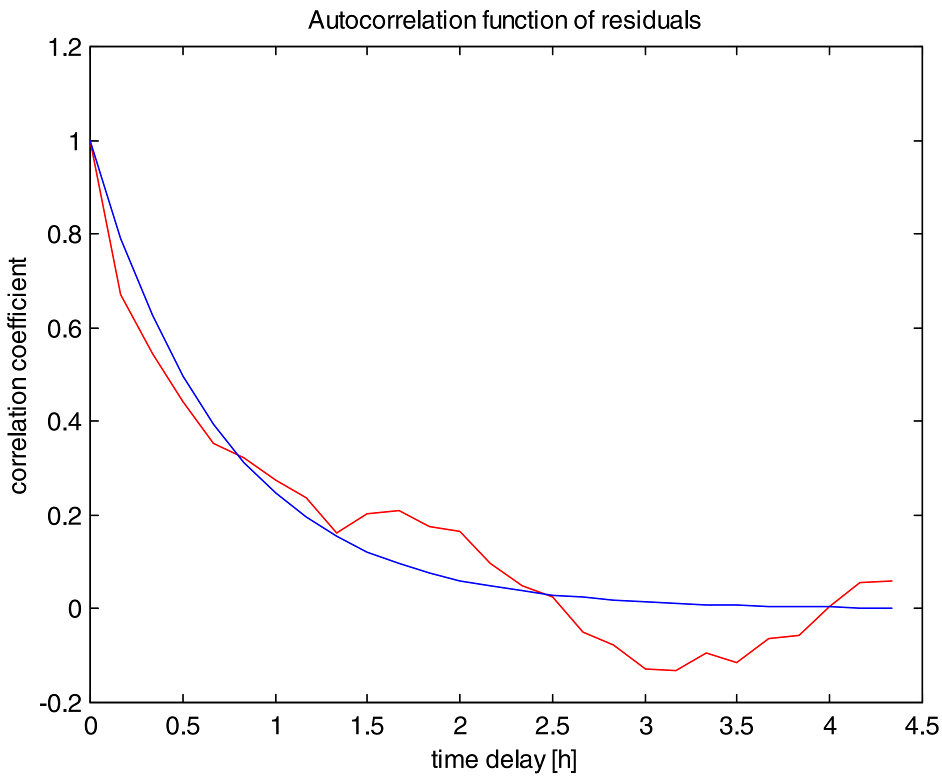

Fig. 4: Auto-correlation function of residuals

The correlation is obviously decreasing slowly. This is a so called Gauß-Markov-process. The continuous auto-correlation function of this process can be described via

$$K(\tau) = e^{-\alpha|\tau|}, \quad \alpha > 0\,. \tag{17}$$

The function depends on one parameter α. The estimated value $\hat{\alpha} = 0.234$ is calculated by the empiric auto-correlation function, figure 4 shows the result. As shown, there is high correlation between neighbored values of distance maesurement in time domain, due to similar parts of systematic errors. An example for this is an incorrect first velocity correction due to not representative temperature of the air which puts similar effect on contemporary distances.

3.3. Modified regression approach with correlated observations

The correlation between contemporarly neighbored measurements proven in the last chapter has to be taken into consideration in two different ways.

At first one has to consider the proven auto-correlation of the distance measurement in the stochastic model (3) by adequate values in the cofactor matrix. E.g. the first secondary diagonal contains the correlation coefficient between immediately neighbored values [WELSCH et. al., 2000]. The result is shown in (18).

$$Q_{\Delta s \Delta s} = \begin{vmatrix} 1 & K(1) & K(2) & \cdots & K(n-1) \\ K(1) & 1 & K(1) & \ddots & \vdots \\ K(2) & K(1) & 1 & \ddots & K(2) \\ \vdots & \ddots & \ddots & \ddots & K(1) \\ K(n-1) & \cdots & K(2) & K(1) & 1 \end{vmatrix} \tag{18}$$

Besides this the auto-correlation has an influence on the degree of freedom. A new measurement is not an independent observation because consecutive values are highly correlated. One can predict a new observation because it is „almost identical“ with the last one. Thus the degree of freedom is not $f = n - u$, but only n_{eff} independend measurements [TAUBENHEIM, 1969]. According to [BARTELS, 1935] this value is called „effective number of measurments“. It can be calculated by

$$n_{eff} = \frac{n}{1 + 2 \cdot \sum_{k=1}^{n-1} \frac{n-k}{n} K(k)} . \tag{19}$$

Regarding the consideration first mentioned with a modified cofactor matrix, the algorithm (4) to (13) leads to values shown in tab. 3. For comparison the last row contains once again the values calculated without auto-correlation.

	n_{eff}	b	s_0	$s_{\hat{b}}$	B
with auto-correlation	32	0.029 mm/°C	0.10 mm	0.0060 mm/°C	0.08
without auto-correlation	266	0.032 mm/°C	0.08 mm	0.0018 mm/°C	0.55

Tab. 3: Comparison of the results of adjustment

Comparing the results shown in tab. 3 there are the following statements:

- The result of the multiple value of definiteness did change substantially due to the auto-correlation. For the decision, if B is significant or not, see below.
- The same is valid for the standard deviation of the parameter $s_{\hat{b}}$. There is an increase by factor 3.
- The empiric standard deviation of unit of weight did change only to a small extend. It is still optimistic in relation to the value expected. Obviously in this case the Leica DI2002 has higher accuracy for difference measurements then expected.
- Remarkable is the changing of the parameter $\hat{b}$ by approximate 10%.

The degree of freedom changed to $n_{eff} = 32$. This fact has an influence on the quantiles used for the statistic tests (14) and (15).

value	test value	quantile	significant?
$\hat{b}$	$T_{\hat{b}} = \hat{b} / s_{\hat{b}} = 4.83$ (14)	$t_{n_{eff},1-\alpha/2} = 2.04$	yes
B	$T_B = f \cdot \frac{B}{1-B} = 2.8$ (15)	$F_{1,n_{eff},1-\alpha} = 4.15$	no

Tab. 4: Test values with $\boldsymbol{Q}_{\Delta s \Delta s} \neq \boldsymbol{I}$

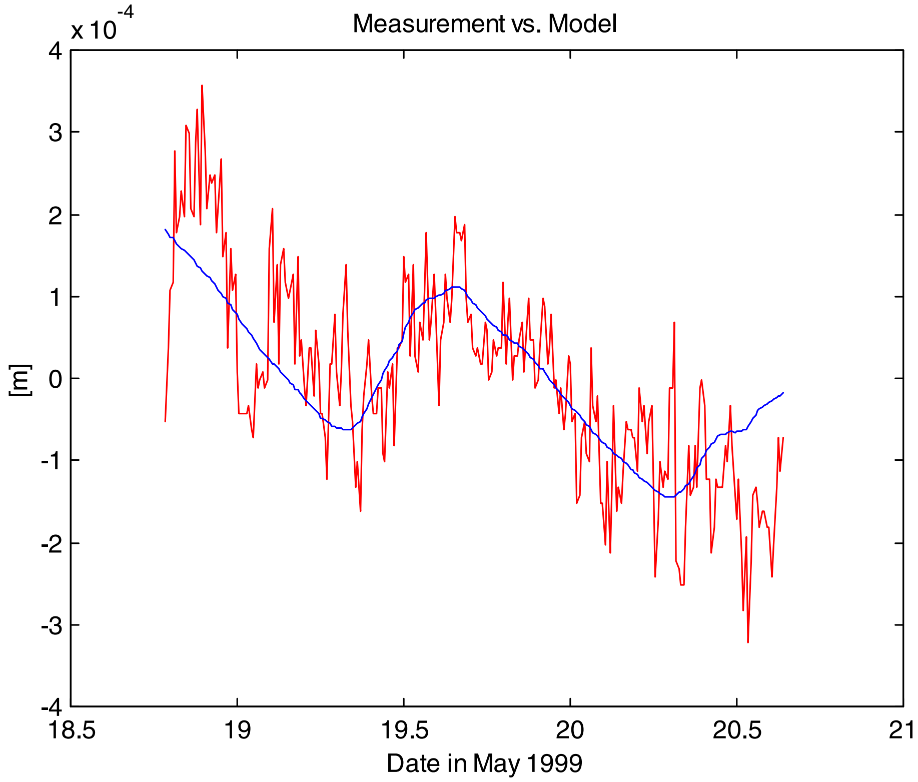

Fig. 5: measured (red) vs. calculated (blue) deformation with $\boldsymbol{Q}_{\Delta s \Delta s} \neq \boldsymbol{I}$

As seen in tab. 4, the paramter $\hat{b}$ is significant once again, but not the multiple value of definiteness B. These phenomena could be explained by figure 5. It shows the measured distance differences at the calculated model: the general course of the distance differences could be explained by the deformation model (2) and the estimated parameter $\hat{b}$. Thus the distance differences are a result of temperature changing of the building. But there are still considerable deviations in relation to the general course which are not modeled in (2). Thus B is not significant. The order of magnitude of the deviation is only a tenth of a millimeter and has to be compared with the accuracy of the sensor. Thus no other result is expected.

4. Conclusion

A movement of the abutment due to temperature changing is proved by the realized measurements. But an assumption of a temperature changing of $\Delta\vartheta = 50°C$ leads only to a distance difference of Δs= 1.5 mm. For the static behaviour of the abutment this magnitude is not critical.

An auto-correlation between different realizations of the distance could be proved by the residuals. If this influence is neglected in computation, there result too optimistic measures of accuracy. In this case some stochastic parameter will be erroneously assumed to be significant. The estimated parameter $\hat{b}$ did change by 10%.

Especially with regard to the statistic tests of adjustment results this influence is important. With auto-correlation the measures of accuracy will increase while the degree of freedom will decrease.

Finally the calculation of auto-correlation function for a correct stochastic model neccesitates an iteration, as the auto-correlation function can only be calculated by the residuals. These are accessible after estimation of parameters. A calculation of auto-correaltion by raw observations is not suitable, because the observations include the influence of acting forces -in this case the concrete temperature-, which is described in functional model.

5. References

Bartels, J. (1935): Zur Morphologie geophysikalischer Zeitfunktionen. Sitzungsbericht der physikalisch-mathematischen Klasse der Preußischen Akadademie der Wissenschaften, Band 30, S. 504-522.

Engelsmann, S.; Schlaich, J.; Schäfer, K. (1999): Entwerfen und Bemessen von Betonbrücken ohne Fugen und Lager. Beuth Verlag, Berlin.

Höpcke, W. (1980): Fehlerlehre und Ausgleichungsrechnung. Verlag Walter de Gruyter, Berlin, New York.

Pelzer, H. (1985): Geodätische Netze in Landes- und Ingenieurvermessung II. Verlag Konrad Wittwer, Stuttgart.

Taubenheim, J. (1969): Statistische Auswertung geophysikalischer und meteorologischer Daten. Akademische Verlagsgesellschaft Geest und Portig, Leipzig.

Welsch, W.; Heunecke, O.; Kuhlmann, H. (2000): Auswertung geodätischer Überwachungsmessungen. In: Möser et.al.: Handbuch Ingnieurgeodäsie. Wichmann Verlag, Heidelberg.