Using Bayesian Methods for the Parameter Estimation of Deformation Monitoring Networks

Tanir[1,3] E., Felsenstein[2] K. and Yalcinkaya[3] M.

[1]Institute of Geodesy and Geophysics, Vienna University of Technology, Vienna, Austria
[2]Institute of Statistics and Probability Theory, Vienna University of Technology, Vienna, Austria
[3]Department of Geodesy and Photogrammetry Engineering, Karadeniz Technical University, Trabzon, Turkey
etanir@mars.hg.tuwien.ac.at

Abstract

In order to investigate the deformations of an area or an object the geodetic observations are repeated at different time periods and then these observations of each period are adjusted independently. From the coordinate differences between the periods the input parameters of a deformation model are estimated. The decision about the deformation is given by appropriate deformation model using the parameter estimation results from each observation period. So, we have to be sure that we use accurately taken observations (assessing the quality of observations) and also we use appropriate mathematical model for both network adjustment period measurements and for the deformation modelling. All inaccuracies of the model, especially systematic and gross errors in the observations, as well as incorrectly evaluated a priori variances will contaminate the results and lead to apparent deformations. Therefore, it is of prime importance to employ all known methods which can contribute to the development of a realistic model (Caspary 2000).

Within this study, we want to introduce alternative parameter estimation to least-squares estimation for an deformation monitoring network which is constructed for landslides monitoring in Macka in the province of Trabzon in north eastern Turkey. We use Bayes theorem for this new parameter estimation tool. The Bayesian inference allows evaluating the probability of an event by available prior evidences and of the collected observations. Bayes theorem underlines that the observations modify through the likelihood function the prior knowledge of the parameters, thus leading to the posterior density function of the parameters themselves.

1. Introduction

Bayes theorem leads to posterior distribution of unknown parameters given the data. All inferential problems concerning the parameters can be solved by means of posterior distributions. Based on this distribution we will estimate the unknown parameters, establish confidence regions for them and test hypothesis concerning parameters. In this study, it is aimed to apply Bayes-Updating (BU) and Gibbs-Sampling (GS) algorithms in a geodetic parameter estimation problem. In addition, the task is to offer the user in which algorithm is more useful in the aspect of accuracy and the methodology. Concerning this issue, it is also aimed to offer some remarks in using these algorithms. For example, what can be done to improve the accuracy of these algorithms in the aspect of data?

To apply the Bayes-Updating we introduce the prior information for station coordinates. We considered different situations: prior information coming from different epoch observations and non-informative prior. In Gibbs-Sampling algorithm we introduce different starting points w.r.t. point positioning errors. Usually, measurements at different points of a network are of different quality concerning statistical variance. We test the quality of network points by comparing the Bayesian estimates derived from a Bayes-Updating (conjugate updating) algorithm or a Gibbs-sampling procedure. The total variation (L1-distance) of the resulting estimates serves as measure of discrepancy for reasons of robustness. The information contained in a network point is evaluated by taking that point as starting point for Gibbs-sampling. The distances (L1-measure) between

Bayesian updating and Gibbs-sampling indicate how much imprecision the point adds to the estimation procedure.

As well as the influence one single network point we consider the information constrained in different data sets in the same way. The differences of estimation taking data sets in exchanging order (concerning prior data set and actual data set) serves as method to evaluate the information contained in the data set. We applied the method to a concrete network successfully and constructed a concrete spectrum of informational points.

2. Bayesian Parameter Estimation Methods

2.1. Bayes Updating

The basic instrument of Bayesian procedures consists of updating algorithm via Bayes theorem

$$\pi(\theta / D) \ \alpha \ l(\theta, D)\pi(\theta) \tag{1}$$

where $\pi(.)$ denotes the prior distribution and $\pi(./D)$ the posterior distribution of the model parameter. $l(.)$ denotes the likelihood function of the data. For the network data we use a correlated normal vector *x* with the mean vector μ and covariance matrix Σ. In a conjugate prior approach the prior knowledge about parameters is assessed in the same form as the likelihood. In case of correlated normal observations the corresponding prior is a *Normal-Wishart* distribution (Felsenstein 1996). The common prior of μ and the precision matrix $P = \Sigma^{-1}$ is split into a normal part

$$\mu / P \sim N(m, (bP)^{-1}) \tag{2}$$

with $b>0$ and Wishart part

$$P = \Sigma^{-1} \sim Wis(a, \Lambda) \tag{3}$$

The weights $b,a>0$ play the same role as the number of observation in the likelihood. μ is the prior guess for the mean and Λ a prior guess for covariance matrix. For the data of the geometric network it is important to specify the prior covariance according to the geometric structure (euclidian distances on surfaces). Once the prior hyper parameters are specified the posterior can be achieved through the updating of the hyper parameters only. These are

$$m^* = \frac{bm + n\bar{x}}{b+n} \tag{4}$$

$$\Lambda^* = \Lambda + \frac{nb}{b+n}(m - \bar{x})(m - \bar{x})^T + R \tag{5}$$

where *R* is the empirical covariance

$$R = \sum_i x_i x_i^T - n\bar{x}\bar{x}^T \tag{6}$$

which is computed out of data vectors. Note that any observation x_i contains least-squares estimation for coordinates of all points of the network. The modelling through a conjugate family of priors is indeed attractive even in case of little prior information. Prior distributions carrying as little

information as possible can be considered as limits of conjugate priors in the sense

$$\Lambda \to 0 \tag{7}$$

and $b \to 0$. The conjugate property is not restricted to a single distribution. Mixtures of Normal-Wisharts are conjugate as well allowing more flexibility to the model. As weights we choose several values of *b* in our calculations to analyze the sensitivity of the estimates upon the prior distributions (Robert 2001). Bayesian estimates of the different parameters of the network are calculated as posterior means leading to weighted mixtures of the means

$$m_j^* = \frac{bm_j + n\bar{x}}{b+n} \tag{8}$$

as the posterior expectations of μ. Here j indicates one Normal-Wishart distribution out of the mixture. a_j and b_j are altered according to $a^* = a+n$ and $b^* = b+n$. The Bayesian estimate of the covariance is

$$\mathrm{E}(\Sigma / D) = \frac{1}{a+n-k-1}\Lambda^* \tag{9}$$

where *k* is the dimension (*k* = 14 number of network points in this study). Since the specification of a complete informational prior turns out to be practically impossible for the covariance structure we choose an approach which can be compared to a modified empirical Bayes setup. First we take one data set as a learning sample equipped with non-informative prior. By this first step we achieve a reliable prior guess of Σ used as Λ. Namely, we insert the empirical covariance matrix of the learning sample as prior covariance matrix. In alternating these starting sets and comparing the results we analyze the sensitivity of the model concerning the prior covariance structure. Note that Bayesian estimates coincide with standard least square estimates in case of non-informative priors.

2.2. Gibbs-Sampling

Since all measurements in the network are highly correlated the distributions (posterior as well as prior) of the parameters become high dimensional. Therefore an algorithm is needed to handle the posterior distributions which incorporate the special covariance structure of the data. The generic Gibbs algorithm aims at reproducing the posterior density and associated estimates in an automatic manner. Gibbs sampling represents a specific application of general MCMC (Markov Chain Monte Carlo) processes (Robert and Casella 2004).

The Gibbs sampling algorithm works as follows. The distribution of the *k* dimensional stochastic variable $x = (x_1, ..., x_k)$ is simulated by the following transition from step *s* to *s+1*. The basic idea of the algorithm is the simulation of the full conditional distributions where the components of the stochastic vector are replaced in turn. Let $x_1^{[0]}, ..., x_k^{[0]}$ denote a proper starting value for network point coordinates and $x_1^{[s]}, ..., x_k^{[s]}$ the result of the sth step. The components of $x_1^{[s+1]}, ..., x_k^{[s+1]}$ are generated according the conditional distributions *D*,

$$\begin{aligned} Z_1 &\sim D(Z / x_2^{[s]}, ..., x_k^{[s]}) \\ Z_2 &\sim D(Z / Z_1, x_3^{[s]}, ..., x_k^{[s]}) \end{aligned} \tag{10}$$

Iteration at for the kth component reads $Z_k \sim D(Z / Z_1, ...Z_{k-1})$. The vector $(Z_1, ..., Z_k) = (x_1^{[s+1]}, ...x_k^{[s+1]})$ serves as realization in the *s+1* th step. The conditional distributions

are one dimensional and therefore the algorithm doesn't change in structure if *k* increases. While working with normal distributions, a data augmentation leads to conditionally independent variables. That ensures several convergence properties of the procedure as well.

In our network we introduced non-informative priors. These priors are based on the Fisher information and are improper. Since the conditional distributions can be obtained directly the Gibbs sampler works if we start the procedure within a certain region. Standard arguments of convergence of the Gibbs sampler fail in the case of non-informative priors. A monitoring of the simulation and the corresponding estimates is needed and we dealt with this problem by using data sets as training samples. Therefore we reach a state of the procedure where a proper prior is given.

If the Bayesian estimate of the parameter is needed only instead of the complete posterior distribution a version of Gibbs sampling is carried out for computing the parameter estimates. Such algorithms are called EM-algorithm and are originally introduced to calculate maximum likelihood estimates. The stepwise calculation of a maximum likelihood estimate involves the cyclic exchange of the parameters as in the above calculation combined with a maximization step (Felsenstein 2001). If the prior distribution is non-informative maximum likelihood and Bayes estimation coincide for normally distributed data. The calculation of Bayesian estimates reduces to the calculation of the least square estimates if non-informative priors are used only.

3. Application

3.1.Information about study area and individual adjustment results for different periods

In this study, repeated GPS surveys at certain time periods, belonging to a geodetic deformation network in Macka in the province of Trabzon in north eastern Turkey (Figure 1) are used. This network was established in the year 2000 for landslide monitoring within the project of "Determination of landslides by dynamic models" by researchers from Geodesy and Photogrammetry Engineering at Karadeniz Technical University.

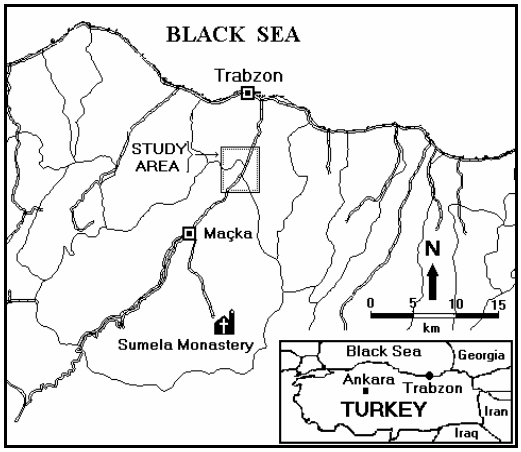


Figure 1. Location map of study area (Yalcinkaya and Bayrak 2003)

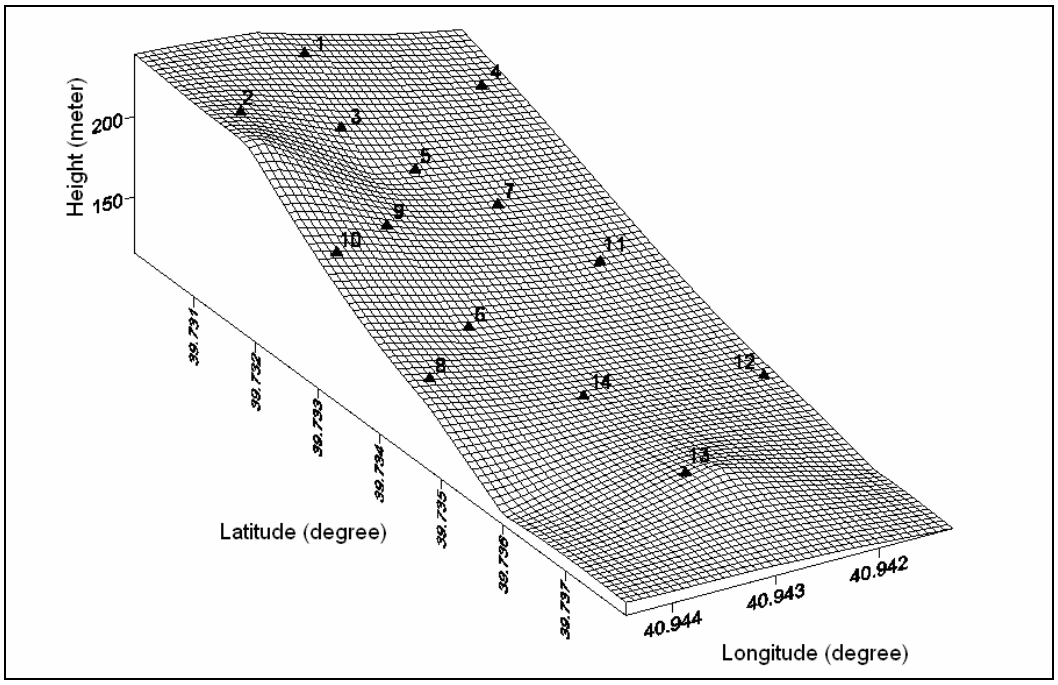


Figure 2. 3D-view of study area

The network consists of 14 points, with four of them (2, 8, 10, and 13) in solid ground. The other points (1, 3, 4, 5, 6, 7, 9, 11, 12, and 14) were placed into moving material (Figure 2). All of them were built with pillars. Geodetic deformation measurements made between November 2000, February 2001, and May 2001 with Ashtech GPS receivers in a static mode, as periodic data were evaluated by Geogenius-2000 software. The outputs of this software are baselines (l) between network points and cofactor matrices of each baseline (Q) (Yalcinkaya and Bayrak 2003; 2005). With these outputs from GPS software and approximate coordinates for network points (see Table 1), we can write our observation equations in Gauss Markov Model (GMM) to estimate network point positions by least-squares estimation. The results from least-squares estimation are used as input data for Bayes-Updating and Gibbs-Sampling algorithms.

Table 1. Approximate values for station coordinates

No	X(m)	Y(m)	Z(m)
1	3710656.9537	3083935.0998	4157818.3182
2	3710501.7100	3084005.1000	4157907.9600
3	3710542.0823	3084033.1843	4157788.0175
4	3710709.5388	3084028.6271	4157648.6441
5	3710480.9725	3084117.3938	4157741.7704
6	3710274.8591	3084218.3811	4157799.2949
7	3710479.6402	3084171.0299	4157677.5811
8	3710205.6283	3084212.3195	4157845.5018
9	3710389.3907	3084138.4578	4157788.8566
10	3710292.4179	3084153.5430	4157861.5579
11	3710442.6596	3084257.8606	4157623.1696
12	3710342.5861	3084401.6709	4157575.3479
13	3710193.7381	3084419.1729	4157702.3676
14	3710280.1590	3084340.5894	4157711.3664

The GMM is a linear mathematical model consisting of functional and stochastic relations. It relates observations to parameters. In matrix notation it takes the following form

$$E(l) = Ax \quad l = Ax + \varepsilon \tag{11}$$

$$E(\varepsilon\varepsilon^T) = \Sigma = \sigma_0^2 Q \tag{12}$$

This theoretical model is rewritten to estimate of parameters (network station coordinates) from real data (baselines between network points) in the following form

$$l + V = A\hat{x} \quad P = Q^{-1} \tag{13}$$

where l, $n-$vector of observations; $E(.)$, expectation operator; x, $u-$vector of unknown parameters; A, $nxu-$matrix of known coefficients; ε, $n-$vector of errors; Σ, $nxn-$covariance matrix; σ_0^2 a priori variance factor; Q, $nxn-$cofactor matrix of observations; P, $nxn-$weight matrix of observations; v, $n-$vector of residuals. The estimation function $\hat{x} = (A^T PA)^{-1}(A^T Pl)$ is used to get estimated values of network point coordinates (see Table 2).

Table 2. Estimation values for coordinates

November 2000 (1. period)			February 2001 (2. period)			May 2001 (3. period)		
dx (cm)	**dy(cm)**	**dz(cm)**	**dx(cm)**	**dy(cm)**	**dz(cm)**	**dx(cm)**	**dy(cm)**	**dz(cm)**
0.0060	0.0075	0.0259	5.6056	-5.3574	2.6970	13.8253	-16.2334	4.9553
0.0009	0.0020	0.0063	-0.4475	0.5205	-0.0910	.2.3584	2.2657	-0.8791
-0.0041	0.0025	-0.0305	-0.5343	-0.3207	0.2538	0.6371	-2.9547	6.4969
0.0051	0.0003	0.0132	-1.7004	0.0568	-1.4145	-4.0736	2.5040	-2.9982
0.0002	-0.0071	0.0016	0.1053	0.0840	-0.1360	2.4213	-0.9610	0.6777
-0.0028	-0.0010	-0.0022	-0.3131	1.0402	0.1334	-2.0397	2.2949	-1.2151
0.0077	-0.0006	-0.0007	-0.3501	0.1160	0.3149	-0.0892	-0.4882	0.7741
-0.0171	-0.0091	-0.0146	-0.4678	0.8057	-0..0393	-2.3664	2.7812	-1.2505
-0.0044	0.0007	-0.0018	0.2791	0.6208	-0.0446	1.6777	0.9220	0.3467
-0.0008	0.0014	-0.0077	-0.9562	-0.2630	-0.9319	-2.5536	2.7336	-1.9813
-0.0043	-0.0016	0.0034	0.3606	1.5477	0.6525	-0.8679	2.5048	0.0052
0.0033	0.0051	0.0105	-0.7386	0.1630	-0.4080	-2.1785	1.6472	-0.8665
0.0024	-0.0017	-0.0074	0.2527	0.7206	-0.2644	-1.3375	1.5609	-2.3790
0.0080	0.0017	0.0040	-0.5899	0.2656	-0.7217	-0.6966	1.4230	-1.6862

$Q_{xx} = (A^T PA)^{-1}$, $uxu-$ cofactor matrix of estimated parameters; $K_{xx} = \sigma^2 Q_{xx}$, $uxu-$variance-covariance matrix of estimated parameters; $m_{(xyz)} = \sqrt{Q_{xx}}$, $u-$mean-square error for estimated coordinates; and $m_p = \sqrt{m_x^2 + m_y^2 + m_z^2}$, $\frac{u}{3}-$ point positioning error for network points. All network points divided in groups looking at their point positioning error values. By looking at the m_p values of points, three different groups in each period are determined as accurate, more accurate and the most accurate. The starting points are selected from these groups. i.e, point 4 are fallen in the first group with the smallest mp value, point 3 in the second group, and point 13 in the third group with nearly the biggest mp value at every period. The starting points are selected as 1, 3, 4, and 13 (see Table 3). These types of grouping for network points are used for Gibbs Sampling algorithm later.

Table 3. m_p(cm) values for the all network points calculated from individual adjustment results in each period

Points	1. period	2. period	3. period
1	0.178	0.205	0.237
2	0.171	0.225	0.266
3	0.183	0.247	0.269
4	**0.150**	**0.175**	**0.207**
5	0.172	0.257	0.269
6	0.160	0.243	0.263
7	**0.154**	**0.208**	**0.246**
8	0.265	0.340	0.383
9	0.182	0.249	0.309
10	0.217	0.344	0.286
11	0.196	0.201	0.251
12	0.273	0.283	0.324
13	0.312	0.277	0.371
14	0.179	0.231	0.273

3.2.Comparison between different Bayes Updating and Gibss Sampling Algorithms

The parameter estimation with Gibbs-Sampling is applied with different starting points and with Bayes updating with different prior information (Roberts and Rosenthal 1998). The results getting with different information (starting points and prior information) in two different algorithms are compared to understand the importance of using different input parameters in a parameter estimation problem. Coordinate parameters estimated from Gibbs-Sampling (GS) algorithm and Bayes-Updating (BU) are called as respectively (X_{GS}, Y_{GS}, Z_{GS}) and (X_{BU}, Y_{BU}, Z_{BU}). Parameters of the difference are calculated; one is the difference between Bayes-Updating and Gibbs-Sampling algorithms (BU_GS), one is between Gibbs-sampling algorithms (GS) and the other is Bayes-Updating algorithms (BU). These differences are calculated as

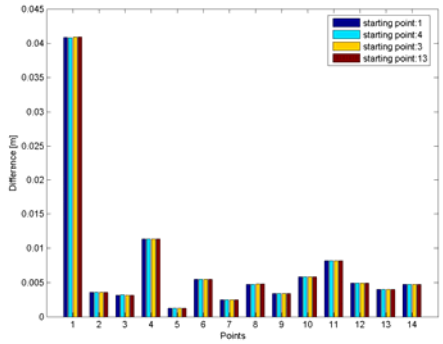
$$
\begin{aligned}
(BU_GS)_n &= \sqrt{(X_{BU(j)} - X_{GS(i)})_n^2 + (Y_{BU(j)} - Y_{GS(i)})_n^2 + (Z_{BU(j)} - Z_{GS(i)})_n^2} \\
(GS)_n &= \sqrt{(X_{GS(i)} - X_{GS(k)})_n^2 + (Y_{GS(i)} - Y_{GS(k)})_n^2 + (Z_{GS(i)} - Z_{GS(k)})_n^2} \\
(BU)_n &= \sqrt{(X_{BU(j)} - X_{BU(l)})_n^2 + (Y_{BU(j)} - Y_{BU(l)})_n^2 + (Z_{BU(j)} - Z_{BU(l)})_n^2}
\end{aligned}
\qquad (14)
$$

Here, *i* and *k* indicates the starting point in Gibbs-Sampling algorithm (e.g., 1, 3, 4, or 13); *j* and *l* indicates prior information (e.g., the first, second or third period) and n indicates the points in network. For making comparison between different algorithms, the smoothness parameters *(VR)* are calculated on values of different algorithms. $f(1), f(2), \ldots, f(n)$ values are taken as *(BU_GS)* and *(BU)* values and *VR* values for all plots are calculated from $VR = \sum_{i=2}^{n} |f(i) - f(i-1)|$. The results are drawn (in Figure 3 and Figure 4) except for GS. Comparisons are made by using different types of prior information in Bayes updating and taking different starting points (4, 3, 13) in Gibbs sampling algorithm.

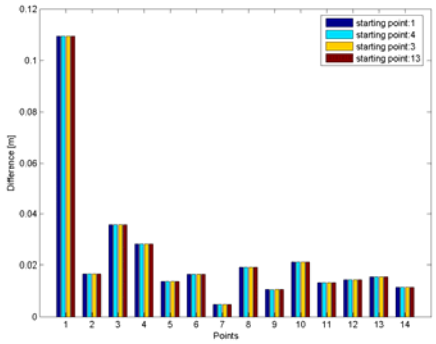
Table 4. VR values for differences between BU and GS Results (cm)

prior period	start point	prior period	start point	prior period	start point	prior period	start point
VR values: BU-GS differences (m) for 1. period							
2.period	**1**	**2.period**	**4**	**2.period**	**3**	**2.period**	**13**
0.07653		0.07636		0.07659		0.07659	
3.period	**1**	**3.period**	**4**	**3.period**	**3**	**3.period**	**13**
0.19574		0.19577		0.19576		0.19576	
VR values: BU-GS differences (m) for 2. period							
1.period	**1**	**1.period**	**4**	**1.period**	**3**	**1.period**	**13**
0.0780		0.0511		0.0780		0.0770	
3.period	**1**	**3.period**	**4**	**3.period**	**3**	**3.period**	**13**
0.1761		0.2870		0.1777		0.1775	
VR values: BU-GS differences (m) for 3. period							
1.period	**1**	**1.period**	**4**	**1.period**	**3**	**1.period**	**13**
0.3064		0.2646		0.3956		0.3713	
2.period	**1**	**2.period**	**4**	**2.period**	**3**	**2.period**	**13**
0.2862		0.2160		0.3113		0.3066	

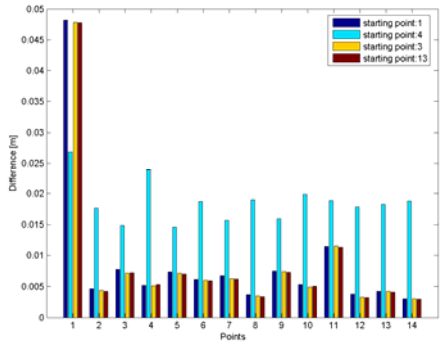
In case using 2. period as a prior in updating for 1. period, the differences on all points are approximately same except of point 1. In addition, the difference on point 4 is bigger than the others on 1.14 cm level after the point 1. The differences on the other points are slightly the same and between the values 0.1 and 0.8 cm. It can be seen that the effect of the starting point from most to least effective respectively 3, 13, 1, 4 (see Table 4 and Figure 4). Because VR values are nearly the same as for all situation, it can bee concluded for the 1. period that there is no difference to use different starting points in BU algorithm. In case of using 3. period as a prior in updating for 1. period, the differences on all points are different from each other and the point 1 has the biggest difference value on 11 cm level. The differences on the points 11, 12, 13, 14 are nearly the same and between 1.16 and 1.56 cm. The point 3 has the biggest difference value on 3.57 cm level after the point 1. It can bee seen from the VR values that there is nearly no difference between Bayes updating and Gibbs sampling with the starting points 1, 3, 4, and 13 (see Table 4 and Figure 3).



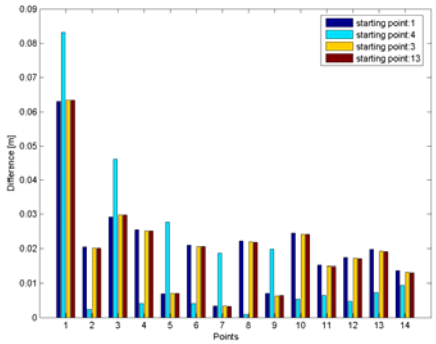
(a) BU with the 2.period as prior and GS with different starting points in the 1. period



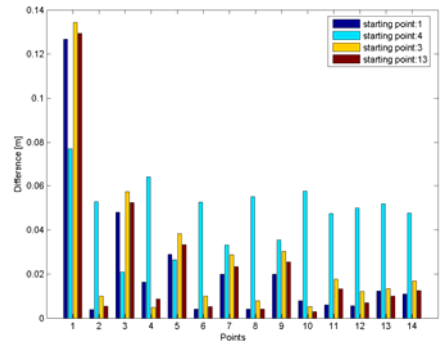
(b) BU with the 3.period prior as a prior and GS with different starting points in the 1. period



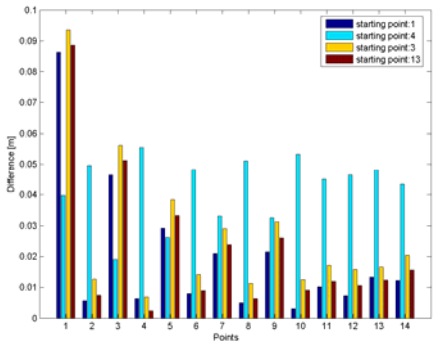
(c) BU with 1.period as prior and GS with different starting points in 2. period



(d) BU with 3. period as prior and GS with different starting points in 2.period



(e) BU with 1. period as prior and GS with different starting points in 3. period



(f) BU with 2. period as prior and GS with different starting points in 3. period

Figure 3. Differences between BU and GS Results

In case using 1. period as a prior in updating for 2. period, the differences on all points are approximately same except the point 1. The difference on point 11 is biggest after the point 1 while point 14 has the smallest difference compared to other points. When looking at the diagram indicating the differences between Gibbs sampling, with the starting point 4, and Bayes updating, the differences become higher on the points while difference becomes smaller on point 1. Apart from these, the difference on point 4 becomes bigger compared to other points. VR values for point 4 is the smallest. However, the other VR values are the same and bigger than the VR for point 4. There is nearly no difference between Bayes updating and Gibbs sampling with the starting points respectively 1, 3 and 13. There is nearly no difference between Bayes updating and Gibbs sampling with the starting points respectively 1, 3 and 13. In case of using 3. period as a prior in updating for

2. period, the differences on points are the most different in using the other priors. The difference on point 1 has also the biggest value on 6.00 cm level while point 7 has the smallest value on 0.3 cm level. Apart from this, differences on points respectively 2, 3, 4 are nearly the same nearly on 2.00 cm level and differences on points respectively 11, 12, 13, 14 are also approximately same on 1.00 cm level compared the other point differences. When looking at the diagram indicating the differences between Gibbs sampling, with the starting point 4, and Bayes updating, the difference on points become higher all points including point 1. VR values for points respectively 1, 3, 13 are nearly the same and smaller than for point 4. The reason of this may be because of using of the most effective informative prior (3. period) and the least effective starting point (point 4) at the same time. VR values for points respectively 1, 3, 13 are from biggest to smallest respectively point 4, point 13 and point 1. It can be said that the effectiveness of the starting point is from most to least respectively 3, 13, 1. Apart from these, differences on points respectively 2, 3, 4 have different structure. When looking at the point 7, difference becomes higher. Differences on points respectively 10, 11, 12, 13, 14 are nearly the same between the values 0.5 and 0.9 cm in compared to other points.

In case using 1. period as a prior in updating for 3. period, the differences on point 1 has the biggest value on level approximately 13 cm. The next biggest difference on point 3 with the value 5.74 cm. Points respectively 10, 11, 12, 13, 14 has slightly the same and small difference values on 0.5 and 1.7 cm level. The point 2 has the smallest value with the value 0.4 cm. In case using point 4 as a starting point, the difference on points becomes higher except point 1 with the small value 7,7 cm. The difference on point 2 becomes higher on 5 cm level. Points respectively 10, 11, 12, 13, 14 has slightly the same and small difference values on approximately 4 and 5 cm level. It can be said that the effectiveness of the starting point is from most to least respectively 3, 13, 1. VR value for point 4 is the smallest. The other VR values for the other point are from the biggest to smallest points are respectively 3, 13 and 1. It can be said that the effectiveness of the starting point is from most to least respectively 3, 13, 1. When using 2. period as a prior in updating for 3. period, the biggest difference also on point 1 with the value nearly 9 cm. Points 3 has the biggest difference with the approximate value 5 cm after points respectively 1, 10, 11, 12, 13, 14 has approximately the same difference value between 0.9 and 2 cm level. In case starting point 4, the difference on point 1 becomes smaller with the value 4.00 cm value. However the general differences on all points becomes higher. In addition small value points become high value points. e.g. point 3 becomes the smallest value point with the value 2 cm compared to other points. VR value for the point 4 is the smallest. The other VR values from the biggest to smallest respectively are points 3, 13 and 1 (see Table 4 and Figure 3).

3.3.Comparison between different Gibbs Sampling algorithms

As indicated Section 3.1., all network points are divided into groups according to their point positioning errors (m_p) values and this grouping is used for GS Algorithm, to decide on which point to be started. For the 1. period, there is no difference using different starting points. For the 2. period, there is a big difference between the normal case (take the points in ascending order respectively 1, 2, 3,..,14) and the case 4 (using point 4 as starting). On the other hand, there is slight difference among the normal case and 3 and 13 cases. The same inferences can be made for the 3. period. However, the differences are bigger than the 2. period. e.g. The difference the normal case and the case 4 is nearly 2 cm in the 2. period. This difference is nearly 5cm in the 3. period. As a result, it can be said that from the most to lest effective starting points are respectively 4, 3, 13. However, the degree of effectiveness of point 3 and point 13 are nearly the same.

3.4. Comparision between different Bayes Updating Algorithms

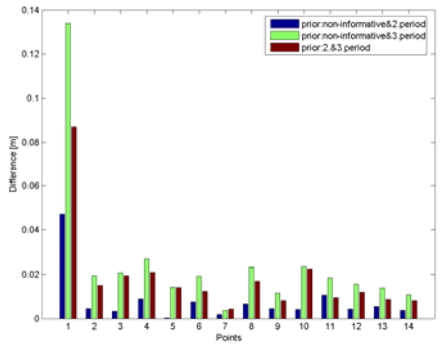
In Bayes Updating algorithm, it is a important question which prior information is most informative than others. To answer to this question in this study, inference should be made in different periods with different prior information. For every algorithm, VR are also calculated. Differences between these algorithms in the aspects of different prior information are following:

By using of non-informative information and the 2. period as a prior in 1. period, the biggest difference is determined on point 1 with the value 4cm. The smallest difference on point 5 with the value 0.03 cm is determined. For the same period, by using of non-informative information and the 3. period as a prior, the biggest difference is determined on point 1 with the value 13cm. The smallest difference on point 7 with the value 0.4cm. Finally, 2. period and the 3. period uses as priors, the biggest difference is determined on point 1 with the value 8.7cm and the smallest difference on point 7 with the value 0.43cm (see Table 5 and Figure 4) .

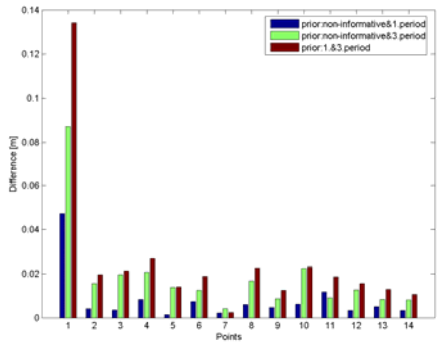
Table 5. VR values for differences between BU algorithms (cm)

BU differences for 1.period					
Prior period		**Prior period**		**Prior period**	
non-informative	**2.period**	**non-informative**	**3.period**	**1.period**	**3.period**
0.0935		0.2103		0.1495	
BU differences for 2.period					
non-informative	**1.period**	**non-informative**	**3.period**	**1.period**	**3.period**
0.0902		0.1494		0.2094	
BU differences for 3.period					
non-informative	**1.period**	**non-informative**	**2.period**	**1.period**	**2.period**
0.2138		0.1551		0.0855	

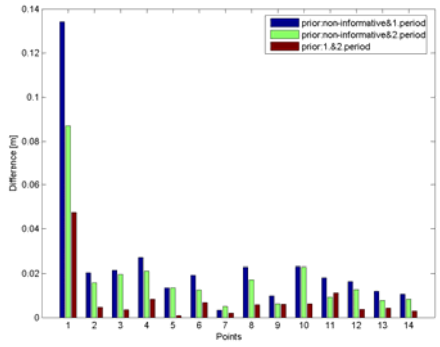
For the 2. period by using of non-informative information and the 1. period as a prior let the biggest difference on point 1 with the value 4.72cm and the smallest difference on point 5 with the value 0.14cm. By using of non-informative information and the 3. period as a prior, the biggest difference is determined on point 1 with the value 8.73cm and the smallest difference on point 7 with the value 0.42cmm. For the period, 1. period and the 3. period are used as prior, the biggest difference is determined on point 1 with the value 13.42cm and the smallest difference on point 7 with the value 0.24cm. Almost the same procedures are done for the 3. period to evaluate Bayes Updating algorithm with different prior information. When we use non-informative information and the 1. period as a prior, we calculate the biggest difference on point 1 with the value 13.41cm and the smallest difference on point 7 with the value 0.33cm. In case of using non-informative information and the 2. period as a prior, the biggest difference is determined on point 1 with the value 8.71cm. The smallest difference on point 7 with the value is 0.50cm. For the same period, 1. period and the 2. period used as priors. This prior information let the biggest difference on point 1 with the value 4.74cm and the smallest difference on point 5 with the value 0.08cm (see Table 5 and Figure 4).



(a) BU by different prior information in 1. period



(b) BU by different prior information in 2. period



(c) BU by different prior information in 3. period

Figure 4. Differences between BU algorithms

4. Conclusions

In this paper we introduced estimation procedures Bayes-Updating and Gibbs-Sampling for deformation monitoring networks based on Bayesian theory, which allows accounting for prior information about the coordinate parameters. Gibbs-Sampling applies the simulation for the conditional distributions of parameter estimation of deformation network points. In Bayes-Updating algorithm, we can analyze the sensitivity of the model concerning the prior information coming from different epoch measurements in deformation network. Concerning usage of these two algorithms, Gibbs-Sapling Algorithm take more time compared to Bayes-Updating because simulation algorithm of Gibbs-Sampling depends on the information about the quality of starting points itself. In this study we use network with 14 points, when the user increase the number of network points, Gibbs-Sampling becomes complicated for such networks.

The results of two parameter estimation procedures were compared to each other and we get the some differences on estimation of network point coordinates on cm level. When we consider all these parameter estimation results as input data for deformation analysis, it should be point out here that the estimation procedure which is used in individual deformation epochs has also big importance as deformation analysis procedures itself.

5. Acknowledgements: This study was done during the first author's research stay in the Institute of Statistics and Probability Theory at Vienna University of Technology (TU Vienna) and financially supported by Natural Science Institute at Karadeniz Technical University and TU Vienna. The first author is thankful to Prof. Reinhard Viertl, head of the Statistics and Probability Theory Institute in TU Vienna, for providing her the full scientific facilities during her research stay in this institute.

6. References

Albertella A., Cazzaniga N., Sansò F., Sacerdote F., Crespi M., Luzietti L.(2005). Deformations detection by a Bayesian approach: prior information representation and testing criteria definition, ISGDM2005 - IAG Symposium volume n. 131.

Caspary W.F. (2000). Concept of network and deformation analysis, Monograph 11, School of Geomatics Engineering, The University of New South Wales, Australia.

Felsenstein K. (2004). Mathematische methoden für die interpretation von risiken. Imago Hominis, Vol. 11, 261-269.

Felsenstein K. (2001). Bayesian interpolation schemes for monitoring systems. In: Advances in model-oriented design and analysis, Vol. 6, Physica-Verlag, Heidelberg.

Felsenstein K. (1996). Bayes'sche statistik für kontrollierte experimente. Vandenhoeck& Ruprecht, Göttingen.

Robert C. and Casella G. (2004). Monte Carlo statistical methods Springer, New York.

Robert C. (2001). The Bayesian Choice, Springer, New York.

Roberts G. and Rosenthal J.(1998). Markov chain Monte Carlo: Some practical implications of theoretical results. Canadian J. Statist., Vol. 25, 5-32.

Yalcinkaya M. and Bayrak T. (2005). Comparison of static, kinematic and dynamic geodetic deformation models for Kutlugün landslide in northeastern Turkey, Natural Hazards, January 2005, Volume 34, Issue 1, pp. 91-110.

Yalcinkaya M. and Bayrak T. (2003). Dynamic model for monitoring landslides with emphasis on underground water in Trabzon province, Northeastern Turkey, Journal of Surveying Engineering, August 2003, Volume 129, Issue 3, pp. 115-124