

# Modelling local GPS/levelling geoid undulations using Support Vector Machines

Piroska Zaletnyik / Lajos Völgyesi / Béla Paláncz

Received 2007-07-24

## Abstract

Support vector machines (SVM) with wavelet kernel has been applied to the correcting gravimetric geoid using GPS/levelling data. These data were divided into a training and a validation set in order to ensure the extendability of the approximation of the corrector surface. The optimal parameters of the SVM were considered as a trade-off between accuracy and extendability of the solution in order to avoid overlearning. Employing 194 training points and 110 validation points, SVM provided an approximation with less than 3 cm standard deviation of the error and nearly perfect extendability.

## Keywords

geoid, corrector surface, GPS, support vector regression, wavelet kernel

## Acknowledgement

The authors wish to thank A. Kenyeres and the Hungarian Institute of Geodesy, Cartography and Remote Sensing for providing GPS/levelling data of Hungary and also the Hungarian Research Fund OTKA project T046718.

## Piroska Zaletnyik

Department of Geodesy and Surveying, Faculty of Civil Engineering, Budapest University of Technology and Economics. PO Box 91, H-1521, Hungary  
e-mail: [piri@agt.bme.hu](mailto:piri@agt.bme.hu)

## Lajos Völgyesi

Department of Geodesy and Surveying, Faculty of Civil Engineering, Budapest University of Technology and Economics; Research Group of Physical Geodesy and Geodynamics of the Hungarian Academy of Sciences. PO Box 91, H-1521, Hungary  
e-mail: [volgyesi@eik.bme.hu](mailto:volgyesi@eik.bme.hu)

## Béla Paláncz

Department of Photogrammetry and Geoinformatics, Faculty of Civil Engineering, Budapest University of Technology and Economics. PO Box 91, H-1521, Hungary  
e-mail: [palancz@epito.bme.hu](mailto:palancz@epito.bme.hu)

## 1 Introduction

The accuracy of the gravimetric geoid can be significantly improved using GPS/levelling measurements. The new, adjusted geoid can be constructed as the old one plus the so called corrector surface, the difference between the new and old geoid. Recently, wide variety of higher-order parametric and non-parametric surfaces have been used as corrector surfaces, such as polynomial models Fotopoulos and Sideris 2005 [4], spline interpolation Featherstone 2000 [2] and Zaletnyik et al 2006 [13], least squares collocation (LSC) Iliffe et al 2003 [5], kriging Nahavandchi and Soltanpour 2004 [8], combined least squares adjustments Fotopoulos 2005 [3], and various other surfaces. Most recently Zaletnyik et al. 2007 [14] employed thin plate spline (TPS) surface, solving the problem via finite element method. Suffice it to say, there are numerous surface – fitting options, each with their own advantages and disadvantages, which will not be discussed nor debate here.

Concerning application of soft computing technique Kavzoglu and Saka 2005 [6] and Lao-Sheng Lin 2006 [7] employed artificial neural network (ANN) for approximating the GPS/levelling geoid instead of the corrector surface itself. Both of them applied feed-forward ANN with the standard sigmoid activation functions and different number of hidden layers. Zaletnyik et al 2006 [13] also used ANN but with radial bases activation function (RBF) and regularization in the training phase. Soltanpour et al 2006 [11] used second generation wavelets to approximate corrector surface directly. This technique let extend the classical wavelet approximation, which requires regularly spaced/sampled data, for unregularly spaced dataset, Sweldens 1997 [12].

Another soft computing technique is represented by the support vector machines (SVM), which are learning algorithms that have many applications in pattern recognition and nonlinear regression. In this study we propose to apply support vector machine with wavelet kernel for modelling the corrector surface.

## 2 Support Vector Machines for Regression

The problem of regression is that of finding a function which approximates mapping from an input domain to the real numbers based on a training sample. We refer to the difference between the hypothesis output and its training value as the residual of the output, an indication of the accuracy of the fit at this point. We must decide how to measure the importance of this accuracy, as small residuals may be inevitable while we wish to avoid large ones. The loss function determines this measure. Each choice of loss function will result in a different overall strategy for performing regression. For example least square regression uses the sum of the squares of the residuals.

Although several different approaches are possible, we will provide an analysis for generalization of regression by introducing a threshold test accuracy, beyond which we consider a mistake to

have been made. We therefore aim to provide a bound on the probability that a randomly drawn validation point will have accuracy less than  $\epsilon$ . One way of visualizing this method of assessing performance is to consider a band of size  $\pm \epsilon$  around the hypothesis function any training points lying outside this band are considered to be training mistakes, see Fig. 1.

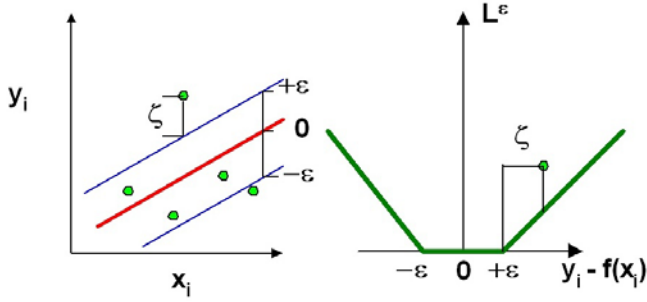


Fig. 1. Linear  $\epsilon$ -insensitive loss function  $L^\epsilon(x, y, f)$

Therefore we can define a so called  $\epsilon$ -insensitive loss function. The linear  $\epsilon$ -insensitive loss function  $L^\epsilon(x, y, f)$  is defined by

$$L^\epsilon(x, y, f) = \left( |y - f(x)| \right)_\epsilon = \max(0, |y - f(x)| - \epsilon) \quad (1)$$

where  $f$  is a real - valued function on a domain  $X \subset \mathfrak{R}^n$ ,  $x \in X$  and  $y \in \mathfrak{R}$ . Similarly the quadratic  $\epsilon$ -insensitive loss is given by

$$L_2^\epsilon(x, y, f) = \left( |y - f(x)| \right)_\epsilon^2. \quad (2)$$

Support vector regression (SVR) uses an admissible kernel, which satisfies the Mercer's condition to map the data in input space to a highdimensional feature space in which we can process a regression problem in linear form. Let  $x \in \mathfrak{R}^n$  and  $y \in \mathfrak{R}$ , where  $\mathfrak{R}^n$  represents input space, see Cristianini and Shawe-Taylor 2003 [1]. By some nonlinear mapping  $\Phi$ , the vector  $x$  is mapped into a feature space in which a linear regressor function is defined,

$$y = f(x, w) = \langle w, \Phi(x) \rangle + b. \quad (3)$$

We seek to estimate this  $f$  function based on independent uniformly distributed data  $\{\{x_1, y_1\}, \dots, \{x_m, y_m\}\}$ , by finding  $w$  which minimizing the quadratic  $\epsilon$ -insensitive losses, with  $\epsilon$ , namely the following function should be minimize

$$c \sum_{i=1}^m L_2^\epsilon(x_i, y_i, f) + \frac{1}{2} (\|w\|)^2 \rightarrow \min \quad (4)$$

where  $w$  is weight vector and  $c$  is a dimensionless constant parameter. Considering dual representation of a linear regressor in (3),  $f(x)$  can be expressed as

$$f(x) = \sum_{i=1}^m \beta_i y_i \langle \Phi(x_i), \Phi(x) \rangle + b \quad (5)$$

which means that the regressor can be expressed as a linear combination of the training points. Consequently using an admissible kernel, a kernel satisfying the Mercer's condition, Paláncz et al 2005 [10], we get

$$f(x) = \sum_{i=1}^m \beta_i y_i K(x_i, x) + b = \sum_{i=1}^m \alpha_i K(x_i, x) + b. \quad (6)$$

By using Lagrange multiplier techniques, the minimization problem of (4) leads to the following dual optimization problem

$$\begin{aligned} \text{maximize } W(\alpha) &= \sum_{i=1}^m y_i \alpha_i - \epsilon \sum_{i=1}^m |\alpha_i| \\ &\quad - \frac{1}{2} \sum_{i,j=1}^m \alpha_i \alpha_j \left( K(x_i, x_j) + \frac{1}{c} \delta_{ij} \right) \\ \text{subject to } &\sum_{i=1}^m \alpha_i = 0. \end{aligned} \quad (7)$$

Let

$$f(x) = \sum_{i=1}^m \alpha_i^* K(x_i, x) + b^*, \quad (8)$$

where  $\alpha^*$  is the solution of the quadratic optimization problem and  $b^*$  is chosen so that  $f(x_i) = y_i - \epsilon - \frac{\alpha_i^*}{c}$  for any  $\alpha_i^* > 0$ .

For samples are inside the  $\epsilon$ -tube,  $\{x_i : |f(x_i) - y_i| < \epsilon\}$ , the corresponding  $\alpha^*$  is zero. It means we do not need these samples to describe the weight vector  $w$ . Consequently

$$f(x) = \sum_{i \in SV} \alpha_i^* K(x_i, x) + b^* \quad (9)$$

where

$$SV = \{i : |f(x_i) - y_i| \geq \epsilon\}. \quad (10)$$

These  $x_i$  sample vectors,  $\{x_i : i \in SV\}$ , that come with nonvanishing coefficient  $\alpha^*$  are called support vectors.

### 3 Wavelet Kernel

In our case, we select wavelet kernel for  $n = 2$ , which provides better local characterization than other kernels, see Zhang et al. 2004 [15] was proved to be very efficient in many regression problems, e.g. Paláncz et al. 2005 [10].

Wavelet kernel with  $a \in \mathfrak{R}^1$  and all compact  $X \subset \mathfrak{R}^n$ ,

$$K(x, z) = \prod_{i=1}^n \left( \cos \left[ 1.75 \frac{x_i - z_i}{a} \right] \exp \left[ -\frac{(x_i - z_i)^2}{2a^2} \right] \right). \quad (11)$$

### 4 Dataset for the numerical computations

The original gravimetric geoid was modelled via third order spline approximation, which provides a fairly good approximation, a fitting with 1 - 2 cm error in height, see Zaletnyik et al 2006 [13].

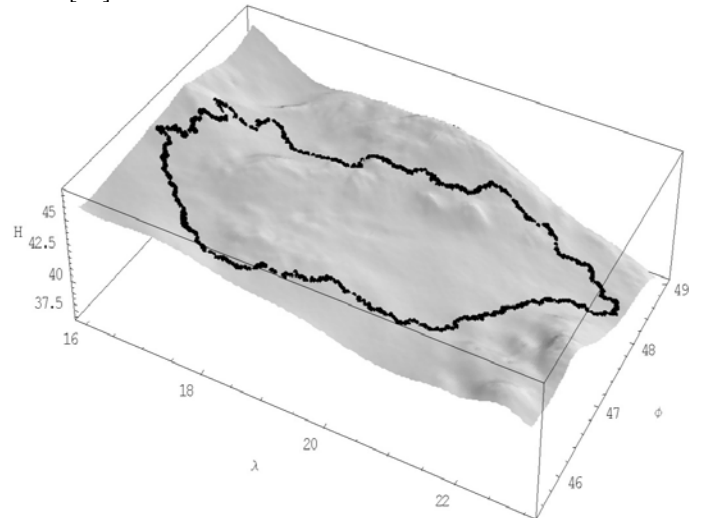


Fig.2. The Hungarian gravimetric geoid

For modelling the corrector surface, there are 304 GPS/levelling data available. One part of these data was employed as training set (194 measurements), the other part was used as validation set (110 measurement), see Fig. 3.

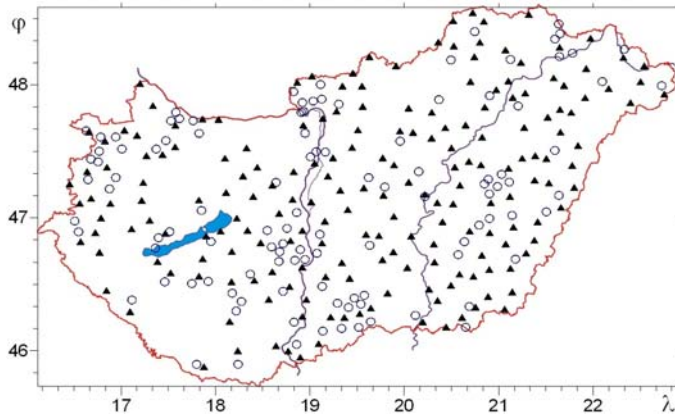


Fig. 3. GPS/levelling data - the training set (circles), and the validation set (triangles)

### 5 Parameter Study

In order to achieve an efficient fitting, one should find the optimal parameter of the applied kernel function ( $a$ ) as well as the proper values of  $c$  and  $\epsilon$ . Parameter investigations showed that with increasing values of  $c$  and  $\epsilon$ , the regression error (root mean square error,  $RMSE$ ) decreases on the training and the validation set, too. In our case  $c = 400$  and  $\epsilon = 10^{-3}$  proved to be reasonable values, while for greater values the changes of  $RMSE$ 's are negligible.

However the value of the parameter  $a$  has a strong influence on the quality of the approximation. Table 1. shows the change of the sum of  $RMSE$ 's (that of training and validation set, respectively) as well as the ratio of these  $RMSE$ 's, namely introducing

$$\eta = \frac{RMSE_V}{RMSE_t} \quad (12)$$

a ratio indicates how reliably can we extend our regression model for not measured data. The ideal value is 1. If  $\eta \gg 1$ , then so called overlearning effect takes place.

Table 1. The result of the parameter study in case of  $c = 400$  and  $\epsilon = 10^{-3}$

$a$	$RMSE_t$ [cm]	$RMSE_V$ [cm]	$RMSE_{total}$ [cm]	$\eta$
0.50	0.50	2.06	1.28	4.12
1.00	1.64	1.94	1.79	1.18
1.50	2.15	2.37	2.26	1.10
2.00	2.31	2.57	2.44	1.11
2.50	2.59	2.74	2.66	1.06

Table 2. Corrector Surface approximated by SVM

Method	Training set				Validation set				$\eta$
	SD [cm]	Min [cm]	Max [cm]	$RMSE_V$ [cm]	SD [cm]	Min [cm]	Max [cm]	$RMSE_V$ [cm]	
SVM regression	2.60	-7.74	6.47	2.59	2.75	-7.68	5.59	2.74	1.06

Fig. 4 shows the corrector surface in case of  $a = 0.5$  when the total error is small  $RMSE_{total} (= 1.28)$  but  $\eta = 4.12$  is high.

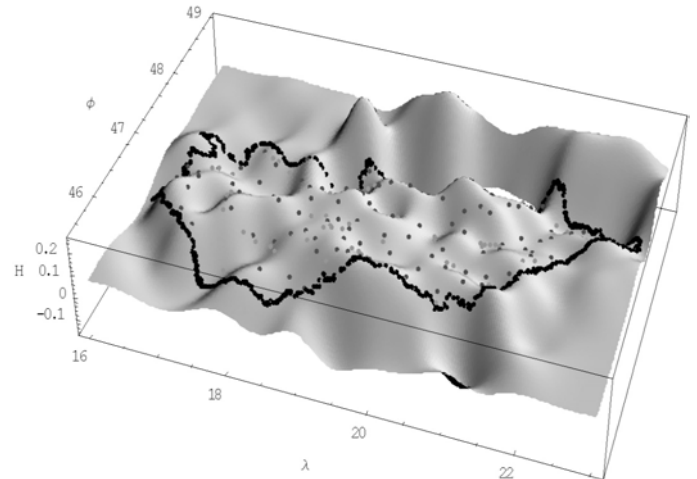


Fig. 4. Corrector surface in case of typical overlearning ( $a = 0.5$ )

This result indicates that one should make a trade – off between extendability ( $\eta$ ) and regression error ( $RMSE_{total}$ ). In our case we selected  $a = 2.5$ , which ensures small  $\eta$  and acceptable  $RMSE$  as well as smooth regression surface, see Fig. 5.

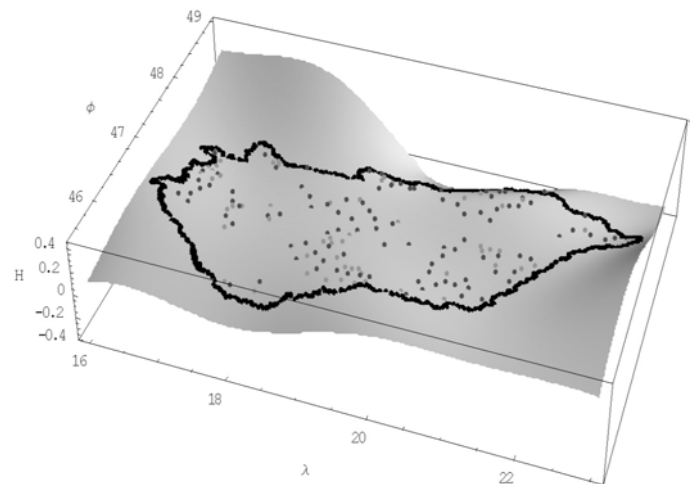


Fig. 5. Smooth and extendable regression surface

### 6 The model for the corrector surface

Using these parameter values ( $c = 400$  and  $\epsilon = 10^{-3}$  m and  $a = 2.5$ ) the computation with the wavelet kernel Eq. (11) was carried out. We used the *Mathematica* implementation of SVM regression, see Paláncz 2005 [9].

The analytical form of the corrector surface is

$$\begin{aligned} \Delta H = & -0.0370989 \\ & - 8.52042 e^{-0.08(47.2429-\varphi)^2 - 0.08(16.4483-\lambda)^2} \\ & \cdot \cos[0.7(47.2429-\varphi)] \cos[0.7(16.4483-\lambda)] \\ & - 11.4187 e^{-0.08(47.0987-\varphi)^2 - 0.08(16.5562-\lambda)^2} \\ & \cdot \cos[0.7(47.0987-\varphi)] \cos[0.7(16.5562-\lambda)] \\ & + \dots \\ & - 6.64333 e^{-0.08(48.0101-\varphi)^2 - 0.08(22.5098-\lambda)^2} \\ & \cdot \cos[0.7(48.0101-\varphi)] \cos[0.7(22.5098-\lambda)] \\ & - 4.17322 e^{-0.08(48.1296-\varphi)^2 - 0.08(22.5491-\lambda)^2} \\ & \cdot \cos[0.7(48.1296-\varphi)] \cos[0.7(22.5491-\lambda)] \\ & - 1.67954 e^{-0.08(47.9176-\varphi)^2 - 0.08(22.762-\lambda)^2} \\ & \cdot \cos[0.7(47.9176-\varphi)] \cos[0.7(22.762-\lambda)]. \end{aligned} \quad (13)$$

Table 2 shows that the standard deviation (SD) on the training as well as on the validation set is about 2.6 - 2.8 cm, which from practical point of view is acceptable and which is also important, that extendability coefficient is very good, near to unity. In the table RMSE- root mean squared error - is the square root of the mean of the error vector of the measurement points,

$$RMSE = \sqrt{\text{Mean} \sum_i (\Delta H(\varphi_i, \lambda_i) - \Delta H_i)^2} \quad (14)$$

The resulted corrector surface now is very smooth, see Fig. 5.

## 7 Adjusted Geoid

In order to get the corrected geoid, the corrector surface should be added to the original geoid surface, see Fig. 6.

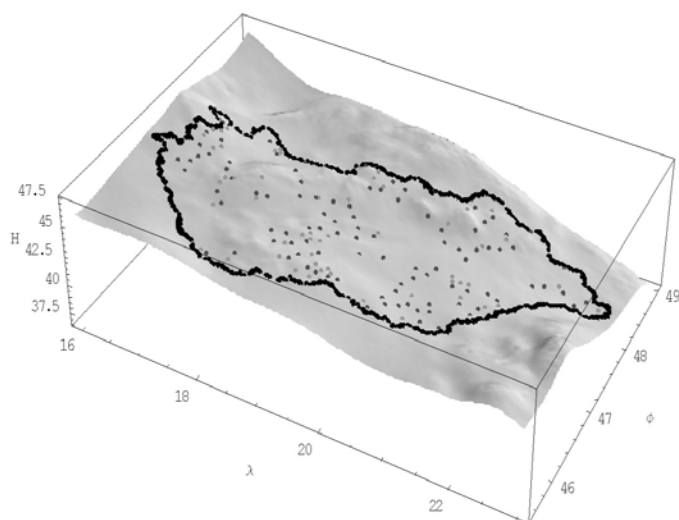


Fig.6 Adjusted geoid with the training and validation points

## 8 Conclusions

SVM with quadratic  $\epsilon$ -insensitive loss function was applied to constructing corrector surface for gravimetrical geoid, using GPS/levelling data. Employing wavelet kernel it turned out, that only the kernel parameter  $a$  has considerable influence on the quality of the approximation, while the SVM parameters  $c$  and  $\epsilon$  do not play important role in this case. The optimal parameters of the SVM were considered as a trade-off between accuracy and extendability of the solution in order to avoid overlearning. Employing 194 training points and 110 validation points, SVM provided an approximation with less than 3 cm standard deviation of the error and nearly perfect extendability. The corrector surface can be described via analytical form and directly implemented in a high level language, like C, in order to get high performance evaluation.

In the future, the investigation of the application of other type of kernel can be reasonable.

## References

- 1 **Cristianini N, Shawe-Taylor J**, An introduction to Support Vector Machines and other kernel - based learning methods, Cambridge, University Press, 2003.
- 2 **Featherstone W E**, Refinement of a gravimetric geoid using GPS and levelling data, J. Surv. Eng. **126/2** (2000), 27-56.
- 3 **Fotopoulos G**, Calibration of geoid error models via a combined adjustment of ellipsoidal, orthometric and gravimetric geoid height data. J. of Geod. **79/1-3** (2005), 111-123.
- 4 **Fotopoulos G, Sideris M G**, Spatial Modeling and Analysis of Adjusted Residuals over a network of GPS-levelling Benchmarks, GEOMATICA **59/3** (2005), 251-262.
- 5 **Illife J C, Ziebart M, Cross P A, Forsberg R, Strykowski G, Tscherning C C**, OGS02: a new model for converting GPS-derived heights to local height datums in Great Britain and Ireland. Surv. Rev. **37/290** (2003), 276-293.
- 6 **Kavzoglu T, Saka M H**, Modelling local GPS/levelling geoid undulations using artificial neural networks. J. Geod. **78** (2005), 520-527.
- 7 **Lao-Sheng Lin**, Application of a Back-Propagation Artificial Neural Network to Regional Grid-Based Geoid Model Generation Using GPS and Levelling Data. J. of Surv. Eng. **133/2** (2007) 81-89.
- 8 **Nahavandchi H, Soltanpour A**, An attempt to define a new height datum in Norway. The geodesy and Hydrography days 2004, 4-5 Nov. Sandnes, Norway.
- 9 **Paláncz B**, Support Vector Regression via Mathematica, <http://library.wolfram.com/infocenter/Mathsource/5270/> (2005)
- 10 **Paláncz B, Völgyesi L, Popper Gy**, Support Vector Regression via Mathematica, Periodica Polytechnica, Civil Eng. **49/1** (2005) 57-84.
- 11 **Soltanpour A, Nahavandchi H, Featherstone W E**, Geoid-type surface determination using wavelet-based combination of gravimetric quasi/geoid and GPS/levelling data, Geophysical Research Abstract, **8** (2006), 04612.
- 12 **Sweldens W**, The lifting scheme: a construction of second generation wavelets. SIAM J Math. Anal **29/2** (1997) 511-546.
- 13 **Zaletnyik P, Paláncz B, Völgyesi L, Kenyeres A**, Correction of gravimetric geoid using GPS/levelling data. (Accepted for publication in Geomatica) 2006. (In Hungarian)
- 14 **Zaletnyik P, Völgyesi L, Kirchner I, Paláncz B**, Thin Plate Spline Interpolation of the Corrector Surface via Finite Element Method, J. of Appl. Geod, 2007 (submitted)
- 15 **Zhang L, Zhou W, Jiao L**, Wavelet Support Vector Machine, IEEE Trans. Systems, Man and Cybernetics - Part B: Cybernetics, **4/1** (2004) 34-39.