Journal of Southeast University (English Edition) Vol.19 No.1 Mar. 2003 ISSN 1003—7985

Converting GPS height by a new method based on neural networks

Hu Wusheng Sha Yuejin

(College of Transportation, Southeast University, Nanjing 210096, China)

Abstract: The adjusted GPS height is the height above the surface of the WGS-84 ellipsoid. It is necessary to convert a GPS height into a normal height in engineering. The conicoid fitting method (CFM) and the neural networks method (NNM) are used for this purpose, but each of them has its advantages and disadvantages. After studying these two methods, a new method (abbr. CF&NNM) is conceived. The procedure of the CF&NNM is introduced. A practical engineering example is used to study these three different methods. The results by the three methods are listed. The CF&NNM method can produce better results than either the CFM or the NNM in deriving normal height from GPS height. The theory of the CF&NNM method is analyzed.

Key words: neural networks; algorithm of BP; GPS height; CF&NNM

Global positioning system (GPS) has been widely used in precise engineering surveys. It is well known that GPS provides more accurate horizontal positions than vertical positions. Much research has been done to improve the accuracy of GPS height.

The adjusted GPS height H_{GPS} is the height above the WGS-84 ellipsoid. In China, however, the normal height H_{Nor}, which is the height above the geoid calculated by using the mean normal gravity along the plumb line, is used in engineering applications. Therefore, it is necessary to convert H_{GPS} into H_{Nor}. The difference between them is called elevation abnormality ξ[1]:

$$\xi = H_{GPS} - H_{Nor} \tag{1}$$

If accurate H_{Nor} can be achieved by adjusting and converting H_{GPS}, it can be used to replace laborious geodetic leveling work. The CFM and NNM are often used for converting H_{GPS}.

CFM (conicoid fitting method) The main idea of CFM is to design a set of control points of which the H_{Nor} and H_{GPS} are known, and then the elevation abnormality ξ can be modeled by a polynomial of second degree[1]:

$$\xi(x,y) = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 xy + a_5 y^2 \tag{2}$$

where x and y are the horizontal coordinates of the control point; $a_0, a_1, \cdots, a_5$ are the unknown coefficients. At least six control points with known H_{GPS} and H_{Nor} are needed.

NNM (neural networks method) Artificial neural network is a relatively new branch of science. It is a highly simplified model of a complicated bio-neural system. Since the 1980's, scientists in many fields have spent a tremendous effort in studying artificial neural networks and made remarkable accomplishments[2]. Its accuracy is better than that of CFM[3].

CF&NNM After studying these two methods, a new method, called CF&NNM, to convert H_{GPS} is conceived[4]. An engineering example is given to demonstrate the three different methods.

1 The Structure and Algorithm of BP Neural Networks

Currently there are more than 40 types of neural network models. For converting H_{GPS}, a multi-layer fore-feedback BP structure is selected. The BP algorithm is widely used in engineering.

1.1 The structure of BP neural networks

The structure of BP neural networks is shown in Fig.1. It can be divided into five layers. For an ordinary engineering application, the input and output transformation layers are needed because the Sigmoid standard active function $f(x)$ ranges from 0 to 1.

1.2 The algorithm of BP neural networks

The computation equation for the input transformation layer and the output transformation layer is different depending on the specific engineering application. It can be coded using a computer language to automate the computation. The formulae of the 1st

Received 2002-12-17.

Biography: Hu Wusheng (1965—), male, doctor, professor, ws.hu@jlonline.com.

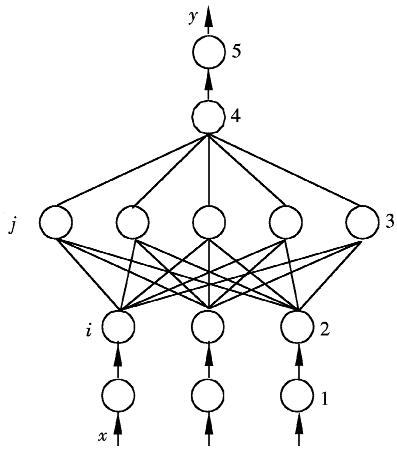

Fig.1 A single hidden layer model of BP networks. 1—Input transformation layer; 2—Input layer; 3—Hidden layer; 4—Output layer; 5—Output transformation layer

and the 5th layers are omitted.

Now, the formulae of the 2nd, the 3rd and the 4th layers are listed below[4].

1) The defined matrix

$$\boldsymbol{X}_{n\times1} = \{x_1, x_2, \cdots, x_n\}^{\mathrm{T}}$$

$$\boldsymbol{Y}_{m\times1} = \{y_1, y_2, \cdots, y_m\}^{\mathrm{T}}$$

$$\boldsymbol{W}_{\boldsymbol{A},n\times p} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}$$

$$\boldsymbol{W}_{\boldsymbol{B},p\times m} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ \vdots & \vdots & & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pm} \end{bmatrix}$$

where n is the number of units in the 2nd layer (i.e. the input layer); p is the number of units in the 3rd layer (i.e. the hidden layer); m is the number of units in the 4th layer (i.e. the output layer); $\boldsymbol{X}$ is the input vector of the 2nd layer; $\boldsymbol{Y}$ is the output vector of the 4th layer; $\boldsymbol{W_A}$ is the matrix of connecting weights between the 2nd and the 3rd layers; $\boldsymbol{W_B}$ is the matrix of connecting weights between the 3rd and the 4th layers.

2) The defined active functions

$$g(x) = x, \quad f(x) = \frac{1}{1 + \mathrm{e}^{-x}} \tag{3}$$

where $g(x)$ is the active function of the 2nd layer; $f(x)$ is the active function of the 3rd and the 4th layers, it is the Sigmoid standard active function.

3) Forward process

At first, initialize weights $\boldsymbol{W}_{ji}(0)$ of $\boldsymbol{W_A}$ and $\boldsymbol{W_B}$ with smaller non-zero random values.

In the 2nd layer, the input is $\boldsymbol{X}$, and the output is $g(\boldsymbol{X})$.

In the 3rd layer, the input is $\boldsymbol{A}$, and the output is $\boldsymbol{O_A}$.

$$\boldsymbol{A} = \boldsymbol{W}_{\boldsymbol{A}}^{\mathrm{T}} \cdot g(\boldsymbol{X}) = \boldsymbol{W}_{\boldsymbol{A}}^{\mathrm{T}} \cdot \boldsymbol{X}, \ \boldsymbol{O_A} = f(\boldsymbol{A}) \tag{4}$$

In the 4th layer, the input is $\boldsymbol{B}$, and the output is $\boldsymbol{Y}'$.

$$\boldsymbol{B} = \boldsymbol{W}_{\boldsymbol{B}}^{\mathrm{T}} \cdot \boldsymbol{O_A} = \boldsymbol{W}_{\boldsymbol{B}}^{\mathrm{T}} \cdot f(\boldsymbol{A}), \ \boldsymbol{Y}' = f(\boldsymbol{B}) \tag{5}$$

Then, the error of output can be calculated by

$$\Delta \boldsymbol{Y} = \boldsymbol{Y} - \boldsymbol{Y}' \tag{6}$$

where $\boldsymbol{Y}$ is the units expected output of the samples used for study; $\boldsymbol{Y}'$ is the units true output of the BP network. An objective function can be defined as

$$E = \frac{1}{2}\sum_{j=1}^{m}(y_j - y_j')^2 \tag{7}$$

4) Backward process

The aim of the backward process is to adjust all the coefficients in $\boldsymbol{W_A}$ and $\boldsymbol{W_B}$ based on the error matrix $\Delta\boldsymbol{Y}$.

From the 4th layer to the 3rd layer:

$$\delta B_j = -(y_j - y_j')f'(B_j) \qquad j = 1,2,\cdots,m \tag{8}$$

$$\left.\begin{aligned} &\Delta W_{B_{ij}} = \delta B_j O_{A_i} = \delta B_j f(A_i) \\ &W_{B_{ij}}^{(t+1)} = W_{B_{ij}}^{(t)} - \eta\,\Delta W_{B_{ij}}^{(t)} + \alpha\Delta W_{B_{ij}}^{(t-1)} \\ &\qquad i = 1,2,\cdots,p;\ j = 1,2,\cdots,m \end{aligned}\right\} \tag{9}$$

where η is training speed; α is the momentum coefficient; t is the cycle time[2,5].

From the 3rd layer to the 2nd layer:

$$\delta A_i = \sum_{j=1}^{m}(W_{B_{ij}}\delta B_j)f'(A_i) \quad i = 1,2,\cdots,p \tag{10}$$

$$\left.\begin{aligned} &\Delta W_{A_{ij}} = \delta A_j g(x_i) = \delta A_j x_i \\ &W_{A_{ij}}^{(t+1)} = W_{A_{ij}}^{(t)} - \eta\Delta W_{A_{ij}}^{(t)} + \alpha\Delta W_{A_{ij}}^{(t-1)} \\ &\qquad i = 1,2,\cdots,n;\ j = 1,2,\cdots,p \end{aligned}\right\} \tag{11}$$

Then, the forward process is reworked with the adjusted $\boldsymbol{W_A}$ and $\boldsymbol{W_B}$, and the iteration is to be continued until the E in Eq.(7) is less than one small value ε. Therefore, the term "train the BP network" means that all the coefficients in $\boldsymbol{W_A}$ and $\boldsymbol{W_B}$ are computed with the known data of samples by the BP network.

2 The Idea of CF&NNM

The procedure of the new method named CF&NNM is as follows:

1) Assume there are n points, of which n_1 point values of H_{GPS} and H_{Nor} are known, and n_2 ($n_2 = n - n_1$) point values of H_{Nor} need to be calculated. (Note: n_1 must be greater than 8.);

2) Based on n_1 points, elevation abnormality ξ can be calculated for all points by CFM;

3) Calculate the error of elevation abnormality of CFM for all n_1 points by

$$\Delta\xi = \xi_0 - \xi \tag{12}$$

where $\xi_0 = H_{GPS} - H_{Nor}$ means the known value of elevation abnormality; ξ means the elevation abnormality by CFM;

4) Use the information of above n_1 points ($x_i, y_i, \xi_i; \Delta\xi_i; i = 1,2,\cdots,n_1$) as a set of samples for study. The BP networks are trained by these samples. (Note: x_i, y_i, ξ_i are the three units in the input layer; $\Delta\xi_i$ is one unit in the output layer);

5) The error of elevation abnormality $\Delta\xi$ can be calculated by the trained BP networks for all n_2 points. The normal height can be calculated by

$$H_{Nor} = H_{GPS} - \xi_0 = H_{GPS} - (\xi + \Delta\xi) \tag{13}$$

where ξ is calculated by CFM and $\Delta\xi$ by NNM.

3 An Example

A city's D-order GPS network (about 300 km^2) has 96 observation points, among which 40 GPS points have third-order elevations obtained by geodetic leveling. The new method is tested with the 40 points for which values of H_{GPS} and H_{Nor} are known. In Fig.2, 10 evenly scattered points are selected as "a study group" to train the neural networks while the other 30 points form "a test group" used to check the effectiveness of the trained neural networks. Comparison of the three methods is discussed below.

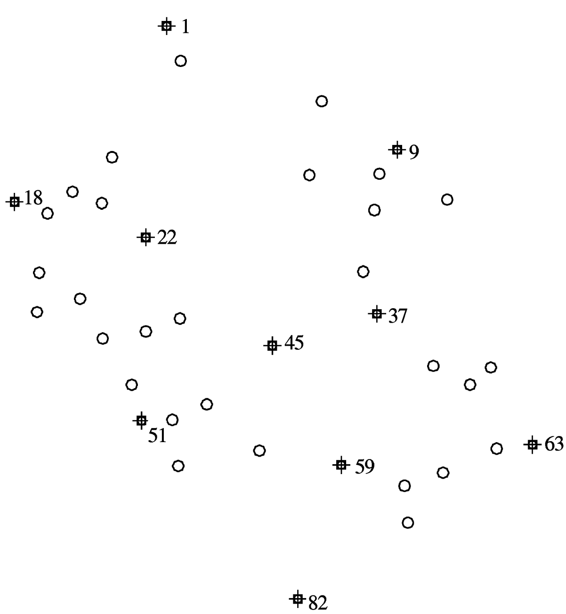

Fig.2 The 40 benchmarks in the D-order GPS network

3.1 CFM method

The unknown coefficients $a_0, a_1, \cdots, a_5$ in Eq.(2) can be obtained with the study group by the CFM, and then the elevation abnormality ξ of all 40 points can be computed by Eq.(2). The results of ξ are shown in Tab.1.

Tab.1 The results of ξ by CFM

Study group		Test group					
Point No.	$\Delta\xi$/mm	Point No.	$\Delta\xi$/mm	Point No.	$\Delta\xi$/mm	Point No.	$\Delta\xi$/mm
1	−1.7	14	−4.0	20	−5.7	6	−6.0
18	0.0	80	3.6	26	0.1	10	−5.3
82	3.5	31	−5.3	30	−5.7	15	−7.3
63	0.9	24	5.2	32	1.1	16	−8.5
9	0.6	48	−12.8	34	−0.7	17	−6.3
51	−3.5	69	−12.2	39	−9.3	62	−11.4
45	5.1	11	2.9	50	2.8	64	−6.9
37	−1.3	33	−2.1	52	−6.3	65	7.4
22	3.2	61	−10.0	54	−15.1	66	−3.3
59	−6.9	5	12.6	56	−10.8	210	11.5

Note: ① $n_1 = 10, m_1 = \pm 3.4$ mm; $n_2 = 30$, $m_2 = \pm 7.8$ mm; ② m_1 is the mean square error of the study group; ③ m_2 is the mean square error of the test group.

3.2 NNM method

After using NNM to convert H_{GPS} over one thousand tests, the ideal BP neural networks structure is the one that has two units (x, y) in the input layer, 15 units in the hidden layer and one unit ξ in the output layer. In the training process, the training mean square error of the study group is treated as the convergence standard. The process results of NNM can be seen in Tab.2.

Tab.2 The process results of NNM

Training mean square error of study group/mm	Cycle times	The mean square error of test group/mm
± 10.0	1 874	± 9.6
± 5.0	3 256	± 8.4
± 4.0	3 727	± 8.1
± 2.0	7 001	± 7.3
± 1.0	11 042	± 6.9

From Tab.2, it is shown that when the training mean square error of the study group is ± 1.0 mm, the neural networks precision can reach ± 6.9 mm, which is higher than that of CFM. The biggest shortcoming of NNM is that the results are unstable and the final results are largely influenced by the initial weights $W_{ji}(0)$ of $\boldsymbol{W_A}$ and $\boldsymbol{W_B}$.

3.3 CF&NNM method

The training mean square error of the study group is taken as the convergence standard in the CF&NNM method. The neural network of CF&NNM is trained for more than one hundred tests with different initial weights $W_{ji}(0)$. The difference of every two tests' results is very small, and in all cases, the computation converged. The process results of CF&NNM are listed in Tab.3.

From Tab.3, we can find that: ① When the training mean square error of the study group is ± 2.0

mm, the mean square error of the test group reaches the minimum value of ± 5.5 mm; ② When the training mean square error is less than ± 2.0 mm, the smaller the training mean square error of the study group is, the bigger the mean square error of the test group is.

Tab.3 The process results of CF&NNM

Training mean square error/mm	Cycle times	The mean square error of test group/mm
± 2.8	491	± 7.6
± 2.4	703	± 6.4
± 2.0	4 978	± 5.5
± 1.6	6 015	± 5.8
± 1.2	6 700	± 5.9
± 0.8	7 632	± 6.1
± 0.4	8 821	± 6.2

3.4 Theoretical analysis of CF&NNM

Now, let us analyze the second conclusion in detail.

In fact, the CF&NNM method constructed in this paper can be used for detecting the model error of the CFM with the help of the neural networks. This can be explained by the BP networks structure of CF&NNM: among the parameters of the input layer of the BP networks, there is one parameter ξ, which is the result of the elevation abnormality by CFM, and the parameter of the output layer is the difference of elevation abnormality $\Delta\xi = \xi_0 - \xi$, which is the difference between the elevation abnormality ξ by CFM and its true value ξ_0. Some conclusions can be drawn based on Tab.1 and Tab.3. By the CFM, the mean square error of the study group is about ± 3.4 mm, of which, about 40%, i.e., about ± 1.4 mm, is the model error of CFM, and the remaining (about ± 2.0 mm) is the observation error. So, when the neural network's training error reaches ± 2.0 mm by CF&NNM, the model error of CFM is removed. Here the mean square error of the test group reaches the minimum, i.e., about ± 5.5 mm, which accounts for 70% of the mean square error (± 7.8 mm) of the test group by CFM. The new method is very effective. Now if one continues to reduce the training mean error of the study group, the results by CF&NNM will be worse, due to taking the observation error as the model error.

Well then, how much should the training mean square error of the study group acting as the convergence standard be taken? In practical applications, when it is difficult to determine the ration between model error and observation error, experiences should be consulted.

When the training mean square error of the study group is ± 2.0 mm, the results of ξ by CF&NNM can be seen in Tab.4. The equation of the difference of elevation abnormality in Tab.4 is as

$$\Delta\xi = \xi_0 - \xi'_0 \tag{14}$$

where ξ_0 is the known elevation abnormality; ξ'_0 is the elevation abnormality calculated by CF&NNM.

Tab.4 The results of ξ by CF&NNM

Study group		Test group					
Point No.	$\Delta\xi$/mm	Point No.	$\Delta\xi$/mm	Point No.	$\Delta\xi$/mm	Point No.	$\Delta\xi$/mm
1	− 0.7	14	− 4.3	20	− 4.9	6	− 6.4
18	1.8	80	1.2	26	− 4.7	10	8.2
82	1.6	31	− 4.4	30	− 6.2	15	− 7.7
63	0.9	24	0.9	32	− 0.6	16	− 8.5
9	− 2.2	48	− 9.6	34	− 4.2	17	− 5.8
51	− 3.0	69	− 2.0	39	0.7	62	− 1.8
45	1.6	11	0.7	50	0.2	64	− 6.0
37	0.9	33	− 5.1	52	− 2.5	65	8.6
22	2.9	61	− 0.6	54	− 4.8	66	− 1.1
59	− 2.7	5	11.5	56	− 0.6	210	7.7

Note: ① $n_1 = 10$, $m_1 = \pm 2.0$ mm; $n_2 = 30$, $m_2 = \pm 5.5$ mm; ② m_1 is the mean square error of the study group; ③ m_2 is the mean square error of the test group.

Comparing Tab.4 and Tab.1, it is obvious that the CF&NNM method can produce better results of ξ than the CFM method. For example, the total number of points, whose difference of elevation abnormality is larger than 10 mm, is 8 by CFM, and is only 1 by CF&NNM.

4 Conclusion

Compared with CFM and NNM, the CF&NNM can produce more accurate results in converting GPS height H_{GPS} into normal height H_{Nor}. More work should be done to evaluate its effectiveness for larger engineering projects with more complicated topography. The disadvantage of the method based on neural networks is that the result is unstable and the final results are largely influenced by the initial weight $W_{ji}(0)$. But the BP structure of the CF&NNM discussed in this paper is stable in calculating $\Delta\xi$ (the error of the elevation abnormality) with no effect on the results by the initial weight $W_{ji}(0)$. Thus the CF&NNM is recommended for use in certain engineering projects.

References

[1] Liu Dajie, Shi Yimin, Guo Jingjun. *The theory of GPS and its application*[M]. Shanghai: Tongji University Press, 1996.224 – 229. (in Chinese)

[2] Yuan Zengren. *Artificial neural networks and their applications* [M]. Beijing: Tsinghua University Press, 1999. 66 – 130. (in Chinese)

[3] Yang Mingqing, Jin Fan, Zhu Dacheng, et al. Conversion of

GPS height by artificial neural network method [J]. *Acta Geodaetica et Cartogra Phica Sinica*, 1999, **28**(4):301 - 307. (in Chinese)

[4] Hu Wusheng. *The research of the theory and the method and the application of precise GPS height survey*[D]. Nanjing: Hohai University, 2001.33 - 78. (in Chinese)

[5] Leonard J, Kramer M A. Improvement of the back propagation algorithm for training neural networks [J]. *Computers Chem Engng*, 1990, **14**(3):337 - 341.

基于神经网络的转换 GPS 高程的新方法

胡伍生　　沙月进

(东南大学交通学院测绘工程系,南京 210096)

摘　要　GPS 高程是相对于 WGS - 84 椭球体的大地高,因此,在工程应用中,GPS 高程需要转换为正常高.转换 GPS 高程通常采用二次曲面拟合法(CFM)和神经网络方法(NNM),但这 2 种方法各有优缺点.在研究了这 2 种方法之后,提出了一种转换 GPS 高程的新方法,该方法综合了上述 2 种方法的优点,故取名为“CF&NNM”方法.介绍了 CF&NNM 方法的思路和计算过程.通过一个工程实例,列出了上述 3 种方法的数据处理结果,新方法效果最好.对 CF&NNM 方法进行了理论分析.

关键词　神经网络; BP 算法; GPS 高程; CF&NNM 方法

中图分类号　P228