Application of a Back-Propagation Artificial Neural Network to Regional Grid-Based Geoid Model Generation Using GPS and Leveling Data

Lao-Sheng Lin[1]

Abstract: The height difference between the ellipsoidal height h and the orthometric height H is called undulation N. The key issue in transforming the global positioning system (GPS)-derived ellipsoidal height to the orthometric height is to determine the undulation value accurately. If the undulation N for a point whose position is determined by a GPS receiver can be estimated in the field, then the GPS-derived three-dimensional geocentric coordinate in WGS-84 can be transformed into a local coordinate system and the orthometric height in real-time. In this paper, algorithms of applying a back-propagation artificial neural network (BP ANN) to develop a regional grid-based geoid model using GPS data (e.g., ellipsoidal height) and geodetic leveling data (e.g., orthometric height) are proposed. In brief, the proposed algorithms include the following steps: (1) establish the functional relationship between the point's plane coordinates and its undulation using the BP ANN according to the measured GPS data and leveling data; (2) develop a regional grid-based geoid model using the imaginary grid plane coordinates with a fixed grid interval and the trained BP ANN; (3) develop an undulation interpolation algorithm to estimate a specific point's undulation using the generated grid-based geoid model; and (4) estimate the point's undulation in the field and transform the GPS ellipsoidal height into the orthometric height in real-time. Three data sets from the Taiwan region are used to test the proposed algorithms. The test results show that the undulation interpolation estimation accuracy using the generated grid-based geoid is in the order of 2–4 cm. The proposed algorithms and the detailed test results are presented in this paper.

DOI: 10.1061/(ASCE)0733-9453(2007)133:2(81)

CE Database subject headings: Geoid; Global positioning; Height; Neural networks; Surveys.

Introduction

The coordinate system of the global positioning system (GPS) is the world geodetic system of 1984 (WGS-84). The positions determined by GPS receivers are expressed in geocentric coordinates or geodetic coordinates defined by a WGS-84 ellipsoid, but in engineering applications, these coordinates need to be transformed into a local plane coordinate system (such as Northing and Easting), and ellipsoidal (geodetic) heights (h) need to be transformed into physical heights, such as orthometric heights (H). The height difference between the ellipsoidal height h and the orthometric height H is called undulation (geoid height) N. If a point's position is determined by a GPS receiver and its undulation N is available, then its three-dimensional geocentric coordinate in WGS-84 can accordingly be transformed into a local plane coordinate system and the orthometric height, since the orthometric height H is the difference between its geodetic height h and the undulation N (Hu et al. 2002, 2004; Kavzoglu and Saka 2005; Kuhar et al. 2001; Stopar et al. 2006; Yang and Chen 1999; Zaletnyik et al. 2004).

In the Taiwan region, a new national vertical datum, Taiwan Vertical Datum 2001, was established between 2000 and 2003. The datum was established using the observations of geodetic leveling, the GPS, and gravity collected at 2,065 newly established benchmarks within a network of 4,500 km first-order leveling lines. These 2,065 benchmarks are well distributed around Taiwan and their point-to-point distances are about 2 km. Hence, each benchmark has two types of heights, namely orthometric height H and ellipsoidal height h. The orthometric height H accuracies from geodetic leveling are at the level of ±8.8 mm and the ellipsoidal height h accuracies from GPS surveying are at the level of ±36 mm (Chen et al. 2004; Yang et al. 2003).

Nowadays, the gravimetric method is the most commonly used technique for precise determination of the geoid. The necessary condition for its use is existence of high resolution gravity data set. With the lack of gravity data, the geoid determination is possible by means of various geometric methods, the astrogeodetic method and the method of determining the undulations from GPS in combination with geometric leveling (Stopar et al. 2006). As the undulations of those 2,065 benchmarks of the Taiwan region can be calculated from the observed GPS and leveling data, it is possible to generate a regional geoid model using these calculated undulations. Further, geodetic leveling work is time consuming, whereas GPS surveying work is relatively time saving. Hence, if a regional geoid model can be generated from these 2,065 benchmarks with an adequate degree of accuracy, then by integrating the generated geoid model and the GPS, it is possible to transform the ellipsoidal height h from the GPS to the orthometric height H in an office or in the field in real-time.

In general, the shape of the geoid is very complex and the task of approximating the geoid surface by a relatively simple math-

[1]Associate Professor, Dept. of Land Economics, National Chengchi Univ., 64, Section 2, Chihnan Rd., Taipei 116, Taiwan.

Note. Discussion open until October 1, 2007. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on April 24, 2006; approved on October 24, 2006. This paper is part of the ***Journal of Surveying Engineering***, Vol. 133, No. 2, May 1, 2007. ©ASCE, ISSN 0733-9453/2007/2-81–89/$25.00.

ematical expression is hardly easy. For the local geoid approximation various interpolation surfaces may be used, e.g., low-order polynomials. Alternatively, artificial neural network (ANN) used as a geoid approximator was introduced by Ambrozic et al. in 1999 (Stopar et al. 2006). Since then, there are many studies on geoid approximation using ANN (Hu et al. 2002, 2004; Kavzoglu and Saka 2005; Kuhar et al. 2001), or using ANN and least squares collocation (Stopar et al. 2006). According to their study results, ANNs may efficiently approximate real undulation on the chosen area.

In order to fulfill the engineering application requirement for the orthometric height and take advantage of GPS surveying, algorithms to generate a Taiwan regional grid-based geoid model using the GPS and leveling data were proposed by the writer. In this paper, the concept and methodology of the proposed grid-based geoid model generating algorithm are first introduced. Then, test results from the data of (1) 2,065 benchmarks around the Taiwan region; (2) 78 benchmarks in Taichung City (in the central part of Taiwan); and (3) 9 check points over the campus of National Chengchi University are presented to demonstrate the performance of the proposed grid-based geoid model generating algorithms.

Application of a Back-Propagation Artificial Neural Network to Regional Grid-Based Geoid Model Generation Using GPS and Leveling Data

Undulation Calculation Using GPS and Leveling Data

Let h and H denote, respectively, the GPS-derived ellipsoidal height and the geodetic leveling derived orthometric height at a specific point. It should be noted that the ellipsoidal height h is referred to as the surface of the WGS-84 ellipsoid, while the orthometric height H is referred to as the surface of the local geoid (the mean sea level). Usually, these two surfaces do not coincide from point to point due to the fact that the physical earth is different from the mathematical earth (ellipsoid). Hence, the height difference between the ellipsoidal height h and the orthometric height H is called undulation N. The following equation establishes the relationship between h and H (Hu et al. 2004; Kavzoglu and Saka 2005; Kuhar et al. 2001; Stopar et al. 2006; Yang and Chen 1999; Zaletnyik et al. 2004):

$$h = H + N \tag{1}$$

The undulation at this point can be calculated by the following equation:

$$N = h - H \tag{2}$$

Artificial Neural Networks

ANNs are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. The network is adjusted, based on a comparison of the output and the target, until the network output matches the target. Typically, many such input/target pairs are used, in this supervised learning, to train a network (Demuth and Beale 2002; Hu et al. 2004; Kavzoglu and Saka 2005; Kuhar et al. 2001; Stopar et al. 2006; Zaletnyik et al. 2004).

A neuron with a single scalar input p and a scalar bias b is shown in Fig. 1. The scalar input p is transmitted through a con-

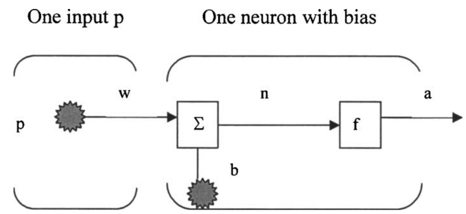

Fig. 1. A neuron with a single input p and a scalar bias b

nection that multiplies its strength by the scalar weight w to form the product wp. Then, the bias b is added to the product wp as shown by the summing junction. The bias is much like a weight. The transfer function net input n, again a scalar, is the sum of the weighted input wp and the bias b. The sum is the argument of the transfer function f. The scalar output of f is a, $a = f(wp + b)$. The central idea of ANN is that such parameters, w and b, can be adjusted so that the network exhibits some desired or interesting behavior (Demuth and Beale 2002).

There are many transfer functions used in ANN. Two of the most commonly used transfer functions for back-propagation ANN are tan-sigmoid transfer function (tansig) and linear transfer function (purelin) (Demuth and Beale 2002). In back-propagation it is important to be able to calculate the derivatives of any transfer functions used. Each of the transfer functions, tansig and purelin, can be called to calculate its own derivative.

Two or more of the neurons can be combined in a neuron layer, and a particular network could contain one or more such layers. There are no connections among neurons in the same layer, whereas every two neurons in neighboring layers are connected. The neuron layer includes the weight matrix, the multiplication operations, the bias vector b, the summer, and the transfer functions.

A three-layer network with input layer, one hidden layer, and output layer is shown in Fig. 2. The network shown in Fig. 2 has two-element input, two neurons in the hidden layer, and one neuron in the output layer. It is common for different layers to have a different number of neurons. The weight matrices being connected to inputs are defined as input weights, such as IW. And the weight matrices coming from neuron layer outputs are defined as layer matrices, such as LW. Further, superscripts are used to iden-

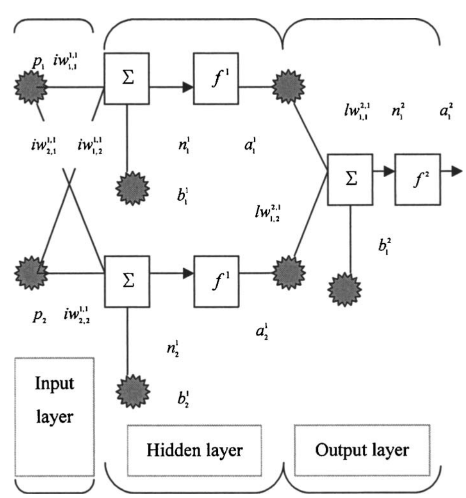

Fig. 2. A three-layer network with two input elements, two neurons, and one output

J. Surv. Eng. 2007.133:81-89.

tify the source (second index) and the destination (first index) for the various weights and other elements of the network. Note that the outputs of each intermediate layer are the inputs to the following layer. Thus the output layer can be analyzed as a two-layer network with two inputs (i.e. a_1^1 and a_2^1), one neuron, and a 1×2 weight matrix LW. The output of hidden layer is: $a^1 = f^1(\mathrm{IW}^{1,1}p + b^1)$. And the output of output layer is: $a^2 = f^2(\mathrm{LW}^{2,1}a^1 + b^2) = f^2(\mathrm{LW}^{2,1}f^1(\mathrm{IW}^{1,1}p + b^1) + b^2)$.

Neural network training can be made more efficient if certain preprocessing steps are performed on the inputs and targets of network. Before training, it is often useful to scale the inputs and targets so that they always fall within a specific range. The function premnmx can be used to scale inputs and targets so that they fall in the range $[-1,1]$. If premnmx is used to scale the inputs and targets, then the output of the network will be trained to produce outputs in the range $[-1,1]$. If one wants to convert these outputs back into the same units that were used for the original target, then one should use the routine postmnmx (Demuth and Beale 2002).

ANN used as a geoid approximator was introduced by Ambrozic et al. in 1999 (Stopar et al. 2006). Since then, there are many studies on geoid approximation using ANN (Hu et al. 2002, 2004; Kavzoglu and Saka 2005; Kuhar et al. 2001; Stopar et al. 2006; Zaletnyik et al. 2004). According to their study results, ANNs may efficiently approximate real undulation (geoid height) on the chosen area. According to Kavzoglu and Saka (2005), the ANN-based (geoid) surfaces seem to be the low deviations from the GPS/leveling data surface. Hence, in this paper, a method using ANN for approximation of the Taiwan geoid surface is employed. On the other hand, the use of ANNs is complicated, basically due to problems encountered in their design and implementation. From the design perspective, the specification of the number and size of the hidden layer(s) is critical for the network's capability to learn and generalize. A further difficulty in the use of ANNs is the choice of appropriate values for network parameters that have a major influence on the performance of the learning algorithm. It is often the case that a number of experiments are required to ascertain the selection of the parameter values that give the highest accuracy. A trial-and-error strategy is frequently used to determine appropriate values for these parameters (Kavzoglu and Saka 2005; Kuhar et al. 2001; Stopar et al. 2006).

Back-Propagation Artificial Neural Networks

Back-propagation (BP) was created by generalizing the Widrow-Hoff learning rule to multiple-layer networks and nonlinear differentiable transfer functions. The architecture of back-propagation artificial neural network (BP ANN) is the multilayer feedforward network such as Fig. 2. Feedforward networks often have one input layer, one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons.

Once the BP ANN weights and biases are initialized, the network is ready for training. The training process requires a set of examples of proper network behavior—network inputs p and target outputs t. During training the weights and biases of the network are iteratively adjusted to minimize the network performance function. The default performance function for feedforward networks is mean square error—the average squared error between the network outputs a and target outputs t. There are many variations of the back-propagation algorithm. All these algorithms use the gradient of the performance function to determine how to adjust the weights to minimize performance. The gradient is determined using a technique called back-propagation, which involves performing computations backward through the network. The back-propagation computation is derived using the chain rule of calculus (Demuth and Beale 2002; Hu et al. 2002, 2004; Kavzoglu and Saka 2005; Kuhar et al. 2001; Stopar et al. 2006).

According to Demuth and Beale (2002), a three-layer back-propagation artificial neural network (BP ANN), i.e., input layer, hidden layer, and output layer, can be used as a general approximator. It can approximate any function with a finite number of discontinuities arbitrarily, given sufficient neurons in the hidden layer. From Kavzoglu and Saka (2005), the BP learning algorithm has been used in about 70% of all ANN applications. Besides, the BP learning algorithm has been used in recent studies on geoid approximation using ANN (Hu et al. 2002, 2004; Kavzoglu and Saka 2005; Kuhar et al. 2001; Stopar et al. 2006). Hence, in this paper the BP ANN was selected to model Taiwan geoid.

MATLAB Artificial Neural Network Toolbox

MATLAB (The Mathworks Inc., Natick, Mass.) is a high-performance language for technical computing. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. Of great importance to most users of MATLAB, toolboxes allow one to learn and apply specialized technology. Toolboxes are comprehensive collections of MATLAB functions (M-files) that extend the MATLAB environment to solve particular classes of problems. Areas in which toolboxes are available include signal processing, control systems, neural networks, fuzzy logic, wavelets, simulation, and many others. There are generally four steps in the training process using the back-propagation artificial neural network toolbox: (1) assemble the training data; (2) create the network object; (3) train the network; and (4) simulate the network response to new inputs (Demuth and Beale 2002). It should be noted that the artificial neural network toolbox is applied in the development of the proposed algorithms.

Application of a Back-Propagation Artificial Neural Network to Undulation Estimation

If a point's coordinates are determined by the GPS, then its ellipsoidal height h can be computed accordingly. Also, its orthometric height H is determined by the geodetic leveling method, so its undulation N can be calculated by Eq. (2). The points with known ellipsoidal height h and orthometric height H are defined as reference points.

Suppose there are n reference points in a region of interest. We can use the calculated undulations from those reference points to generate a local geoid model, applying some kinds of geoid model generating algorithms, such as the curve fitting method or artificial neural networks. The undulations at any other points can be predicted from the generated geoid model using an interpolation technique.

Taking the curve fitting method as an example, suppose there are n reference points with known undulations in one area. A polynomial surface can be used to fit these known undulations, and the coefficient terms of the polynomial can be found by the least squares adjustment method. The following equation is an example of the curve fitting method (Hu et al. 2002, 2004):

J. Surv. Eng. 2007.133:81-89.

$$N(x,y) = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 \quad (3)$$

where (x,y)=plane coordinates of a point; $N(x,y)$=corresponding undulation to point (x,y); and $a_0 \cdots a_5$=polynomial coefficients. For a new point $p(x,y)$ determined by the GPS, its undulation N can be determined by Eq. (3) if these polynomial coefficients are known. Then, its orthometric height H can be computed by the following equation:

$$H = h - N \quad (4)$$

On the other hand, an artificial neural network can also be applied to regional geoid model generation (Hu et al. 2002; 2004; Kavzoglu and Saka 2005; Kuhar et al. 2001; Stopar et al. 2006; Zaletnyik et al. 2004). Suppose there are n reference points in a specific region. The reference point set $P=\{P_1,P_2,\ldots,P_n\}$ can be used to train the BP ANN

$$P_i = (x_i, y_i, N_i), \quad i = 1,2,\ldots,n \quad (5)$$

where (x_i,y_i)=plane coordinates of reference point i; N_i =corresponding undulation to reference point (x_i,y_i); and i indicates the reference point number. It should be noted that a three-layer BP ANN, with one input layer, one hidden layer, and one output layer, was adopted in this paper to generate a regional geoid model. Hence, the input vector of the BP ANN consists of (x_i,y_i) and the output vector consists of N_i.

After being trained by the reference point set $P=\{P_1,P_2,\ldots,P_n\}$, the BP ANN establishes the functional relationship between input layer (x_i,y_i) and output layer N_i

$$N_i = F(x_i, y_i), \quad i = 1,2,\ldots,n \quad (6)$$

where F=function, which associates input vectors (x_i,y_i) with specific output vectors N_i.

It should be noted that the main function of F is similar to that of the polynomial coefficients of Eq. (3). However, the main function of F is described implicitly by the neurons in the hidden layer of the BP ANN. Hence, the better way would be "to store" the "trained BP ANN" and to compute the undulation of any point directly from the trained BP ANN. On the other hand, an alternative approach is to apply the generated regional grid-based geoid model using the following proposed algorithms.

Regional Grid-Based Geoid Model Generation Procedure

In order to fulfill the engineering application requirement for the orthometric height and take advantage of GPS surveying, algorithms to generate a Taiwan regional grid-based geoid model using GPS and leveling data were proposed. The procedure of the proposed algorithms is summarized in the following.

Step 1: Establish the functional relationship between the point's plane coordinates and its undulation using a back-propagation artificial neural network according to the observed GPS and leveling data. The observed GPS and leveling data of the 2,065 benchmarks, i.e., the reference points, are used to calculate their undulations using Eq. (2). The plane coordinates (x_i,y_i) and the calculated undulations N_i of the 2,065 reference points are treated as input vectors and output vectors, respectively. A three-layer back-propagation artificial neural network is trained using the above-mentioned input vectors and output vectors. The transfer functions for the hidden layer and the output layer are tansig (hyperbolic tangent sigmoid transfer function) and purelin (linear transfer function) respectively (Demuth and Beale 2002). After the back-propagation artificial neural network is trained successfully, the functional relationship between the point's plane coordinates and its undulation, such as Eq. (6), is established.

Step 2: Develop a regional grid-based geoid model using the imaginary grid plane coordinates with a fixed grid interval and the trained back-propagation artificial neural network. At this step, an imaginary fixed grid (e.g., 1,000 m in Easting and Northing) on the surface of the ellipsoid over the Taiwan region is generated. Assuming that the grid interval and the plane coordinates of the left-bottom point of the imaginary grid are given, the plane coordinates (x_i,y_i) of the imaginary grid nodes can be calculated and stored in a file. Once, the back-propagation artificial neural network has been trained using the reference points of the input and output vectors, the plane coordinates of the imaginary grid nodes will be treated as input vectors. The undulations of all grid nodes will be estimated using Eq. (6) via the trained back-propagation artificial neural network. Then, the results of this step are included in a file consisting of the plane coordinates (x_i,y_i) and the estimated undulations N_i of all grid nodes. Finally, the regional grid-based geoid model is generated.

Undulation Interpolation Procedure

Once the regional grid-based geoid model is generated, it can be used to interpolate a specific point's undulation if the point's plane coordinates are given. The procedure is summarized in the following.

Step 1: Develop an undulation interpolation algorithm to estimate a specific point's undulation using the generated grid-based geoid model. If the grid-based geoid model has been generated, then the undulation of any specific point can be computed using the following weighting function approach (Junkins et al. 1973; Lin 1998)

$$N^p(x_p,y_p) = \sum_{i=1}^{4} W_i(\delta x_p, \delta y_p) N^i \quad (7)$$

where (x_p,y_p)=plane coordinates of any specific point p; N^p=estimated undulation of a point with coordinates (x_p,y_p); $i=1,2,3,4$ represents the sequence of four nodes of the cell, including the point with coordinates (x_p,y_p), starting from the right-top node, and in the counterclockwise direction; and N^i=estimated undulation value of node i from the trained BP ANN. The general equation for the weighing function is (Junkins et al. 1973)

$$W(\delta x_p, \delta y_p) = \delta x_p^2 \delta y_p^2 (9 - 6\delta x_p - 6\delta y_p + 4\delta x_p \delta y_p) \quad (8)$$

The other quantities are

$$W_1(\delta x_p, \delta y_p) = W(\delta x_p, \delta y_p) \quad (9)$$

$$W_2(\delta x_p, \delta y_p) = W(1 - \delta x_p, \delta y_p) \quad (10)$$

$$W_3(\delta x_p, \delta y_p) = W(1 - \delta x_p, 1 - \delta y_p) \quad (11)$$

$$W_4(\delta x_p, \delta y_p) = W(\delta x_p, 1 - \delta y_p) \quad (12)$$

$$\Delta x_p = x_p - x_1 \quad (13)$$

$$\Delta y_p = y_p - y_1 \quad (14)$$

$$\delta x_p = \frac{\Delta x_p}{x_2 - x_1} \quad (15)$$

J. Surv. Eng. 2007.133:81-89.

$$\delta y_p = \frac{\Delta y_p}{y_2 - y_1} \tag{16}$$

where the plane coordinates of any specific point p are (x_p, y_p); (x_1, y_1)=coordinates of the left-bottom node of the cell, namely Node 3; and (x_2, y_2)=coordinates of the right-top node of the cell, namely Node 1.

Step 2: Estimate the undulation in the field of the point of interest and transform the GPS ellipsoidal height into the orthometric height in real time. According to Step 1, the undulation interpolation algorithm can be developed into a computer program using certain kinds of computer languages, such as C++ (Microsoft Corporation, Redmond, WA), MATLAB, and BASIC (Microsoft Corporation, Redmond, WA). Then, after installing the developed interpolation program and the generated grid-based geoid model on a computer, such as a desktop personal computer (PC) or a notebook, it is easy to estimate the undulation of the point of interest. However, if a pocket PC is used together with a GPS receiver in the field, then the interpolation program will develop differently since the operation system of a pocket PC is different from that of a PC. Further, the hard disk capacity of a pocket PC is much smaller than that of a PC. Therefore, a PC-based program is not executable on a pocket PC system. One of the solutions to this problem is to develop the interpolation program using a program developing package, such as Microsoft (Microsoft Corporation, Redmond, Wash.) Studio Net 2003, and then transform the interpolation program into one suitable for a pocket PC (Deitel et al. 2004).

Test Results and Discussion

Test Data

Three data sets were used to test the proposed algorithms in this paper. The first data set, which included the GPS and leveling data of the 2,065 benchmarks of the Taiwan region, was collected between 2000 and 2003 by the Satellite Survey Center, Department of Land Administration, Ministry of Interior, Taiwan. The test region size is about 373 km (north–south) by 140 km (east–west). The distance between two consecutive benchmarks is about 2 km. The GPS data were collected by the static GPS surveying method in an observation session lasting from 2 to 3 h, whereas the leveling data were collected by the first-order geodetic leveling method. The orthometric height H accuracies from the geodetic leveling are at the level of ±8.8 mm and the ellipsoidal height h accuracies from the GPS surveying are at the level of ±36 mm (Chen et al. 2004; Yang et al. 2003).

The point distribution map of the 2,065 benchmarks (Data Set 1) is shown in Fig. 3. From Fig. 3, it can be seen that there were several areas with no benchmarks. The main reason for this is that Taiwan is an island with small plains in the west and east, and mountains in the central region. Therefore, in the central mountain region, the geodetic leveling surveying was carried out only on several highways going from east to west. According to the observed ellipsoidal heights h and the orthometric heights H of the 2,065 benchmarks, the undulations of these 2,065 benchmarks could be calculated using Eq. (2). The maximum, minimum, and average calculated undulations of the 2,065 benchmarks were +27.891, +18.776, and +21.841 m, respectively. The contour map of the calculated undulations of the 2,065 benchmarks is shown in Fig. 4. From Fig. 4, it can be seen that the minimum calculated undulation appeared in the west part of Taiwan, and the maximum calculated undulation appeared in the central mountain region.

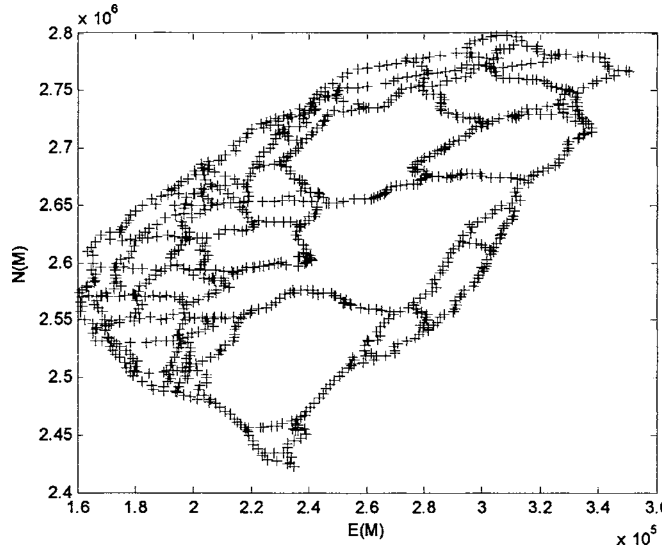

Fig. 3. Point distribution map of the 2,065 benchmarks of the Taiwan region, "+" represents the location of the benchmark

The second data set, including GPS and leveling data of 78 benchmarks, was collected by Taichung City government in 2004. The test region size is about 14 km (north–south) by 25 km (east–west). The distance between two consecutive benchmarks is about 0.3–2 km. The GPS data were collected by the static GPS surveying method in an observation session lasting from 2 to 3 h, whereas the leveling data were collected by the first-order geodetic leveling method. The maximum, minimum, and average calculated undulations of the 78 benchmarks +20.818, +19.213, +19.798 m, respectively.

The third data set, including nine check points, was collected from the campus of National Chengchi University (NCCU). The NCCU test region is about 400 m by 850 m. The ellipsoidal heights of these check points were derived from the Real Time Kinematics (RTK) GPS Survey, whereas the orthometric heights of these check points were observed by a three-order geodetic leveling survey.

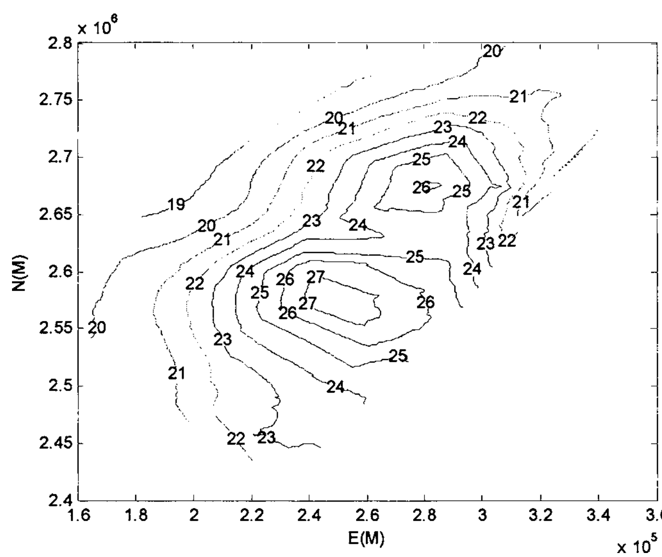

Fig. 4. Contour map of the calculated undulations (m) of the 2,065 benchmarks of the Taiwan region

J. Surv. Eng. 2007.133:81-89.

Data Processing

In order to test the proposed grid-based geoid model generating algorithms, a set of software was developed and revised based on the above-mentioned procedure. The artificial neural network program was developed using the MATLAB artificial neural network toolbox. A three-layer BP ANN, with one input layer, one hidden layer, and one output layer, was adopted to establish the functional relationship between the reference points' plane coordinates and their calculated undulations. The input vector consisted of plane coordinates (x_i, y_i) of each reference point, while the output vector consisted of each reference point's undulation (N_i). The transfer functions for the hidden layer and the output layer were tansig (hyperbolic tangent sigmoid transfer function) and purelin (linear transfer function) respectively (Demuth and Beale 2002).

The undulation interpolation programs were developed using different computer languages. For the desktop PC, C++ was used and for the pocket PC, Microsoft Studio Net 2003 (Deitel et al. 2004) was used.

In order to test the accuracies of the proposed algorithms, some points were used to train the artificial neural network, and others were used to evaluate the performance of the proposed algorithms. The points used to train the artificial neural network were defined as reference points, while the other points used to evaluate the performance of the proposed algorithms were defined as check points. The ellipsoidal height h and the orthometric height H of both the reference points and the check points were known.

The term "ΔN" shown in Tables 1–6 is defined as:

$$\Delta N_i = N_i^{\text{calculated}} - N_i^{\text{estimated}}, \quad i = 1,2,\ldots,n \tag{17}$$

where $N_i^{\text{calculated}}$=calculated undulation N of check point i from its known ellipsoidal height h and orthometric height H; $N_i^{\text{estimated}}$ =either estimated undulation N using the trained BP ANN or interpolated undulation using the generated grid-based geoid model; and ΔN_i=undulation difference between the calculated undulation and the estimated undulation. "$\sigma_{\Delta N}$" indicates the standard deviation of all ΔN_i. "Maximum ΔN" denotes the maximum value of all ΔN values. "Minimum ΔN" denotes the minimum value of all ΔN values. "Iteration number" indicates the epochs taken when the ANN is trained successfully (Demuth and Beale 2002).

Performance Evaluation of the Application of a Back-Propagation Artificial Neural Network to Undulation Estimation Using GPS and Leveling Data

In order to evaluate the performance of undulation estimation when applying a BP ANN, a series of data on 283 benchmarks—part of Data Set 1—was used to test the developed MATLAB BP ANN program. The tested items included: (1) the undulation estimation accuracies versus the training algorithms adopted in the BP ANN program; (2) the undulation estimation accuracies versus the amount of neurons in the hidden layer chosen in the BP ANN program; and (3) the undulation estimation accuracy comparisons with other undulation estimation approaches. The test data, including the 283 benchmarks, were collected from the central part of Taiwan. The test region is about 116 km (east–west) by 103 km (north–south). The distance between two consecutive benchmarks is about 2 km. These 283 benchmarks were divided into two groups. One group, which included 142 points, were treated as reference points to train the BP ANN, whereas the other group, which included 141 points, were treated as check points to evaluate the estimation accuracy of the BP ANN.

Table 1. Performance Statistics of the BP ANN with Varied Training Algorithms

Training algorithm	Iteration number (epochs)	$\sigma_{\Delta N}$ (cm)	Maximum ΔN (cm)	Minimum ΔN (cm)
TRAINBR	1,606	±4.00	21.42	−10.49
TRAINLM	17,785	±5.28	8.94	−48.03
TRAINCGF	489	±5.58	16.22	−17.43
TRAINGDX	50,000	±4.18	9.58	−9.00

In order to test Item 1, first, the amount of neurons in the hidden layer was fixed to 15, and the training algorithms were changed accordingly. The test results are shown in Table 1. The "training algorithm" indicates the adopted BP ANN training algorithm. Four different training algorithms were tested: "trainbr" (Bayesian regularization back-propagation), "trainlm" (Levenberg-Marquardt back-propagation), "traincgf" (conjugate gradient back-propagation with Fletcher-Reeves updates), and "traingdx" (gradient descent with momentum and adaptive learning rate back-propagation) (Demuth and Beale 2002). From Table 1, it can be seen that the accuracy,$\sigma_{\Delta N}$, when using the "trainbr" algorithm is ±4.00 cm, the minimal value of all cases, and its iteration number is 1,606. Hence, after considering the accuracy of the BP ANN, the training algorithm trainbr was adopted in the following tests.

In order to test Item 2, the amount of neurons in the hidden layer was changed from 5 to 50. The test results are summarized in Table 2. From Table 2, it can be seen that the value of $\sigma_{\Delta N}$ decreases from ±5.08 to ±3.60 cm when the amount of neurons increases from 5 to 35. If the amount of neurons in the hidden layer is larger than 35, then the values of $\sigma_{\Delta N}$ increase again. Hence, the best number of neurons in the hidden layer is 35 in this case.

In order to test Item 3, the same data set was tested on other undulation estimation approaches, such as (1) the curve fitting method that simulates the undulation surface using a quadratic polynomial; and (2) the undulation estimation model of the Ministry of Interior, Taiwan. The test results are summarized in Table 3. "BP ANN" denotes that BP ANN, with the trainbr training algorithm and 35 neurons in the hidden layer, was used to estimate the check point's undulation. "Curve fitting" indicates

Table 2. Performance Statistics of the BP ANN with the Varied Amount of Neurons in the Hidden Layer

Amount of neurons	Iteration number (epochs)	$\sigma_{\Delta N}$ (cm)	Maximum ΔN (cm)	Minimum ΔN (cm)
5	161	±5.08	11.30	−15.70
10	948	±4.13	23.53	−10.82
15	1,606	±4.00	21.42	−10.49
20	487	±4.42	11.79	−16.44
25	3,628	±4.22	20.08	−15.35
30	1,459	±3.85	9.67	−13.40
35	667	±3.60	12.94	−10.39
40	1,494	±5.44	12.58	−13.33
45	541	±5.08	11.89	−16.84
50	548	±5.08	11.89	−16.84

J. Surv. Eng. 2007.133:81-89.

Table 3. BP ANN Accuracy Comparisons with Other Undulation Estimation Methods

Estimation method	$\sigma_{\Delta N}$ (cm)	Maximum ΔN (cm)	Minimum ΔN (cm)
BP ANN	±3.60	12.94	−10.39
Curve fitting	±19.88	82.86	−60.33
MOI model	±20.07	34.70	−47.10

that a polynomial of Eq. (3) was used to fit the calculated undulations of the reference points, and that the computed polynomial coefficients of Eq. (3) were used to estimate the check point's undulation. "MOI model" denotes that a local geoid model of Taiwan, determined by least-squares collocation using observations of sea surface gradients derived from satellite altimetry, land gravity anomalies, ship gravity anomalies, and the Earth Gravitational Model 1996 (EGM96) (Hwang 1997), was used to estimate the undulations of those check points. From Table 3, it can be seen that the performance of the undulation estimation using the BP ANN was better than those of the other two estimation methods. The $\sigma_{\Delta N}$ values for BP ANN, curve fitting, and MOI model are ±3.60, ±19.89, and ±20.07 cm, respectively.

Performance Evaluation of the Application of a Back-Propagation Artificial Neural Network to Regional Grid-Based Geoid Model Generation Using GPS and Leveling Data

According to the abovementioned "regional grid-based geoid model generation procedure," the grid-based geoid model of the Taiwan region was generated accordingly. Hence, the tests described in the following concentrated on the issues of performance evaluation of the generated Taiwan grid-based geoid model, such as the grid interval selected, the amount of used reference points, and so on. The test items included: (1) the interpolation accuracies of the grid-based geoid model versus the varied amount of neurons in the hidden layer of the BP ANN if the grid interval was fixed; (2) the interpolation accuracies of the grid-based geoid model versus the varied amount of reference points if the grid interval was fixed; and (3) the interpolation accuracies of the grid-based geoid model versus varied grid intervals if the amount of reference points was fixed.

In order to test Item 1, the data from the 2,065 benchmarks were used. The 2,065 benchmarks were divided into two groups: 1,033 points were treated as reference points to train the BP ANN, and the other 1,032 points were treated as check points to evaluate the undulation estimation/interpolation accuracy. One type of grid-based geoid model was generated with fixed 1,000 m grid interval. The total amount of grid nodes of the generated geoid models with grid intervals of 1,000 m are 56,347 nodes. First, the 1,033 reference points were used to train the BP ANN with the varied amount of neurons (from 5 to 50) in the hidden layer; then the plane coordinates of all imaginary grid nodes were entered into the BP ANN to generate the grid-based geoid model; and finally, the 1032 check points were used to evaluate: (1) the undulation estimation accuracy directly of the trained BP ANN and (2) the corresponding undulation interpolation accuracy of the grid-based geoid model. The test results are shown in Table 4. "BP ANN" denotes that the test results of undulations were estimated from the trained BP ANN directly. "Grid model" indicates that the test results of undulations were interpolated from the generated grid-based geoid model with fixed grid interval (e.g., 1,000 m in this test). The contour map of the differences (cm) between the interpolated undulations from generated grid-based geoid model (with 1,000 m grid interval and 50 neurons in the hidden layer) and the original undulations of 1,032 check points, ΔN's, is shown in Fig. 5. From Table 4 and Fig. 5, it can be seen that: (1) the accuracies of grid model are as accurate as those of BP ANN in all cases, i.e., the generated regional grid-based geoid model could be an alternative approach to the "trained BP ANN;" (2) if the amount of neurons in the BP ANN is larger, then the interpolation accuracy is better; and (3) when the amount of neurons is 50, then the interpolation accuracy, $\sigma_{\Delta N}$ was about ±3.98 cm for a grid interval of 1,000 m.

Table 4. Performance Statistics of the BP ANN and Grid Model with the Varied Amount of Neurons in the Hidden Layer of the BP ANN

Amount of neurons	Estimation method	$\sigma_{\Delta N}$ (cm)	Maximum ΔN (cm)	Minimum ΔN (cm)
5	BP ANN	±27.62	76.96	−79.78
5	Grid Model	±27.64	87.99	−76.95
10	BP ANN	±12.41	37.48	−32.17
10	Grid Model	±12.43	32.42	−37.47
20	BP ANN	±5.65	25.47	−26.21
20	Grid Model	±5.66	26.22	−25.47
30	BP ANN	±4.37	22.92	−22.17
30	Grid Model	±4.38	22.20	−22.92
40	BP ANN	±4.10	22.11	−30.00
40	Grid Model	±4.17	37.24	−22.11
50	BP ANN	±3.96	21.57	−38.31
50	Grid Model	±3.98	38.56	−21.50

In order to test Item 2, the data from the 2,065 benchmarks of the Taiwan region and the 78 benchmarks of Taichung City were used. The 2,065 benchmarks, which were all treated as reference points, were divided into two groups: the whole Taiwan region case (including all 2,065 points) and the central Taiwan region case (including only 283 points). In each case, if the BP ANN (with 50 neurons in the hidden layer) was trained successfully with the varied amount of reference points, then the grid-based

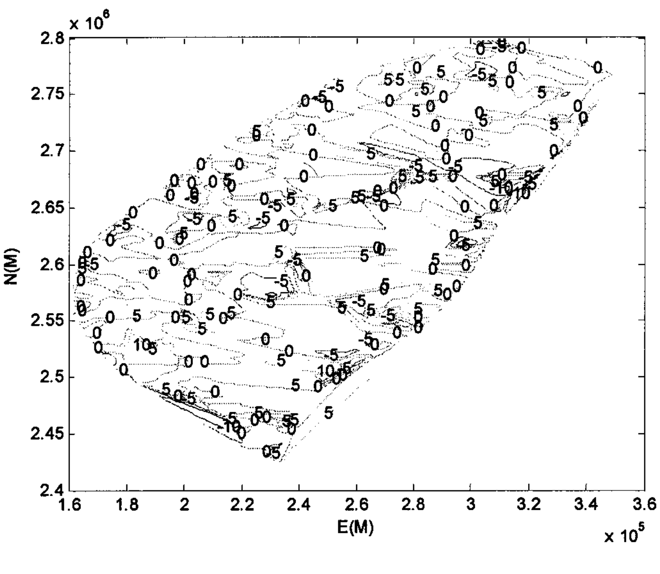

Fig. 5. Contour map of the differences (cm) between the interpolated undulations from grid-based geoid model and the original undulations of 1,032 check points of the Taiwan region

Table 5. Performance Statistics of the BP ANN and Grid Model with the Varied Amount of Reference Points, in the Cases of Whole Taiwan Region (Case A) and Central Part of Taiwan (Case B)

Case	Amount of reference points	Amount of check points	Estimation method	$\sigma_{\Delta N}$ (cm)	Maximum ΔN (cm)	Minimum ΔN (cm)
A	2,065	78	BP ANN	±5.22	16.04	−6.38
A	2,065	78	Grid model	±5.22	6.35	−16.07
A	2,091	52	BP ANN	±3.52	10.27	−3.74
A	2,091	52	Grid model	±3.53	3.73	−10.32
A	2,104	39	BP ANN	±3.37	10.00	−3.48
A	2,104	39	Grid model	±3.38	3.47	−10.06
B	283	78	BP ANN	±4.27	14.83	−4.42
B	283	78	Grid model	±4.28	4.41	−14.86
B	309	52	BP ANN	±3.13	7.98	−7.58
B	309	52	Grid model	±3.14	7.55	−7.98
B	322	39	BP ANN	±2.16	6.81	−13.5
B	322	39	Grid model	±2.17	2.17	−6.88

geoid model with a grid interval of 1,000 m was generated. Then the check points were used to evaluate: (1) the undulation estimation accuracy of the trained BP ANN; and (2) the interpolation accuracy of the grid-based geoid model. The test results of the whole Taiwan region case and the central Taiwan region case are shown in Table 5. It should be noted that the 78 benchmarks of Taichung City were divided into two groups: Reference points and check points. For example, from Table 5, it can be seen that if the amount of check points was 78, then all 78 benchmarks were used as check points. However, if the amount of check points was 52, then only 52 points from the 78 benchmarks were used as check points and the remaining 26 points were treated as reference points. Therefore, the corresponding amount of reference points became 2,091; that is, 2,065+26=2,091. From Table 5, the following can be seen: (1) the accuracies of grid model are as accurate as those of BP ANN in all cases; (2) if the amount of reference points increases, for example, from 2,065 to 2,104, then the $\sigma_{\Delta N}$ value decreases, e.g., from ±5.22 to ±3.38 cm, respectively, for the whole Taiwan region case (i.e., case A in Table 5). The same conclusion can be made about the central Taiwan region case. The smallest $\sigma_{\Delta N}$ values were ±3.37 and ±2.16 cm for the whole Taiwan region case and the central Taiwan region case respectively; and (3) if the conditions, such as the grid interval and the amount of check points, were the same, then if the smaller region of the geoid model was generated, the interpolation accuracy was better. In other words, the interpolation accuracy of the central Taiwan region case was better than that of the whole Taiwan region case.

In order to test Item 3, the data from the 283 benchmarks of the central Taiwan region and the 78 benchmarks of Taichung City were used. Three kinds of tests were carried out. In the first test, the amounts of reference points and check points were 283 and 78, respectively, and the grid intervals varied from 500, through 1,000 to 1500 m. In the second test, the amounts of reference points and check points were 309 and 52, respectively, and the grid intervals varied from 500, through 1,000 to 1500 m. In the third test, the amounts of reference points and check points were 322 and 39, respectively, and the grid intervals varied from 500, through 1,000 to 1500 m. The test results are shown in Table 6. Grid model (N m) denotes that the test results of grid-based geoid model with grid interval N m (e.g., 500 m). From Table 6, it can be seen that: (1) the accuracies of Grid model (N m) are almost as accurate as those of BP ANN in all cases; (2) if the amount of reference points was fixed, then the interpolation accuracies of varied grid intervals (e.g. 500, 1,000, and 1,500 m in this test) were almost the same; (3) If the grid intervals were the same, then the larger the amount of reference points, the better the interpolation accuracy; and (4) the best accuracy in this test is ±2.16 cm in case of 322 reference points fixed.

Performance Evaluation of Undulation Interpolation Using a Pocket PC Program

In order to fulfill the requirement of transforming the GPS-derived ellipsoidal height h to the orthometric height H in the field, the undulation interpolation algorithm from the generated regional grid-based geoid model was developed for a pocket PC operation system by means of Microsoft Studio Net 2003 (Deitel et al. 2004). Then, the interpolation program and the grid-based geoid model were transformed into ones suitable for a pocket PC. The interface of the interpolation program on a pocket PC is shown in Fig. 6. The "open file" key denotes that the system is waiting for the user to "open" a geoid model file. The "read file" key indicates that the system is asked to read the opened file. The

Table 6. Performance Statistics of the BP ANN and Grid Model with the Varied Grid Interval If the Amount of Reference Points Was Fixed

Amount of reference points	Amount of check points	Estimation method	$\sigma_{\Delta N}$ (cm)	Maximum ΔN (cm)	Minimum ΔN (cm)
283	78	BP ANN	±4.27	14.83	−4.42
283	78	Grid model (500 m)	±4.27	4.42	−14.83
283	78	Grid model (1,000 m)	±4.28	4.41	−14.86
283	78	Grid model (1,500 m)	±4.28	4.39	−14.83
309	52	BP ANN	±3.13	7.98	−7.58
309	52	Grid model (500 m)	±3.13	7.56	−7.98
309	52	Grid model (1,000 m)	±3.14	7.55	−7.98
309	52	Grid model (1,500 m)	±3.14	7.49	−7.98
322	39	BP ANN	±2.16	6.81	−1.35
322	39	Grid model (500 m)	±2.16	2.16	−6.84
322	39	Grid model (1,000 m)	±2.17	2.17	−6.88
322	39	Grid model (1,500 m)	±2.18	2.18	−7.02

J. Surv. Eng. 2007.133:81-89.

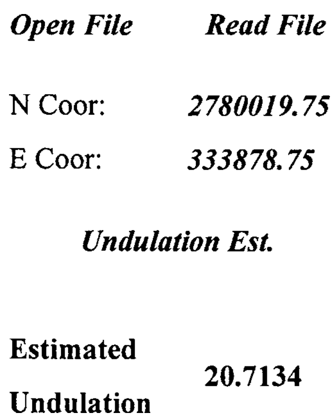

Fig. 6. Interface of the undulation interpolation program on a pocket PC

"N Coor.:" and "E Coor.:" keys indicate the entered plane coordinates of a specific point from the user. Then, the "undulation est." key is pressed to run the undulation interpolation program, and the estimated undulation is shown after the "estimated undulation" key.

In order to test the performance of undulation interpolation using the pocket PC program, nine control points collected from the campus of NCCU were used as check points. The NCCU test region is about 400 m by 850 m. The ellipsoidal heights of these check points were derived from the RTK GPS survey, while the orthometric heights of these check points were observed by the three-order geodetic leveling survey. The grid-based geoid model with a grid interval of 1000 m was generated by the data from the 2,065 benchmarks. The pocket PC with the developed interpolation program was taken to those control points in the field with an RTK GPS receiver. According to the test results, the $\sigma_{\Delta N}$ value, maximum ΔN, and minimum ΔN, of these nine check points were ±2.72, +2.79, and −4.31 cm, respectively.

Conclusions

In order to fulfill the requirement of transforming the GPS-derived ellipsoidal height h to the orthometric height H and take advantage of GPS surveying, algorithms of applying a back-propagation artificial neural network (BP ANN) to generate a Taiwan regional grid-based geoid model using GPS and leveling data were proposed. Three data sets, including 2,065 benchmarks around the Taiwan region, 78 benchmarks in Taichung City, and 9 check points in National Chengchi University, were used to test the proposed algorithms.

Based on the test results, the following comments can be made: (1) the training algorithm trainbr should be adopted in a BP ANN program; (2) the undulation estimation accuracy using the BP ANN is better than the other two estimation methods, namely the curve fitting method and the MOI model; (3) the undulation interpolation accuracies of the generated grid-based geoid model are as accurate as those undulation estimation accuracies directly from the trained BP ANN, i.e., the proposed grid-based geoid model algorithm is an alternative approach to the trained BP ANN; (4) the undulation interpolation accuracies of the generated grid-based geoid model with a larger amount of reference points are better than those of the grid-based geoid model with a smaller amount of reference points, if the grid interval is fixed; (5) the Taiwan region grid-based geoid model and the undulation interpolation program were developed and installed in a pocket PC; and (6) this paper presents a method using BP ANN for approximation of the regional geoid surface, however, further investigations should be done to refine the approximation geoid surface, such as using the least squares collocation, etc.

Acknowledgments

The work presented in this paper was funded, in part, by the National Science Council, Taiwan, R.O.C. (Project No. NSC 93-2211-E-004-001). The test data sets were kindly provided by the Satellite Survey Center, Department of Land Administration, Ministry of Interior, Taiwan, R.O.C.

References

Ambrozic, T., Kuhar, M., Stopar, B., and Turk, G. (1999). "Approximation of local geoid surface by artificial neural network." *Proc., FIG (International Federation of Surveyors) Commission 5 Symp. Geodesy and Surveying in the Future—The Importance of Heights*, NLS, Gavle, 273–283.

Chen, C. H., Yang, M., and Tseng, C. L. (2004). "Improving local geoid model of Taiwan using TWVD2001 leveling and GPS data: With application to GPS heighting." *J. Surv. Eng.*, 46(1), 1–20 (in Chinese).

Deitel, H. M., Deitel, P. J., and Nieto, T. R. (2004). *Visual Basic.Net: How-to program*, 2nd Ed., Prentice-Hall, Upper Saddle River, N.J.

Demuth, H., and Beale, M. (2002). *User's guide of neural network toolbox for use with MATLAB*, Version 4, The MathWorks, Inc., Natick, Mass.

Hu, W., Kong, Y., Deng, Y., and Kuang, S. (2002). "A new method to derive normal height from GPS height based on neural network." *FIG 22th Int. Congress*, Washington, D.C., 1–11.

Hu, W., Sha, Y., and Kuang, S. (2004). "New method for transforming global positioning system height into normal height based on neural network." *J. Surv. Eng.*, 130(1), 36–39.

Hwang, C. (1997). "Analysis of some errors affecting altimeter-derived sea surface gradient with application to geoid determination over Taiwan." *J. Geodesy, Berlin*, 71(2), 113–130.

Junkins, J. L., Miller, G. W., and Jancaitis, J. R. (1973). "A weighting function approach to modeling of irregular surfaces." *J. Geophys. Res.*, 78(11), 1794–1803.

Kavzoglu, T., and Saka, M. H. (2005). "Modeling local GPS/leveling geoid undulations using artificial neural networks." *J. Geodesy, Berlin,* 78(9), 520–527.

Kuhar, M., Stopar, B., Turk, G., and Ambrozic, T. (2001). "The use of artificial neural network in geoid surface approximation." *Allgemeine Vermessungs-Nachrichten*, 108(1), 22–27.

Lin, L. S. (1998). "Real-time estimation of ionospheric delays using GPS measurements." *Rep. No. UNISURV S-51*, School of Geomatic Engineering, The Univ. of New South Wales, Sydney, Australia.

Stopar, B., Ambrozic, T., Kuhar, M., and Turk, G. (2006). "GPS-derived geoid using artificial neural network and least squares collocation." *Surv. Rev.*, 38(300), 513–524.

Yang, M., Chen, K. H., and Shiao, S. W. (2003). "A new height network in Taiwan." *Surv. Rev.*, 37(290), 260–268.

Yang, Z. J., and Chen, Y. Q. (1999). "Determination of local geoid with geometric method: Case study." *J. Surv. Eng.*, 125(3), 136–146.

Zaletnyik, P., Volgyesi, L., and Palancz, B. (2004). "Approach of the Hungarian geoid surface with sequence of neural networks." *Geo-Imagery Bridging Continents, 20th ISPRS Congress*, Vol. 35, Istanbul, Turkey, Part B8, 119–122.