

A Hybrid Model to Predict Cartesian Planimetric Coordinates

Bernard Kumi-Boateng¹ and Yao Yevenyo Ziggah^{1,2}

1. Department of Geomatic Engineering, University of Mines and Technology, Ghana

2. Department of Survey and Mapping, China University of Geosciences, Wuhan, China

Abstract: Coordinates are basic needs for both geospatial and non-geospatial professionals and as a result, geodesists have the responsibility to develop methods that are applicable and practicable for determining cartesian coordinates either through transformation, conversion or prediction for the geo-scientific community. It is therefore necessary to implement mechanisms and systems that can be employed to predict coordinates in either two dimensional (2D) or three dimensional (3D) spaces. Artificial Intelligence (AI) techniques and conventional methods within the last decade have been proposed as an effective tool for modeling and forecasting in various scientific disciplines for solving majority of problems. The primary objective of this work is to compare the efficiency of artificial intelligence technique (Feed Forward Back propagation Neural Network (FFBPNN)) and conventional methods (Ordinary Least Squares (OLS), General Least Squares (GLS), and Total Least Squares (TLS)) in cartesian planimetric coordinate's prediction. In addition, a hybrid approach of conventional and artificial intelligence method thus, TLS-FFBPNN has been proposed in this study for 2D cartesian coordinates prediction. It was observed from the results obtained that FFBPNN performed significantly better than the conventional methods. However, the TLS-FFBPNN when compared with FFBPNN, OLS, GLS and TLS gave stronger and better performance and superior predictions. To further confirm the superiority of the TLS-FFBPNN the Bayesian Information Criterion was introduced. The BIC selected the TLS-FFBPNN as the optimum model for prediction.

Key words: Artificial intelligence, ordinary least squares, total least squares, general least squares, cartesian coordinates.

1. Introduction

In surveying and mapping, one of the primary tasks for geodesist and surveyors is to determine precise coordinates for accurate positioning on the Earth surface. These coordinates may range from the global to local sphere and are pertinent for a variety of applications in both applied science and geodesy. Over the years, the issues of precise point positioning for geodetic applications have become a real, persistent and serious problem facing the developing countries like Ghana, calling for pragmatic solutions. The increasing use of the Global Navigation Satellite System (GNSS) such as Global Positioning System (GPS) has significantly advanced the expansion of

global application of geodetic surveys in all areas of the globe, even though many countries, for example, Ghana are still yet to fully utilize the full potentials of GPS.

Ghana is one such post-colonial countries' where the use of non-geocentric coordinate system established through conventional techniques of surveying for national mapping still persist [1, 2]. Moving along with the disheartened political, economic and socio-cultural trend, there are numerous challenges to incorporating GPS derived coordinates into the Ghana National coordinate system, thus, demanding the determination of transformation parameters by scaling, translation and rotation. To solve this, 3D similarity transformation models of three translations, three rotations and one scale factor [3, 4] are used to estimate the transformation parameters. Several transformation models have been put forth to estimate transformation parameters for geodetic reference network for

Corresponding author: Bernard Kumi-Boateng, Ph.D., senior lecturer and vice dean of International Programmes, research fields: geo-spatial land planning and vegetation mapping.

countries [5-8] such as Ghana. The most commonly used among these models in Ghana are the conformal methods of Bursa-Wolf, Molodensky-Badekas and Helmert similarity [8-10]. It is worth noting that the first step in applying any of the above-mentioned conformal similarity transformation models in Ghana is the conversion of the geodetic data of common points into cartesian rectangular coordinates using Bowring's algorithm.

Although Bowring's algorithm [11-15] has established a forward computation for transforming geodetic coordinates to cartesian coordinates, it will take some time before this can be applied directly in Ghana due to the non-geocentric coordinate system. Therefore, for local application to be realized, the iterative Abridged Molodensky transformation model [16, 17] must be applied to the geographic coordinates of common points on the War office ellipsoid [10]. This will then be used to determine the approximate change in ellipsoidal heights which are then used to calculate the ellipsoidal height for the local geodetic network before Bowring's equation can be applied to the data. The reason is because the Ghana geodetic reference network established through classical methods of surveying is referenced on the War Office ellipsoid with data in latitude, longitude and orthometric height without the existence of ellipsoidal height [2, 8, 9, 18].

The above prevailing situation has hampered the use of GPS effectively thereby creating inconsistencies in the transformation parameters derived among researchers in Ghana. As a result, users from all related disciplines adopt different techniques in the post-processing of the GPS observational raw data and these generate a lot of uncertainties in the use of the GPS data from different sources [9, 10]. Also, the lack of geoid model for Ghana War Office datum has also contributed to the inconsistencies in the transformation parameters determined by researchers in Ghana. These have lead to non-existence of standards to regularize GPS usage in its domain, even though it is

acknowledged that the use of GPS for geodetic surveys is on the increase in Ghana and around the world.

Efforts have been made by researchers [2, 8-10, 18, 19] in Ghana and foreign agencies to determine a suitable method and transformation parameters for the Ghana Geodetic Reference Network. No decision has been reached as to which of them is the selected technique and no doubt this is an area where deliberations will continue for some time to come. Consequently, an optimum set of transformation parameters between the War Office datum and WGS 84 datum does not exist; this makes it difficult for Ghana to utilize the massive potential of GNSS services. In order to establish compatibility in data obtained from GPS measurements correctly and effectively in Ghana, it is necessary to determine appropriate transformation parameters that relate the coordinates in the War Office system to the WGS 84 system using relevant transformation models. In view of the above development, it can be inferred that the inconsistencies in the parameters determined among researchers is further creating distortions in the local coordinates generated from these parameters. Therefore, there is a need to model out distortions within the local coordinates in our local geodetic system after applying the national transformation parameters.

Conversely, there exist physically based numerical methods that can serve as practical alternative in modelling out these distortions by way of converting geodetic data to cartesian coordinates for a local geodetic network by way of prediction to a tolerable degree of accuracy and precision. For example, Ref. [20] applied linear regression (OLS) to find a relationship between global and cartesian coordinates. They concluded that the proposed linear regression models were suitable for predictions at 95% confidence interval and do not violate any of the statistical assumptions of a linear model. In Ref. [21] linear regression was also applied for two different regions in the Ondo State, Nigeria whereby their respective

results were compared. It was established by the authors that distances between random distributions of points can be a factor when developing regression model for local coordinate prediction. Also, Ref. [22] developed regression models for coordinate prediction at Makkah Metropolitan area. The simplicity and accuracy of the regression models were considered. Moreover, the authors' proposed that these regression models could be used for surveying and mapping activities such as GIS, data collection, engineering surveying and topographic mapping based upon their prediction capabilities.

Other numerical methods applicable for predicting local cartesian coordinate is the Total Least Squares (TLS) and General Least Squares (GLS). It is worth mentioning that several studies have been carried out by researchers into TLS in 3D coordinate transformation and in robotics [23-26]. Although extensive applications of linear regression (OLS) and TLS has been carried out, limited literature is available in geodesy technical papers on the application of TLS, GLS for two-dimensional (2D) local cartesian coordinate prediction.

This study also considered the artificial intelligence technique of Feed Forward Back Propagation Neural Network (FFBPNN) as an alternative to predicting local cartesian coordinates from global coordinates within the local geodetic system in Ghana. It must be emphasized that artificial intelligence techniques has been widely used in solving problems in geodesy. For example, researchers like [27-30] applied the back propagation and radial basis function neural network as an alternative to 3D coordinate transformation by comparing with conformal similarity models. Moreover, several research works have been carried out using artificial neural network (ANN) for GPS height conversion. This can be found in [31-34]. Yet, a study on the integration of conventional technique into ANN to improve the efficiency of the model developed has not been comprehensively investigated for planimetric coordinate predictions.

Therefore, this study focused on Ghana as a case for application of the conventional methods (OLS, GLS, and TLS), FFBPNN and a proposed hybrid approach of TLS-FFBPNN for the study area. The hybridized model yielded the most precise coordinate values in good agreement with the measured data.

2. Study Area

Ghana is a country located in the Western part of Africa and is bordered by Cote D'Ivoire to the West, Togo to the East, Burkina Faso to the North and the Gulf of Guinea to the South. The country has a 239,460 km² land mass generally consisting of low plains [35] with 2,093 km of international land borders. Lying just above the equator, Ghana has a tropical climate with mean annual temperature ranging between 26 °C and 29 °C but temperature are generally higher in the North than in the South [35]. The country is divided into ten administrative regions as shown in Fig. 1 below.

This study covers the western administrative region in Ghana (Fig. 1). This region have almost all the natural resources such as gold, bauxite, manganese, oil, timber, cocoa and many others found in the country and thus, are of high economic value. It also forms part of the ongoing newly established geodetic reference network referred as the Golden Triangle [19].

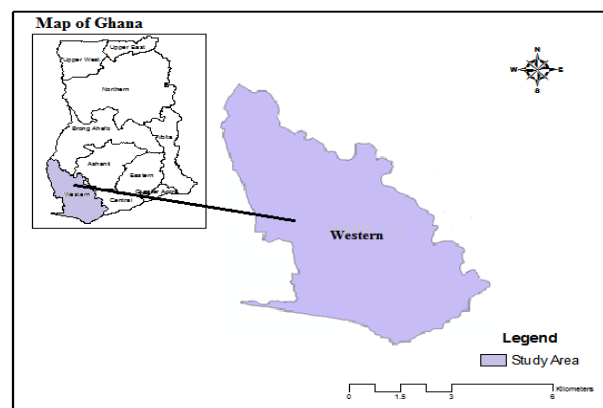


Fig. 1 Map of Ghana showing the study area.

3. Data

Primary control point data of global geodetic coordinates (ϕ, λ) were collected by field work using the Trimble GPS receivers with the static GPS surveying technique adopted. Their corresponding existing cartesian coordinates (E, N) in the Ghana local coordinate system were acquired from a mining company's database in Ghana. A total of 200 above-mentioned datasets were used in this study to develop the prediction models and obtain an approximation of the relationship between (ϕ, λ) and (E, N) respectively. An independent datasets of 26 was then applied to test the model prediction accuracy level.

4. Methods

Normally in applied numerical problems in engineering and in the field of sciences, it is essential to fit a surface or curve to a set of data points with known coordinates. The most commonly used method is the Ordinary Least Squares. Its aim is to minimise the sum of squared errors of only one of the variables but not of the others [36]. However, in surveying and mapping, all observed data (coordinates) may suffer from errors. Fortunately, there exist numerical techniques that incorporate errors in both coordinates in their respective procedures. Two of such methods namely, Total Least Squares and General Least Squares were considered in this study. Artificial intelligence technique namely, Feed Forward Back Propagation Neural Network was also applied. Finally, a hybrid model was proposed.

4.1 Ordinary Least Squares

The Ordinary least Squares (OLS) mathematical model for fitting a straight line to points is of the form $y = mx + c$ where x and y are the coordinates of a point, m is the slope of a line, and b is the y intercept at $x = 0$. The assumption in the OLS is that residuals are applied only to the observation vector (y coordinate) even though both observations (x, y) contain errors. In

an ideal situation, all coordinates points must lie on a least squares line if they are truly linear and no observation errors exist. On the other hand, this is seldom the case in that (1) the observations contain errors; (2) the functional model could be a higher-order curve, or both (Ghilani, 2010). If the desirable best-fitting straight line is found for the data, residuals are added to the least square equation. Hence, accounting for the errors in the observation vector (y). Assuming four control points A, B, C, and D, its observation equation in least squares term can be written as

$$mx_A + c = y_A + v_{yA}$$

$$mx_B + c = y_B + v_{yB}$$

$$mx_C + c = y_C + v_{yC}$$

$$mx_D + c = y_D + v_{yD}$$

The above observation equation contains two unknown parameters, m and b , with four observations. Their matrix representation in the least squares form is $AX = L + V$ where

$$A = \begin{bmatrix} x_A & 1 \\ x_B & 1 \\ x_C & 1 \\ x_D & 1 \end{bmatrix}, \quad X = \begin{bmatrix} m \\ b \end{bmatrix}, \quad L = \begin{bmatrix} y_A \\ y_B \\ y_C \\ y_D \end{bmatrix},$$

$$V = \begin{bmatrix} v_{yA} \\ v_{yB} \\ v_{yC} \\ v_{yD} \end{bmatrix}$$

The equation is then solved using the least squares method $X = (A^T A)^{-1} A^T L$ and the observational residuals are calculated using $V = AX - L$.

4.2 General Least Squares

The General Least Squares (GLS) postulates that both observed variables (x and y) contain errors since

they are observational data. Yet in the case of the OLS mathematical model, the residuals are applied only to the y coordinate. This clearly shows the deficiency in the OLS model by failing to account for the error in the other observed variable (x coordinate). Assuming that four control points A, B, C, and D were surveyed on the field, its observation equation in GLS can be written as [37]

$$F(x, y) = (y + v_y) - m(x + v_x) - b = 0$$

where V_x and V_y are the residuals of the coordinate points (x, y), m is the slope of the line and b is the intercept. The above equation having the unknown parameters is nonlinear and therefore requires linearization.

In this study, Taylors Series for linearizing nonlinear equations was applied. The resulting equation for the four assumed control points after linearization can be written as:

$$\begin{aligned} -m_o V_{x_A} + V_{y_A} - x_A dm - db &= -(m_o x_A + b_o - y_A) \\ -m_o V_{x_B} + V_{y_B} - x_B dm - db &= -(m_o x_B + b_o - y_B) \\ -m_o V_{x_C} + V_{y_C} - x_C dm - db &= -(m_o x_C + b_o - y_C) \\ -m_o V_{x_D} + V_{y_D} - x_D dm - db &= -(m_o x_D + b_o - y_D) \end{aligned}$$

These equations can be represented in matrix form as:

$$BV + JX = K$$

Where

$$B = \begin{bmatrix} -m_o & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -m_o & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -m_o & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -m_o & 1 \end{bmatrix}$$

$$J = \begin{bmatrix} -x_A & -1 \\ -x_B & -1 \\ -x_C & -1 \\ -x_D & -1 \end{bmatrix}, \quad V = \begin{bmatrix} v_{x_A} \\ v_{y_A} \\ v_{x_B} \\ v_{y_B} \\ v_{x_C} \\ v_{y_C} \\ v_{x_D} \\ v_{y_D} \end{bmatrix}, \quad X = \begin{bmatrix} dm \\ db \end{bmatrix},$$

$$K = \begin{bmatrix} -(m_o x_A + b_o - y_A) \\ -(m_o x_B + b_o - y_B) \\ -(m_o x_C + b_o - y_C) \\ -(m_o x_D + b_o - y_D) \end{bmatrix}$$

In solving for the general least squares, the unknown parameters dm and db is estimated using the relationship; $X = (J^T W_e J)^{-1} J^T W_e K$. Where W_e is the equivalent weight matrix given us; $W_e = (B \Sigma B^T)^{-1}$ and Σ is the covariance matrix of the observations. The observational residuals are estimated using the equation $V = \Sigma B^T W_e V_e$ and $V_e = JX - K$ is the equivalent residuals. It should be noted that GLS is a nonlinear equation system and therefore requires iteration until convergence is achieved.

4.3 Total Least Squares (TLS)

The Total Least Squares (TLS) is a technique that is appropriate when there is an existence of errors in both the observation vector and in the data matrix [38]. In the plane, the main goal of TLS is to determine a line that minimises the sum of the squared distances from that line to a given data point in the plane. The total least squares method adopt the symmetric form of the equation of a line. Thus, the distance d from a point say (p, q) to a straight line will have the equation $rx + sy - c = 0$ that acknowledges the following formula [36]:

$$d^2 = \frac{(rp + sq - c)^2}{r^2 + s^2}$$

If the line passes through a point $\vec{X}_0 = (x_0, y_0)$, then $c = rx_0 + sy_0$. Therefore the equation becomes $r(x - x_0) + s(y - y_0) = 0$. So, with given data points $\vec{X}_1 = (x_1, y_1), \dots, \vec{X}_n = (x_n, y_n)$, the problem of the TLS consists of determining a line, defined by a point (x_0, y_0) on the line and by normal vector (r, s) , which minimises the sum D of the squared distances given as

$$D(r, s, \vec{x}_0; \vec{x}_1, \dots, \vec{x}_n) = \sum_{i=1}^n \frac{(r(x_i - x_0) + s(y_i - y_0))^2}{r^2 + s^2}.$$

It must be known that the line of total least squares passes through the centroid of the data points.

4.4 Feed Forward Back Propagation Neural Network (FFBPNN)

Artificial neural networks (ANNs) are computational standards based on mathematical models inspired by the structure, operation and behaviour that imitate the human brain of biological neurons in the nervous system. These networks consist of highly interconnected neurons (Sordo, 2002) via links of variable synaptic weights (Handhel, 2009) that process information through an interaction with the environments. Generally, there are different types of ANN based on their architecture. The most popular network architecture in use today for engineering applications is the Feed Forward Back Propagation Neural network (FFBPNN) (Haykin, 2007). This network consists of three layers; input layer which act as control variables having an influence over the desired output of the neural network and an output layer giving the internal computation results. Between these two layers exist another layer not evident from the outside called the hidden layer liable for executing intermediate calculations (Konaté, 2015). It is pertinent to note that the efficiency of a neural network and its simplicity depends on the appropriate number of neurons applied in the hidden layer.

The number of hidden neurons is determined by trial and error procedure depending on the characteristics of the problem, training samples size and requirement. This is due to some of these factors; the problem under consideration, the choice of neural network structure and proposed theories that are yet to be accepted universally to clarify the number of hidden units needed to approximate a given function. In this study, the optimum number of neurons in the hidden layer was selected by sequential trial and error procedure based on the smallest mean square error (MSE).

The choice for the hidden layer used in this study was 1. This decision was based on [43] which further corroborate that Feed Forward Back Propagation Neural Network (FFBPNN) with one hidden layer together with their hidden and output transfer functions is a universal approximator of any discrete and continuous functions. The hyperbolic tangent was chosen as the hidden neurons activation function while a linear transfer function was used in the output neurons. The hyperbolic tangent can be expressed mathematically as

$$f(t) = \tan \text{sig}(t) = \frac{e^t - e^{-t}}{e^t + e^{-t}}$$

where t is the sum of the weighted inputs.

For the network training, the objective is to find optimal weight connection w^* in such a way that it will provide estimated outputs that match for each example of the desired outputs value. This is a typical nonlinear optimization problem, where w^* is given as [42]

$$w^* = \arg \min E(w),$$

where w is the weight matrix and $E(w)$ is an objective function on w , which is to be minimize. The $E(w)$ is evaluated at any point of w as follows:

$$E(w) = \sum_n E_n(w),$$

where n is the number of examples in the training set and $E_n(w)$ is the output error for each example n and is defined as

$$E_n(w) = \frac{1}{2} \sum_j (d_{nj} - y_{nj}(w))^2,$$

where $y_{nj}(w)$ and d_{nj} are the calculated and desired network outputs of the j -th output neuron for n -th example, respectively. Substituting $E_n(w)$ into $E(w)$ gives

$$E(w) = \frac{1}{2} \sum_n \sum_j (d_{nj} - y_{nj}(w))^2$$

In this study, Levenberg-Marquardt algorithm (LMA) was chosen to train the neural network because it is faster and has more stable convergence than the classical gradient descent algorithm. This was proven in the work of [44]. The LMA can be thought of as a combination of steepest descent and the Gauss-Newton method. The algorithm behaves like steepest descent method when the current solution is far from the correct one. Thus, slow but guaranteed to converge. On the other hand, when the algorithm approaches the correct solution, it becomes a Gauss-newton method. Detailed analysis of the LMA is beyond the scope of this study and interested reader is referred to [45] for more comprehensive treatments.

4.5 Concept of the Proposed Method

Three conventional techniques namely OLS, GLS, and TLS were employed in this study. Having considered the strengths and weaknesses as well as the results obtained for each method the TLS was the preferred method to OLS and GLS for the hybridization with FFBPNN. Hence, the proposed hybrid technique of TLS-FFBPNN was developed. The concept is elaborated below.

(1) First the TLS models were determined for Eastings and Northings coordinates respectively. However, like in every survey measurements there exist random errors after any least squares adjustment.

This therefore limited the prediction potential of the TLS models developed. Also, jackknifing validation approach being implemented further show that the model is not in good agreement with the measured data.

(2) In order to improve the efficiency of the model, the residuals obtained from the TLS adjustment and its predicted coordinates values was used as a training sample in conjunction with the global coordinates in the FFBPNN.

(3) The TLS-FFBPNN hybrid model was then used to predict the 2D cartesian coordinates.

5. Performance Criteria

The evaluation of the performance of the implemented methods was focused on the residuals generated from test data by each prediction models developed for predicting (E, N) from (ϕ, λ) . In order to check the validity of the solutions obtained from the prediction models, the following statistical estimation accuracy measures were used as a performance criteria index (PCI) namely: Mean Square error (MSE), Root Mean Square error (RMSE), Nash-Sutcliffe Efficiency Index (η), Index of Agreement (D) and Modified Index of Agreement (D). Mathematical representations of the PCIs are given by:

$$MSE = \frac{1}{n} \sum_{i=1}^n (t_i - o_i)^2,$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (t_i - o_i)^2},$$

$$\eta = 1 - \frac{\sum_{i=1}^n (t_i - o_i)^2}{\sum_{i=1}^n (t_i - \bar{t})^2}, \quad D = 1 - \frac{\sum_{i=1}^n (t_i - o_i)^2}{\sum_{i=1}^n (t_i' + o_i')^2},$$

$$\text{where } t_i' = |t_i - \bar{o}|, \quad o_i' = |o_i - \bar{o}|,$$

$$MID = 1 - \frac{\sum_{i=1}^n |t_i - o_i|}{\sum_{i=1}^n (|o_i - \bar{t}| + |t_i - \bar{t}|)}$$

where n is the total number of examples presented to the learning algorithm, t and o are the observed values and predicted values. \bar{t} is the mean of the observed values and \bar{o} is the mean of the predicted values respectively.

6. Results and Discussion

6.1 Models Developed

Different model equations were determined between geodetic coordinates (latitude, longitude) and cartesian coordinates (Easting (E), Northings (N)) respectively. A total of 200 data points were used to form the models. Tables 1 and 2 below present the various models developed when OLS, GLS, TLS, FFBPNN and TLS-FFBPNN were applied on the datasets.

From Table 1 above, it can be observed that the OLS, GLS and TLS models for both Eastings-Longitude and Northings-Latitude exhibit a positive linear relationship. This indicates that models for Eastings and Longitude as well as Northings and Latitude all move in the same direction. That is, a unit increase in Longitude increases the Easting component by 110932.46313 having a constant of 1499298.88761 for the OLS. In the case of GLS and TLS, there was a unit increasing of 110932.46220 and 110933.24579 with a constant value of 1499298.87783 and 1499307.31603 respectively. The same situation was also observed for the Northings and Latitude models determined (Table 2).

The FFBPNN models accepted for predicting cartesian planimetric coordinates consist of three layers; thus, 1 input layer, 1 hidden layer and 1 output layer. With reference to Tables 1 and 2, it could be

Table 1 Models determined for eastings and longitude.

Method	Models developed
OLS	E = 110932.46313 (Longitude) + 1499298.88761
GLS	E = 110932.46220 (Longitude) + 1499298.87783
TLS	E = 110933.24579 (Longitude) + 1499307.31603
FFBPNN	2-16-1
TLS-FFBPNN	4-5-1

Table 2 Models determined for northings and latitude.

Method	Models developed
OLS	N = 110325.28210 (Latitude) + 1894.85001
GLS	N = 110325.28213 (Latitude) + 1894.85012
TLS	N = 110325.52889 (Latitude) + 1893.07289
FFBPNN	2-8-1
TLS-FFBPNN	4-4-1

seen that the optimum FFBPNN architecture has 2 input neurons corresponding to 2 selected inputs (longitude, latitude), hidden neurons of 16 and 8 with 1 output neuron of Eastings and Northings respectively. However, the proposed hybrid model (TLS-FFBPNN) consist of 4 inputs namely (latitude, longitude, residuals from TLS, TLS predicted coordinates), hidden neurons of 5 and 4 with 1 output neuron of Eastings and Northings respectively.

6.2 Prediction

A closer look at Tables 3 and 4, it is evident that both the OLS, GLS and TLS produced identical results which were closely related to the existing data. This clearly shows that there is not much difference between the three least square approaches. However, the predictions from the FFBPNN showed significant improvement in the predicted coordinate values compared to the least square methods. Nevertheless, there was substantial improvement in the TLS-FFBPNN model predicted values compared with the other methods.

6.3 Model Performance Evaluation

The validity of the models developed was further

accessed based on the performance criteria index (PCI) stipulated in section 5. The MSE and the RMSE values for each method in relation to their respective models were calculated and the results are shown in Tables 5 and 6 below. The MSE and RMSE are useful when comparing different techniques applied to the same

dataset. Hence, the MSE and RMSE was used in this study as a criterion to measure the goodness of fit of the various models developed. The closer the MSE is to 0 the better the prediction capabilities of the model. On the basis of the results (Table 5), the MSE values obtained for OLS, GLS, TLS were identical and was in

Table 3 Predicted northing coordinate values.

EXISTING NORTHING (m)	PREDICTED NORTHING COORDINATES (m)				
	OLS	GLS	TLS	FFBPNN	TLS-FFBPNN
796594.03	796593.45	796593.45	796593.45	796593.96	796594.02
796970.81	796972.31	796972.31	796972.31	796970.71	796970.69
796147.42	796147.52	796147.52	796147.52	796147.37	796147.43
796502.50	796502.32	796502.32	796502.32	796502.44	796502.49
796393.86	796393.87	796393.87	796393.87	796393.86	796393.86
796213.96	796214.04	796214.04	796214.04	796214.05	796213.98
795818.81	795818.86	795818.86	795818.86	795818.63	795818.80
796881.79	796880.96	796880.96	796880.96	796881.32	796881.79
796963.26	796965.03	796965.03	796965.03	796963.66	796963.13
796906.12	796905.34	796905.34	796905.34	796906.17	796906.12
796685.47	796684.69	796684.69	796684.69	796685.68	796685.47
796100.84	796101.07	796101.07	796101.07	796100.74	796100.86
795822.14	795822.39	795822.39	795822.39	795822.11	795822.14
796510.30	796510.05	796510.05	796510.05	796510.56	796510.29
796823.82	796822.82	796822.82	796822.82	796823.68	796823.81
796991.42	796991.62	796991.62	796991.62	796991.67	796991.33
796967.81	796968.89	796968.89	796968.89	796967.81	796967.65
796685.59	796684.80	796684.80	796684.80	796685.58	796685.56
796980.24	796980.14	796980.14	796980.14	796980.21	796980.16
795996.81	795997.03	795997.03	795997.03	795996.75	795996.81
796666.56	796665.83	796665.83	796665.83	796666.58	796666.56
796968.35	796969.88	796969.88	796969.88	796968.33	796968.23
795846.76	795846.88	795846.88	795846.88	795846.58	795846.76
796375.90	796375.89	796375.89	796375.89	796375.95	796375.90
796450.97	796450.80	796450.80	796450.80	796450.88	796450.96
796963.67	796964.81	796964.81	796964.81	796963.64	796963.58

the range of 0.74 m. This means that the models (OLS, GLS, TLS) could predict observations within the accuracy of 0.74 m. In the case of the FFBPNN, the 0.14 m obtained shows that the model predictions could only deviate from the existing observations by not more than 0.14 m. However, the hybrid approach (TLS-FFBPNN) produced values that are in close

agreement with the measured data. This was based on the MSE value of 0.000900 m. It could be seen that the TLS-FFBPNN gave better prediction capabilities among all the methods applied. With reference to Table 6, the MSE values obtained by each least square methods were identical with a prediction accuracy of 0.6 m. The FFBPNN MSE of 0.03 m shows the extent

to which the model fit the observational data. The TLS-FFBPNN MSE value of 0.004 m act as a measure of how well the model explain a given sets of observations.

The RMSEs shown in Tables 5 and 6 above were used in the evaluation process because the RMSEs are

sensitive to even small errors and can determine the quality of a model, making it good in comparing small changes between predicted and observed differences in models. It is evident that the RMSE (Tables 5 and 6)

Table 4 Predicted easting coordinate values.

EXISTING EASTING (m)	PREDICTED EASTING COORDINATES (m)				
	OLS	GLS	TLS	FFBPNN	TLS-FFBPNN
304470.199	304470.842	304470.842	304470.840	304470.277	304470.198
304623.270	304622.819	304622.819	304622.819	304623.255	304623.271
304588.533	304587.210	304587.210	304587.209	304587.480	304588.530
304698.670	304697.255	304697.255	304697.255	304698.130	304698.701
304815.158	304814.067	304814.067	304814.068	304815.290	304815.183
305086.147	305085.962	305085.962	305085.965	305086.147	305086.067
305393.409	305394.687	305394.687	305394.692	305393.547	305393.488
304590.688	304589.539	304589.540	304589.539	304590.632	304590.674
304622.893	304621.488	304621.488	304621.488	304622.006	304622.883
304751.579	304750.281	304750.281	304750.281	304751.668	304751.599
304893.491	304892.607	304892.607	304892.609	304893.339	304893.485
305039.514	305039.149	305039.149	305039.151	305039.437	305039.429
305215.634	305216.086	305216.086	305216.090	305215.000	305215.654
305393.275	305394.576	305394.577	305394.581	305393.201	305393.264
304323.571	304324.522	304324.522	304324.519	304323.441	304323.572
304415.356	304416.152	304416.152	304416.150	304415.553	304415.359
304251.031	304251.861	304251.861	304251.858	304250.956	304251.041
304244.736	304245.538	304245.538	304245.535	304244.821	304244.735
304411.266	304412.048	304412.048	304412.046	304411.416	304411.265
304494.100	304494.803	304494.803	304494.802	304494.256	304494.098
304552.457	304552.821	304552.821	304552.820	304553.124	304552.451
304619.673	304619.713	304619.713	304619.713	304620.166	304619.655
304512.787	304513.218	304513.218	304513.217	304512.762	304512.791
304588.770	304588.985	304588.985	304588.984	304588.736	304588.765
304612.633	304612.613	304612.614	304612.613	304612.536	304612.631
304637.892	304637.684	304637.684	304637.684	304638.271	304637.880

Table 5 Statistical indicators for easting-longitude model.

PCI	METHODS				
	OLS (m)	GLS (m)	TLS (m)	FFBPNN (m)	TLS-FFBPNN (m)
MSE	0.746325	0.746322	0.746279	0.139851	0.000900
RMSE	0.863901	0.863900	0.863875	0.373966	0.030002
η	0.999992	0.999992	0.999992	0.999999	1.000000
D	0.999998	0.999998	0.999998	1.000000	1.000000
MID	0.998441	0.998441	0.998442	0.999484	0.999964

Table 6 Statistical indicators for northing-latitude model.

PCI	METHODS				
	OLS (m)	GLS (m)	TLS (m)	FFBPNN (m)	TLS-FFBPNN (m)
MSE	0.575834	0.575885	0.576120	0.026614	0.003570
RMSE	0.758837	0.758871	0.759025	0.163138	0.059749
η	0.999996	0.999996	0.999996	1.000000	1.000000
D	0.999999	0.999999	0.999999	1.000000	1.000000
MID	0.999175	0.999175	0.999176	0.999832	0.999948

obtained for Eastings-Longitude and Northings-Latitude models identified the variations in observed and predicted values. Visual inspection of Tables 5 and 6 shows that the TLS-FFBPNN attained the smallest RMSE value among the methods. This signify that the TLS-FFBPNN gave better model performance than the other methods because the closer the RMSE value is to 0 the better the prediction accuracy of the model.

The Nash-Sutcliffe Efficiency Index (η) is a commonly used reliable statistic for measuring model efficiency by comparing predicted values to its corresponding observed values. The (η) range from $-\infty$ to 1 with improved model performance approaching 1. It can be inferred from Tables 5 and 6 that the TLS-FFBPNN based model attained the optimum value of 1 for both models unlike the FFBPNN where 0.999999 was achieved for the Easting-Longitude and 1 for the Northing-Latitude model. This shows that the TLS-FFBPNN predictions are closest to the existing coordinates hence performing better than other models. Nevertheless, the OLS, GLS, TLS and FFBPNN indicated good model performance as their respective values were approaching 1 (Tables 5 and 6).

The index of agreement (D) takes values from 0 to 1 with higher index values indicating that the modelled values have better agreement with the observations. Observation of Tables 5 and 6 show good model performance for all models. However, the TLS-FFBPNN and FFBPNN further showed their superiority over the least square methods based on their respective D values. Although, the D values obtained are better indications of the efficiency of the respective

models, research have shown that relatively high values could be obtained even for a poor model fit. Also, D is sensitive to extreme values due to the square differences in the mean square error in the numerator. In addition, the presence of outliers in the dataset may lead to relatively higher values of D due to the squaring of the difference term [46]. Based on these reasons the authors applied the Modified Index of Agreement (MID) as part of the PCIs. The results for MID in Tables 5 and 6 further corroborate the above mentioned deficiencies in the D. The MID varies from 0 to 1 with higher values indicating a better fit of the model. With this in mind, it can be stated that the TLS-FFBPNN models are better than the other models.

6.4 Model Selection Criteria

In order to further confirm the superiority of the hybrid approach (TLS-FFBPNN) as the optimum model in predicting 2D cartesian coordinates over the FFBPNN and the least square methods (OLS, GLS, TLS) the Bayesian Information criterion (BIC) was applied. Also, BIC tends to favour models with fewer parameters because its penalty term is smaller compared to other criteria. The BIC is represented mathematically as

$$BIC = n * \ln\left(\frac{SS_E}{n}\right) + K * \ln(n)$$

where n denotes the number of observations, SS_E the sum of squares of residuals, and K is the penalty term for the number of parameters.

The results in Table 7 illustrates that the BIC selected both the TLS-FFBPNN models of Northing-Latitude

Table 7 Model selection results.

METHOD	BIC VALUES	METHOD	BIC VALUES
OLS	-1.091247	OLS	-7.834144
GLS	-1.091342	GLS	-7.831833
TLS	-1.092840	TLS	-7.821243
FFBPNN	-44.630483	FFBPNN	-87.767933
TLS-FFBPNN	-175.820946	TLS-FFBPNN	-139.996336

and Easting-Longitude as the optimum model for 2D cartesian coordinates prediction. The reason is that the TLS-FFBPNN model obtained the least value of BIC among the other models.

7. Conclusion

Coordinate prediction is crucial in the surveying and mapping industry especially in developing countries like Ghana where a non-geocentric coordinate system and datum are still used. It is worth knowing that coordinates determined for any geodetic application is used by geospatial and non-geospatial professionals for various purposes. For this reason, in order to facilitate setting a standard in practice for coordinate prediction especially in Ghana, a hybrid approach of Total Least Squares-Feed Forward Back Propagation Neural Network (TLS-FFBPNN) is proposed to be used over conventional techniques (Ordinary Least Squares, Total Least Squares, General Least Squares) particularly within local geodetic networks. Hence, building an effective and accurate prediction model has become a major research area for both academia's and field practitioners in recent times. The proposed hybrid method utilizes the prediction tool capabilities of both the total least squares and artificial neural network architecture of a multilayer perceptron. Using this methodology, researchers and surveyors will be able to predict two-dimensional (2D) coordinates using readily available global geodetic data. This methodology can stir up an intelligent front end and also add a whole new dimension to the usage of artificial neural network in the surveying and mapping industry. It further proves that there is a relationship, no matter how complex in nature,

between global geodetic coordinates and 2D cartesian coordinates. It was also shown from the study that a carefully designed artificial neural network is capable to predict 2D coordinates of Eastings and Northings with accuracies comparable to actual measurements than the conventional techniques. The final hybrid prediction models summary in this study have shown that the models gives more than 99% overall accuracy for the prediction of the 2D cartesian coordinates. This has further shown that the computational efficiency of hybridizing artificial neural network and conventional methods cannot be over stated. To conclude, the proposed hybrid model could be used for cadastral surveys, farm compensation surveys, topographic mapping, engineering surveying and other related mapping activities because of the following considerations: the mean-square error and root mean square error is small indicating no significant over fitting occurring; an overall hybrid model prediction error of not more than 15 cm; satisfied the performance criteria evaluations as well as the Bayesian Information Criterion test.

Acknowledgment

The authors acknowledge the support from University of Mines and Technology (UMaT) and China University of Geosciences (Wuhan) during the process of producing this article.

References

- [1] Mugnier, J. C. 2000. OGP-Coordinate conversions and Transformations including formulae. COLUMN, Grids and Datums. The Republic of Ghana. Photogrammetric Engineering and Remote Sensing, pp. 695-7.
- [2] Ayer, J. 2008. "Transformation models and Procedures for Framework Integration of the Ghana National Geodetic Network." The Ghana Surveyor 1 (2): 52-8.
- [3] Featherstone, W., and Vanicek, P. 1999. "The Role of Coordinate Systems, Coordinates and Heights in Horizontal Datum Transformations." The Australian Surveyor 44 (42): 143-9.

- [4] Kutoglu, S. H., Mekik, C., and Akcin, H. A. 2002. "Comparison of Two Well Known Models for 7-Parameter Transformation." *The Australian Surveyor* 47 (1): 1.
- [5] Thomson, D. B. 1994. "A study of the Combination of Terrestrial and Satellite Geodetic Networks." PhD Thesis, Department of Surveying Engineering, University of New Brunswick, Canada.
- [6] Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J. 1997. *GPS Theory and Practice* (4th Ed.). Springer-Verlag, Wien, New York, pp. 279-303.
- [7] Constantin-Octavian, A. 2006. "3D Affine Coordinate Transformations, Master of Science Thesis in Geodesy No. 3091 TRITA-GIT EX 06-004." School of Architecture and the Built Environment, Royal Institute of Technology (KTH), 100 44 Stockholm, Sweden, pp. 1-7.
- [8] Ayer, J., and Tiennah, T. 2008. "Datum Transformation by the Iterative Solution of the Abridging Inverse Molodensky Formulae." *The Ghana surveyor* 2 (1): 59-66.
- [9] Dzidefo, A. 2011. "Determination of Transformation Parameters between the World Geodetic System 1984 and the Ghana Geodetic Network." Master's Thesis, Department of Civil and Geomatic Engineering, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, pp. 1-97.
- [10] Ziggah, Y. Y., Youjian, H., and Odutola, C. A. 2013. "Determination of GPS Coordinate Transformation Parameters of Geodetic Data between Reference Datums—A Case Study of Ghana Geodetic Reference Network." *International Journal of Engineering Sciences and Research Technology* 2 (4): 956-71.
- [11] Hoar, G. J. 1982. *Satellite Surveying*. Magnavox Advanced Products and Systems Company, 2829 Maricopa Street. Torrance, California, pp. 233-50.
- [12] Schofield, W. 2001. *Engineering Surveying: Theory and Examination Problems for Students*, Butterworth-Heinemann (5th Ed.). Linacre House, Jordan Hill, Oxford OX2 8DP, UK, p. 7.
- [13] Leick, A. 2004. *GPS Satellite Surveying*. John Wiley and Sons, Inc., Hoboken, New Jersey, United States of America, p. 36.
- [14] Sickle, J. V. 2010. *Basic GIS Coordinates*, CRC Press, Taylor and Francis Group, 2nd ed., 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742 USA, p. 7.
- [15] Jekeli, C. 2012. "Geometric Reference Systems in Geodesy." Division of Geodetic Science, School of Earth Sciences Ohio State University, USA, p. 15.
- [16] NIMA. 2000. *World Geodetic System 1984: Its Definition and Relationships with Local Geodetic Systems*, Technical Report No. 8350.2 (3rd Ed.). Amendment 1, National Imagery and Mapping Agency, Washington.
- [17] Ayer, J., and Fosu, C. 2008. "Map Coordinates Referencing and the Use of GPS Datasets in Ghana." *Journal of Science and Technology* 28 (1): 116-27.
- [18] Deakin, R. E. 2004. "The standard and Abridged Molodensky Coordinate Transformation Formulae." Department of Mathematical and Geospatial Sciences, RMIT University, pp. 1-21.
- [19] Poku-Gyamfi, Y., and Schueler, T. 2008. "Renewal of Ghana's Geodetic Reference Network." In: 13th FIG Symposium on Deformation Measurement and Analysis, 4th IAG Symposium on Geodesy for Geotechnical and Structural Engineering, LNEC, LISBON, pp. 1-9.
- [20] Ziggah, Y. Y., Youjian, H., Amans, O. C., and Kumi-Boateng, B. 2012. "Regression Models for 2-Dimensional Cartesian Coordinates Prediction: A Case Study at University of Mines and Technology (UMaT), Tarkwa-Ghana." *International Journal of Computer Science & Engineering Survey* 3 (6): 61-79.
- [21] Odutola, A. C., Beiping, W., and Ziggah, Y. Y. 2013. "Testing Simple Regression Model for Coordinate Transformation by Comparing Its Predictive Result for Two Regions." *Academic Research International* 4 (6): 540-50.
- [22] Dawod, G. M., Mirza, M. N., and Al-Ghamdi, K. A. 2011. "Simple Precise Coordinates Transformations for Geomatics Applications in Makkah Metropolitan Area, Saudi Arabia." In: *Bridging the Gap between Cultures FIG Working Week*, pp. 18-22.
- [23] Hollerbach, J., and Nahvi, A. 1997. *Experimental Robotics IV, Lecture Notes in Control Information Sciences*. Springer Berlin Heidelberg publication, pp. 274-82.
- [24] Yang, T. 1997. "Total Least Squares Filter for Robot Localization." In: *Digital Signal Processing Proceedings of 13th International Conference*, Santorini, pp. 1-2.
- [25] Markovsky, I., Sima, D. M., and Huffel, S. V. 2009. "Generalization of the Total Least Squares Problem." *Advanced Reviewed Article*, pp. 1-6.
- [26] Amiri-Simkooei, A., and Jazaeri, S. 2012. "Weighted Total Least Squares Formulated by Standard Least Squares Theory." *Journal of Geodetic Science*, pp 1-12.
- [27] Tierra, A., Dalazoana, R., and De Freitas, S. 2007. "Using an Artificial Neural Network to Improve the Transformation of Coordinates between Classical Geodetic Reference Frames." *Computers and Geosciences* 34 (2008): 181-9.
- [28] Turgut, B. 2010. "A Back-Propagation Artificial Neural Network Approach for Three-Dimensional Coordinate Transformation." *Scientific Research and Essays* 5 (21): 3330-5.
- [29] Gullu, M. 2010. "Coordinate Transformation by Radial Basis Function Neural Network." *Scientific Research and Essays* 5 (20): 3141-3146.

- [30] Gullu, M., Yilmaz, M., Yilmaz I., and Turgut, B. 2011. "Datum Transformation by Artificial Neural Networks for Geographic Information Systems Applications." In: Ayvaz M. (ed.), *Proceedings of the International Symposium on Environmental Protection and Planning: Geographic Information Systems (GIS) and Remote Sensing (RS) Applications (ISEPP)*, Izmir, Vol. 13, No.19, doi: 10.5053/ise, pp. 1-6.
- [31] Hu, W., Sha, Y., and Kuang, S. 2004. "New Method for Transforming Global Positioning System Height into Normal Height Based on Neural Network." *Journal of Surveying Engineering* 130 (1): 36-39, doi: 10.1061/(ASCE)0733-9453(2004)130:1(36).
- [32] Ning, G., and Cai-yun, G. 2010. "Combining the Genetic Algorithms with BP Neural Network for GPS Height Conversion." In: *International Conference on Computer Design and Applications, IEEE*, Vol. 2, pp. 404-8.
- [33] Pikridas, C., and Fotiou, A. 2011. "Estimation and Evaluation of GPS Geoid Heights Using Artificial Neural Network Model." *Applied Geomatics* 3: 183-7.
- [34] Liu, S., Li, J., Liu, S., and Wang, S. 2011. "A Hybrid GPS Height Conversion Approach Considering of Neural Network and Topographic Correction." In: *International Conference on Computer Science and Network Technology, IEEE*, pp. 2108-11.
- [35] Baabereyir, A. 2009. "Urban Environmental Problems in Ghana: Case Study of Social and Environmental Injustice in Solid Waste Management in Accra and Sekondi-Takoradi." PhD Thesis, Department of Geography, University of Nottingham, UK, p. 97.
- [36] Nievorgelt, Y. 1994. "Total least Squares: State of the Art Regression in Numerical Analysis." *Society of Industrial and Applied Mathematics* 36 (2): 256-8.
- [37] Ghilani, C. 2010. *Adjustment Computations: Spatial Data Analysis*. John Wiley and Sons Inc., New York, pp. 464-70.
- [38] Golub, G. H., and Van Loan, C. F. 1980. "An Analysis of the Total Least Squares Problem." *SIAM Journal on Numerical Analysis* 17 (6): 883-93.
- [39] Sordo, M. 2002. *Introduction to Neural Networks in Healthcare*, Open Clinical, p. 3.
- [40] Handhel, A. M. 2009. "Prediction of Reservoir Permeability from Wire Logs Data Using Artificial Neural Networks." *Iraqi Journal of Science* 50 (1): 67-74.
- [41] Haykin, S. 2007. *Neural Networks: A Comprehensive Foundation* (3rd Ed.). Prentice Hall, Inc. Upper Saddle River, New Jersey, p. 842.
- [42] Konaté, A. A., Pan, H., Khan, N. and Ziggah, Y. Y. 2015. "Prediction of Porosity in crystalline rocks using artificial Neural Networks: An example from the Chinese Continental Scientific Drilling Main Hole." *Studia Geophysica et Geodaetica* 59 (1): 113-6.
- [43] Hornik, K., Stinchcombe M., and White, H. 1989. "Multilayer Feed Forward Networks Are Universal Approximators." *Neural Networks* 2: 359-66.
- [44] Hagan, M. T., and Menhaj, M. B. 1994. "Training Feed Forward Techniques with the Marquardt Algorithm." *IEEE Trans Neural Network* 5 (6): 989-93.
- [45] Nocedal, J., and Wright, J. S. 1999. *Numerical Optimization*. Springer Verlag New York, Inc, pp. 250-71.
- [46] Willmott, C. J. 1981. "On the Validation of Models." *Physical Geography* 2 (2): 184-94.