International Journal of *Geo-Information*

MDPI

*Article*

# Short-Term and Long-Term Forecasting for Changing the Position of Points in the space by using Artificial Neural Networks

**Eleni-Georgia Alevizakou [1], George Siolas [2] and George Pantazis [1,*]**

[1] School of Rural and Surveying Engineering, National Technical University of Athens, Zografou, Athens, Greece ; elalevizakou@gmail.com (E.G.A); gpanta@central.ntua.gr (G.P.)

[2] School of Electrical and Computer Engineering, National Technical University of Athens, Zografou, Athens, Greece; gsiolas@islab.ntua.gr

* Correspondence: gpanta@central.ntua.gr; Tel.: +30-210-7722696

Received: date; Accepted: date; Published: date

**Abstract:** Forecasting is one of the most growing areas in most sciences attracting the attention of many researchers for more extensive study. Therefore, the goal of this study is to develop an integrated forecasting methodology based on an Artificial Neural Network (ANN), which is a modern and attractive intelligent technique. The final result is to provide short-term and long-term forecasts for changing the position of points, i.e. the displacement or deformation of the surface they belong to. The motivation was the combination of two thoughts, the introduction of the concept of forecasting in most scientific disciplines (e.g. economics, medicine) and the desire to know the future behavior and location of a construction or an area of the natural earth surface. This methodology was designed to be accurate, stable and general in deferent kind of geodetic data. The basic procedures, which are involved, are the definition of the problem, the data pre-processing, the definition of the most suitable ANN, the evaluation using the proper criteria and the production and use of forecasts. The methodology gives great importance to the stages of the pre-processing and the evaluation. The forecasting intervals are also emphasized. Finally the most appropriate ANN is presented and evaluated also it is proved that the use of ANNs in order to make short-term and long-term forecasts gives more accurate forecasts compared to other conventional-statistical forecasting methods.

**Keywords:** Artificial Neural Networks; NAR; NARX; Geodesy; Position Changing; Short-Term Forecasting; Long-Term Forecasting; Data Mining.

## 1. Introduction

The introduction should briefly place the study in a broad context and highlight why it is important. It should define the purpose of the work and its significance. The current state of the research field should be reviewed carefully and key publications cited. Please highlight controversial and diverging hypotheses when necessary. Finally, briefly mention the main aim of the work and highlight the principal conclusions. As far as possible, please keep the introduction comprehensible to scientists outside your particular field of research. References should be numbered in order of appearance and indicated by a numeral or numerals in square brackets, e.g., [1] or [2,3], or [4–6]. See the end of the document for further details on references.

In the 75 years since the introduction of ANNs basically in Neuroscience [1], their use has expanded to numerous fields, including Economics, Mathematics, Meteorology, Clinical Medicine, Environmental Area etc. Specifically in recent years, ANNs have found a number of applications in the area of recognition and classification [2], water quality [3,4], meteorology [5], politics [6], medical diagnostics [7] etc. Therefore, it appears that the ANNs are used in applications for classification, clustering, association, regression, modeling and forecasting.

Moreover, ANNs have also been introduced to solve various problems in the scientific field of Geodesy. They have been used in the regional mapping of the Geoid [8], sea level prediction [9], coordinate transformation [10], UTC time correction [11, 12] etc.

In brief, an ANN is an artificial intelligence technique that mimics the human brain's biological neural network. It is defined as a system of simple processing elements called "artificial neurons", which are typically organized in layers, so that each ANN includes: an input layer, hidden layers (which can have n neurons linked in different ways to the other hidden layers or to the output layer) and an output layer. The main feature of ANNs is their inherent capacity to learn, a key component of their intelligence. Learning is achieved through training, a repetitive process of gradual adaptation of the network's parameters to the values required to solve a problem [13-19].

Before analyzing the proposed forecasting methodology it is essential to mention, the worldwide used forecasting techniques. The techniques, which are used in order to produce a forecast, can be separated in two main categories: the judgmental forecasting techniques and the statistical forecasting techniques [20]. However, there is another alternative way of separate these techniques in: the judgmental ones, the extrapolative, the econometric and finally the non-linear computer-intensive ones. Also the above mentioned techniques can further split in some other corresponding subcategories [21]:

Finally, they can also be distinguished, according to the method which is used in order to produce the forecast, in the following categories as the Quantitative forecasting (as Time series methods, Causal relationship or explanatory methods and Artificial intelligence methods), Qualitative or judgmental forecasting (as Individual methods and Committee methods) and Technological forecasting (as Exploratory methods and Normative methods).

After the necessary bibliographic research, it was found that in the scientific area of Geodesy the main interest of the Geodesists is focused on modeling and monitoring of such a phenomenon rather than on forecasting. Thus, it was emerged that there is no methodology, which could be used in order to forecast the position change of a point on an artificial structure or on the earth surface. The majority of the proposed techniques or models are focused on monitoring the point's position.

The aim of this paper is to present an automatic and integrated methodology in order to forecast the position of a point, which belongs to a construction or to the earth surface, in a future time. This forecast concerns both the long-term future as well as the short-term.

The use of the term "forecasting" is more accurate than the term "prediction", because the first one is more objective rather than the second, which is more subjective and intuitive. According to Lewis-Beck "forecasting can be seen as being a subset of prediction, with all forecasts being predictions, but not all predictions being forecasts" [22].

The proposed methodology is based on the combination of two disciplines: the Science of Knowledge Discovery in Databases (KDD) for Data Mining and the subset of the Artificial Intelligence Science, namely the ANNs.

Essentially, KDD is a science that emphasizes in the data analysis process, but it does not deal with the methods which will be used to perform this analysis [23- 25] and ANN is considered as one of the modern mathematical - computational methods which are used to solve a large number of different kinds of problems.

## 2. Description of the methodology

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation as well as the experimental conclusions that can be drawn.

The methodology of this research uses an Artificial Intelligence method, namely an ANN, so it is a quantitative based methodology. It is based on a method which is modern but at the same time accurate, stable and general so that it could be used in different but similar problems.

The most important step is the definition of the examined forecasting problem. This means that the aim of the forecasting and the use of the results need to be clear. In addition, it is useful to define the timescale (one-step-ahead forecasting or multi-step-ahead forecasting, short-term or long term) of the desired forecasting.

The problem, which the methodology deals with, is the *forecasting of the position of a point in a future time*. It is important to mention that the desired forecasting depends on the available data. No change in the position of a point of the order of a few mm can be predicted if no such movement is recorded in the existing data, i.e. if the geodetic instruments and methods, which collect the data, cannot detect such a displacement.

The basic steps of the methodology are: Problem definition → finding the available data in time series form → data pre-processing → ANN definition → evaluation → production and use of forecasts.

### *2.1. Data pre-processing*

This stage is of a great importance since it has a significance role in the accuracy of the results. Through this process the mechanism that produces the timeseries can be found.

The first test is to identify if **the timeseries follows a known distribution**. This test can be performed either by the Kolmogorov-Smirnov test [26] or its variation, the Anderson-Darling Gof (goodness of fit) test [27].

If the timeseries does not follow a known distribution, then it is possible to transform it. The transformation that can be used for this purpose is the Box-Cox [28, 29], so that the abnormal timeseries becomes almost normal. The Johnson transformation can also be used as an alternative method [30]. Additionally, the **statistical analysis** of the timeseries (statistics) is then used to find statistical measures.

The computational measures which are used are: the mean, the maximum and minimum values, the standard deviation, the covariance, the autocovariance, the correlation and the autocorrelation.

The next step is to find the **qualitative features** and the **timeseries analysis of the major components**. At this stage the existence of stationarity in the timeseries or any trend, seasonality, cyclicality or randomness is controlled. To convert the timeseries from non stationary to stationary, the diffusion method is used to calculate either the first or the second differences.

Finally, it is sometimes useful to **analyze the frequency** of the timeseries in order to find some information that is not immediately apparent in the original signal (timeseries). Then there is the possibility to use some mathematical transformations such as the Fourier Transformation, the Fourier Short - Time Transformation and the Wavelet Transformation, with the ultimate goal of retrieving some additional useful information.

It is then necessary to manage:

- the missing values presented in the timeseries,
- the possible zero values in the event that it is known that they should not exist,
- the duplicate recordings that may exist and
- any unusual observations (outliers) [31].

With regard to the missing values likely to be presented in the timeseries, if they involve long periods of incomplete data, an effort should be made to find these values from other sources. If there are individual cases (not continuous time) and when no seasonal behavior has been observed from the plot of the timeseries, this missing value (or an unexpected zero value) is defined as the mean of the previous and the next value. In the case of duplicate recordings, they must be detected and deleted.

The most complex case is the existence of outliers. The difficulty of this case lies in detecting these values not in an optical way but in a mathematical way that can be automated. Once those individual values are found, they are replaced by the method used in the previous cases, if that is possible. A definition of the outliers is given as follows [32]:

"*An outlier is a residual which, according to some test rule, is in contradiction to assumptions on the stochastic properties of the residuals*"

There are several methods for finding outliers such as the Z-score, modified z-score, tucked method, boxedot, MADe method, Median rule, and generalized ESD test [33]. Each test should be checked by the researcher for its suitability for the problem under consideration. A proposed method is the following [34]:

- the first differences of the z data are calculated
- the upper (Uz) and lower (Lz) quartile of the series of differences is calculated
- a point is considered as an outlier if the following is valid:
  $z < Lz - 1.5 \cdot (Uz - Lz)$ or $z > Uz + 1.5 \cdot (Uz - Lz)$

Another way of identifying outliers in a timeseries is to determine the standard deviation (SD), ± 3s, which includes 99% of the measurements, excluding only the values that can be considered as outliers [35].

### 2.2. Prediction interval

When attempting to forecast a value of a phenomenon, the result always has an uncertainty and an ambiguity [36], so it is necessary to define the forecasting interval.

Forecasting interval or prediction interval (PI) is defined as the interval within which the value is expected to occur with a specific probability (1-a) % [37]. It is very important to estimate this interval, in order to finally find the most accurate forecasting method [38].

The interval consists of an upper predictive limit (UPL) and a lower predictive limit (LPL) that defines the interval within the future value will appear with a particular probability [39]. In the international literature, they are referred to either as forecast limits [40] or as prediction bounds [41]. There is a parametric and non-parametric approach for estimating the forecasting interval.

In the parametric approach, which was proposed by Box and Jenkins [42], it is considered that forecasting errors $e_t = D_t - P_t$, meaning the series $\{ e_t \}$, is independent and follows the normal distribution with zero mean and variance $\sigma_e^2$. The forecasting error used concerns data that is not used when creating the model or applying a method, and is often referred as an out-of-sample forecast error. The PI, for a probability (1-a)%, is calculated by the following equation [39, 43]:

$$\hat{x}_N(h) \pm z_{a/2} \cdot \sqrt{\mathrm{var}[e_N(h)]} \quad \text{or} \quad \hat{x}_N(h) \pm z_{a/2} \cdot \sigma_e \tag{1}$$

Where the standard deviation of errors is calculated from the equation :

$$\sigma_e = \sqrt{\sigma_e^2} = \sqrt{\mathrm{var}[e_N(h)]} = \sqrt{\frac{\sum_{i=1}^{N} (e_i - \hat{e})^2}{N-1}} \tag{2}$$

and $\hat{e} = \frac{\sum_{i=1}^{N} e_i}{N}$ (Bibliographically referred as bias).

Thus, the prediction limits are as follows : $UPL = z_{\alpha/2} \cdot \sigma_e$ and $LPL = -z_{\alpha/2} \cdot \sigma_e$.

### 2.3. The design of the forecasting ANN

Since the problem relates to forecasting, so that the concept time information is used, the appropriate ANN is a **Dynamic Neural Network**.

In this work, the dynamic network decided upon by an extensive bibliographic study is a **recurrent or feedback dynamic network (RNN)**, because they can process data sequences, time information in our case. They can simulate two kinds of memory elements, long-term and short-term.

In every kind of ANN, the synaptic weights can be considered as long-term memories. However, if the problem has a temporal dimension, some form of short-term memory must be integrated into the neural network. The simplest way to do this is with time delays.

RNNs also allow non-linear relationships to be found, between the elements of a time series. The most important is that these networks have the advantage of recognizing time patterns regardless of their duration. More specifically; we chose to use non-linear **NAR** autoregressive recurrent networks for forecasting a time step, and a **NARX** non-linear autoregressive with eXogenous inputs for multi-step forecasting.

In the case of NAR, the forecasting of the $\ell(t)$ element of the timeseries is made, using only d past values of that timeseries, according to the equation:

$$\ell(t)=f(\ell(t-1),\ldots,\ell(t-d)) \tag{3}$$

Similarly, in the case of NARX, the d past values of the same timeseries ($\ell$) and also d of another exogenous timeseries ($q$) are used, according to the equation:

$$\ell(t)=f(\ell(t-1),\ldots,\ell(t-d),\ q(t-1),\ldots,q(t-d)) \tag{4}$$

The training of the ANN is done through supervised process based on prior knowledge of the right outputs. For this reason, a percentage of the historical data is retained at the end for evaluation and control of the results.

These networks produce a one-step-ahead forecasting only. In order to predict a multi-step-ahead, a loop is added to the networks. That is, once the first forecast is made, it returns to the network as an input to produce the next forecast, and so on. Therefore, the network based on its own outputs (forecasted values), also predicts the following future values.

In order to select the most suitable final network, after its general architecture has been decided, various tests of ANN's implementation are carried out, with the main objective of finding the one that will achieve the best results for forecasting. The basic principle used is that of validation by varying the hyper-parameters of the network: number of hidden layers, the number of neurons of each hidden layer, the learning algorithm, the activation function of the hidden and exit neurons, and the number of delays. This step also includes all the procedures that are necessary to select the most appropriate ANN.

- Normalization of data.

For this methodology the Min-Max normalization is proposed, in order to normalize the values between height=0.1 and low=0.9 using the following equation.

$$\widehat{\ell}=\text{low}+\frac{\ell-\ell_{\min}}{\ell_{\max}-\ell_{\min}}\cdot(\text{hight-low}) \tag{5}$$

Where $\ell_{\min}$ and $\ell_{\max}$ are the min and the max value of the original timeseries

- Separation of training, testing and validation data sets.

  The available historical data (timeseries) is proposed to be separated in chronological order by using the following percentages: training set 70%, test set 15% and validation set 15%.

- Tests on the selection of the training algorithm, the transfer function, the number of hidden neurons and the hidden levels.

The training algorithm, which is found to be appropriate, is the Bayesian regularization and the training function is the logsig function. In this preliminary study we chose to use only one hidden layer in order to keep complexity low and training times small.

*2.4. Forecasting evaluation*

Estimating the uncertainty of the forecasts is an integral part of the forecasting process. In order to evaluate a forecast, the results obtained must be compared with the actual values, which are already known, for the phenomenon under consideration. To do this, some statistical indicators or other evaluation criteria are used.

The appropriate evaluation criteria are the mean square error (**MSE**) , the root of the mean square error (**RMSE**), the mean absolute error (**MAE**) and the coefficient of determination (**$R^2$**) [44].

The **$R^2$** is used because one of the usual evaluation methods is to perform a linear regression in order to check if the selected method has a good performance during the evaluation phase. This index is interpreted as the percentage of the variance of the values of the dependent variable (i.e. the predicted values) from the values of the independent variable (i.e. the actual-known values).

The **MAE** is a failure index of the forecast by its actual desired value. Absolute values are used, so the forecast direction is not taken into account.

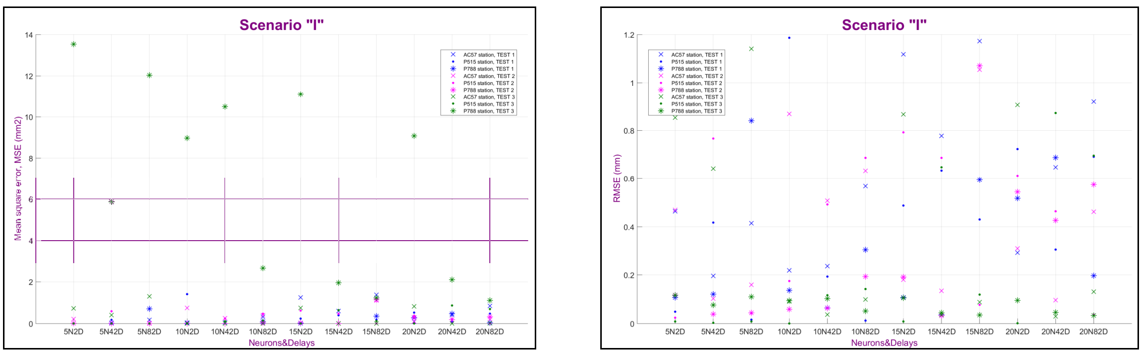
The **MSE** has the possibility to highlight possible extreme deviations. This is the most common index of error that shows the qualitative performance of the forecasting method. Obviously it does not have the same units as the actual values and the forecasts, so instead of this **RMSE** is used. The main difference between using the **MAE** or the **RMSE** is that the **MAE** assigns equal weight to all errors when calculating overall performance, while **RMSE** gives greater weight to errors with greater absolute value. Thus, the **RMSE** value is never lower than the **MAE** value for the same data set. When both metrics are calculated, the **RMSE** is by definition never smaller than the **MAE**.

In order to define the accuracy of the forecasting the mean position change must be calculated. Specifically, a forecast can be considered as accurate if the **MAE** is significantly smaller than the mean predicted position change.

The **$R^2$** index measures the amount of variance in the target variable that can be explained by the model. In some cases, it is necessary to quantifying the error in the same measuring unit of the variable. This problem is solved by the use of the **RMSE**. The issue with the **RMSE**, and with the **MSE**, is that since they square the residuals they tend to be more affected by large residuals. To solve the problem with large residuals **MAE** can be used, as we average absolute value of the residuals is calculated. Taking the square root of the average squared errors has some interesting implications for **RMSE**.

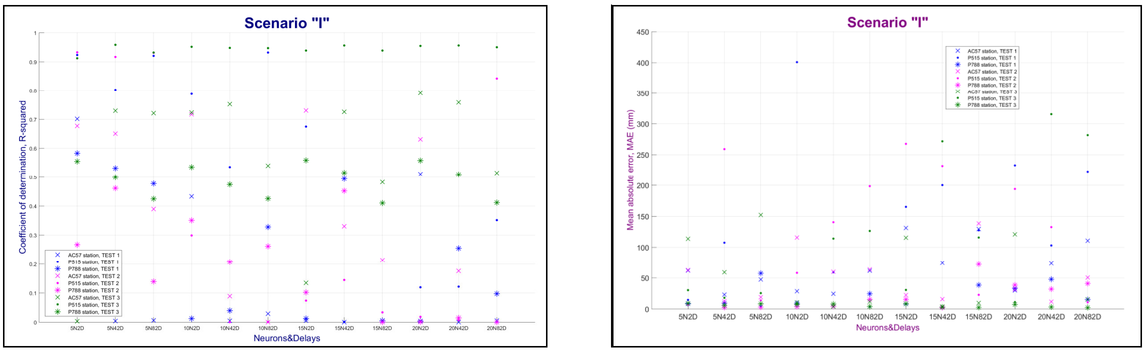
Since the errors are squared before they are averaged, the **RMSE** gives a relatively high weight to large errors. This means the **RMSE** should be more useful when large errors are particularly undesirable. There are cases when the **MAE** may be steady but **RMSE** may increase as the variance associated with the frequency distribution of error also increases [45].

The following figures 1 and 2 present an example of the above mentioned criteria from a case study. The final hyper-parameters are chosen from the examination of those diagrams. Finally, the network, which has the smallest values (for the test set) in these criteria, is selected.




(**a**) (**b**)

**Figure 1.** (**a**) A combined MSE diagram; (**b**) A combined RMSE diagram




(**a**) (**b**)

**Figure 2.** (**a**) A combined $R^2$ diagram; (**b**) A combined MAE diagram

## 3. Discussion

The proposed methodology has been developed to solve the problem both for one-step-ahead and for multi-step-ahead forecasting and therefore expands on the steps which are presented in the figure 3.

To sum up, the ANN resulting as the best for forecasting the change of the points is a non-linear autoregressive NAR recurrent neural network and a NARX non-linear autoregressive with eXogenous inputs in the case of multi-step forecasting. Some preliminary tests on a sample dataset led to the following findings about the ANN design choices and training procedure:

- It consists of one hidden layer and has "logsig" as an activation function of hidden and output neurons.
- Separation of the data is done in chronological order (70%, 15%, 15%) and their normalization in the interval [0.1,0.9]
- The learning algorithm is the Bayesian regularization.
- It is proposed the number of hidden neurons to be 10 and the number of delays 42 for one-step forecasts. While for multi-step forecasts 20 hidden neurons and 2 delays.
- Finally, in cases where time series of geocentric coordinates X, Y, Z (m) are used, it is proposed X Y Z to be used separately as inputs and outputs.

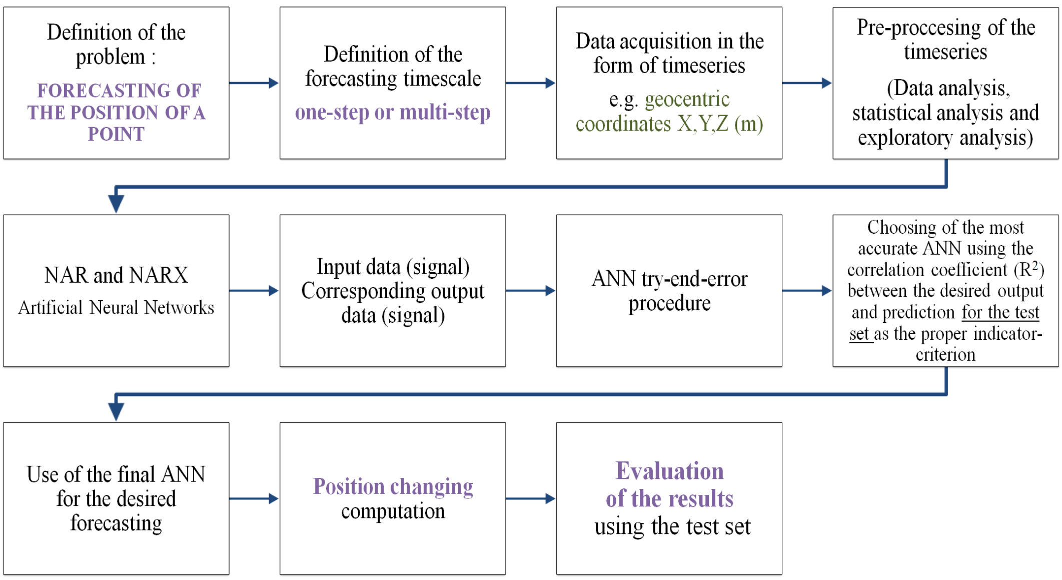



**Figure 3.** Main steps of the proposed forecasting methodology

The best ANNs for one-step and for multi-step forecasting are presented accordingly in the following figure 4.

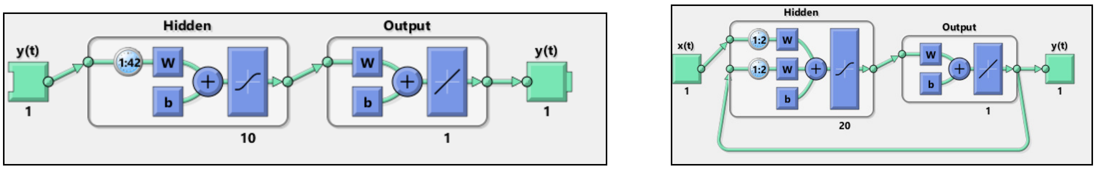



**Figure 4.** The NAR for one-step and NARX for multi-step forecasting suggested networks for the forecasting of points position changing

The number of available data to be used is a very important factor in their training. As the number of data increases, the better the network will be "trained" and will therefore produce better results.

Big data management is very complex and requires special treatment. In such cases it is necessary to use automated techniques rather than optical means as in smaller volume and size data. Despite the difficulties, finding and using large amounts of information is a major asset especially when it comes to forecasting.

During training, special care must be taken to avoid the overfitting phenomenon. If the network is trained for too many epochs and/or in the entire dataset, we could observe small training error but lose the ability to generalize well, i.e. perform well on new, unseen data. Therefore, there is a possibility that there is a limit beyond which, while the training error decreases, the validation error increases. In this situation, the network may have memorized the correct results that arise from the education phase, but has not learned to generalize properly on new data.

As already mentioned, the proposed methodology is based on ANNs. The main drawback of their use is the large number of tests (try-and-error process) that must be carried out with all the possible different combinations of the hyper-parameters.

Therefore, in cases where the data is so many, a representative sample should be selected by which all the experiments will be carried out and, finally, the most suitable ANN will be proposed. The sample could be simple random, systematic or else quasi-random and finally stratified random depending on the available dataset.

Despite this disadvantage, the methodology gives more accurate forecasts compared to other conventional-statistical forecasting methods (i.e. simple mean, simple moving average, simple exponential smoothing, double moving average, Brown's method, Holt's method, ARMA/ARIMA models) [46].

## 4. Conclusions

ANNs have been introduced to the engineering sciences in general and Geodesy in particular. In recent years, an attempt has been made to investigate their use. Several of the geodetic applications in which the ANNs have been used to date have yielded very satisfactory results, giving even better results compared to classical methods for solving these problems. Applications mainly used are local geoid mapping, averaging of the sea level, coordinate conversion to reference systems, mobile mapping models, and so on.

As far as ANNs are concerned, this is a modern and very attractive intelligent technique. For this reason, implementing an ANN to solve any problem is an interesting and original alternative. However, a try-and-error process must always be followed during the implementation process. The more testing done, the more complete the work is to investigate the use of ANNs for any problem. This means that it takes a lot of time and a lot of computational power until one gets the desired result.

The documentation and suggestion of a particular ANN and all its over-parameters arises after each application, which is another significant disadvantage. ANN are a data driven method, so the architecture and hyper-parameters found to be optimal in a dataset, in the general case won't be optimal for another dataset.

This methodology which is based on ANNs is accurate, stable and general and can be used in order to predict the changing of the position of points. It gives better, more accurate, results rather than the classical-conventional methods. It can give forecasts of position change of the order of 2mm with MAE in the order of 0.5mm while the optimum convention method -ARMA model (4,3)- with MAE 1.8mm.

## References


1. McCulloch; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics.* **1943**, 5, 115–133.
2. Yang, C.-C.; Prasher, S.O.; Landry, J.-A.; Ramaswamy, H.S.; Ditommaso, A. Application of artificial neural networks in image recognition and classification of crop and weeds. *Canadian Agricultural Engineering*. **2000**, 42 (3),147-152.
3. Babu, B.V.; Ramakrishna, V; Kalyan Chakravarthy, K. Artificial Neural Networks for Modeling of Adsorption. Second International Conference on Computational Intelligence, Robotics, and Autonomous Systems (CIRAS-2003), Singapore , **2003**.
4. Khalil, B.M.; Awadallah, A.G.; Karaman, H.; El-Sayed A. Application of Artificial Neural Networks for the Prediction of Water Quality Variables in the Nile Delta. *Journal of Water Resource and Protection* **2012**, 4, 388-394. Internet access : http://dx.doi.org/10.4236/jwarp.2012.46044
5. Memarianfard, M.; Hatami; A.M. Artificial neural network forecast application for fine particulate matter concentration using meteorological data. *Global J. Environ. Sci. Manage.* **2017**, 3(3), 333-340. Internet access : http://dx.doi.org/10.22034/gjesm.2017.03.03.010
6. Khaze, S.R.; Masdari, M; Hojjatkhah, S. Application of Artificial Neural Networks in estimating participation in elections. *International Journal of Information Technology, Modeling and Computing (IJITMC).* 2013, 1(3),23-31. Internet access : http://dx.doi.org/10.5121/ijitmc.2013.1301
7. Amato, F.; López, A.; Peña-Méndez, E.; Vaňhara, P.; Hampl, A.; Havel, J. Artificial neural networks in medical diagnosis. *Journal of Applied Biomedicine.* **2013**, 11, 47–58. Internet access : http://dx.doi.org/10.2478/v10136-012-0031-x
8. Veronez, M. R.; de Souza, S.F.; Matsuoka, M.T.; Reinhardt, A.; de Silva, R. M. Regional Mapping of the Geoid Using GNSS (GPS). Measurements and an Artificial Neural Network. *Remote Sensing*. **2011**, 3, 668-683, 2011. Internet access : https://doi.org/10.3390/rs3040668
9. Makarynskyy, O.; Makarynska, D.; Kuhn, M.; Featherstone, W.E., Predicting sea level variations with artificial neural networks at Hillarys Boat Harbour Western Australia. *Estuarine, Coastal and Shelf Science.* **2004**, 61, 351-360. Internet access : https://doi.org/10.1016/j.ecss.2004.06.004
10. Turgut, B.; A back-propagation artificial neural network approach for three-dimensional coordinate transformation. *Scientific Research and Essays*. **2010**, 5(21),3330-3335.
11. Miczulski, W.; Sobolewski L. Influence of the GMDH Neural Network data preparation method on UTC(PL) correction prediction results. *Metrology and Measurement Systems*. **2012**, XIX(1), 123-132.
12. Luzar, M.; Sobolewski, L.; Miczulski, W; Korbicz, J. Prediction of corrections for the Polish time scale UTC(PL) using artificial neural networks. *Bulletin of the Polish Academy of sciences, Technical sciences.* **2012**, 61(3). Internet access : http://dx.doi.org/10.2478/bpasts-2013-0060
13. Minsky, M.L. Steps towards artificial intelligence. In the Proceedings of the Institute of Radio Engineers, 49, 8–30, **1961**.
14. Patterson, D.W. *Artificial Neural Networks : Theory and Applications*. **1996**.Prentice Hall.
15. Haykin, S. *Neural Networks: A Comprehensive Foundation*, second edition. New Jersey, Prentice Hall, **1999**.
16. Basheer, I.A.; Hajmeer M. Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*. **2000**, 43(1), 3-31. Internet access : https://doi.org/10.1016/S0167-7012(00)00201-3
17. Hudson,D.; Cohen M. *Neural Networks and Artificial Intelligence for Biomedical Engineering*. **2000**. The Institute of Electrical and Electronics Engineers.
18. Kriesel D. *A brief Introduction to Neural Network*. **2005**.

19. Haykin, S. *Neural networks and learning Machines*, 3rd edition. Pearson education, 2010. Internet access : http://www.dkriesel.com/en/science/neural_networks
20. Armstrong, J.S.; Green, C.K.; Graefe, A. *Forecasting Principles*. **2010.** University of Pennsylvania, Scholarly Commons.
21. Fildes, R. *Bibliography on business and economic forecasting*. Lancaster Center for Forecasting. Lancaster University Management School. Working paper, **2014**.
22. Lewis-Beck M. Election of Forecasting: Principles and Practice. *The British Journal of Politics and International Relations*. 2005, 7(1), 145-164.
23. Frawley, W.; Piatetsky-Shapiro, G.; Matheus C. Knowledge Discovery in Databases: An Overview. *AAAI, AI Magazine*. **1992**, 13 (3).
24. Fayyad, U.; Piatetsky-Shapiro, G.; Smyth P. From Data Mining to Knowledge Discovery in Databases. **AAAI, AI Magazine**. 1996, 17 (3).
25. Bellazzi, R.; Zupan B. Predictive data mining in clinical medicine: current issues and guidelines. *International journal of medicine informatics Elsevier*. 2008,81-97.
26. Chakravarti; Laha; Roy. *Handbook of Methods of Applied Statistics*. John Wiley and Sons. **1967**, I, 392-394.
27. Stephens, M.A. EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association*. 1974, 69, 730-737.
28. Box, G.E.P.; Cox D.R. An analysis of transformations. *Journal of the Royal Statistical Society*. **1964**, B 26, 211-234.
29. Osborne, J.W. Improving your data transformations: applying the Box-Cox transformation. *Practical Assessment, Research and Evaluation*. **2010**, 15 (12).
30. Kovarik, M.; Sarga L. Process capability indices for non-normal data. *WSEAS transaction on business and economics*. **2014**, 11.
31. Chatfield C. *The Analysis of Timeseries An Introduction*. **2003**. Chapman Hall/CRC, Sixth Edition.
32. Caspary W. Concepts of networks and deformation analysis. **1987**. Monograph No. 11, School of Surveying, University of New South Wales, Australia.
33. Rosner, B. Percentage Points for a Generalized ESD Many-Outlier Procedure. *Technometrics*. **1983**, 25(2), 165-172.
34. Fildes, R.; Hibon, M.; Makridakis, S.; Meade, N. Generalizing about univariate forecasting methods: further empirical evidence. *International journal of Forecasting*. **1998**, 14, 339-358.
35. Wdowinski, S; Bock, Y.; Zhang, J.; Fang P.; Genrich J. Southern California permanent GPS geodetic array: Spatial filtering of daily position for estimating coseismic and postseismic displacements induced by the 1992 Landers earthquake. *Journal of Geophysical Research*. **1997**, 102(B8), 18057-18070.
36. Pinson, P., Estimation of the uncertainty in wind power forecasting, Ph.D. dissertation, Ecole des Mines de Paris, Center for Energy and Processes, Rue Claude Daunesse, Sophia Antipolis, **2006**.
37. Pinson, P.; Kariniotakis G. On-line adaptation of confidence intervals based on weather stability for wind power forecasting. In the Proceeding of the 2004 Global Windpower Conference, Chicago, Illinois (USA), **2004**.
38. Holfmann-Wellenhof, B.; Moritz W.E. *Physical Geodesy*, **2006**. 2nd edition, Springer Wien New York.
39. Chatfield, C., Prediction intervals for time-series forecasting. Department of Mathematical Science. University of Bath, from the book "*Principle of forecasting: A Handbook for Researchers and Practitioners*" editing by Armstrong S.J., Kluwer Academic Publishers, **2001**.
40. Wei, W.W.S. *Timeseries Analysis*. **1990**. Redwood City, CA: Addison-Wesley.
41. Brockwell, P.J.; Davis R.A. *Timeseries: Theory and Methods*. **1991**. Second edition New York: Springer-Verlag, 1991.
42. Box, G.E.P.; Pierce D.A. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of the American Statistical Association*.**1970**, 65, 1509–1526.
43. Turyna, M.; Hrdina Th. Calculating Interval Forecasts. Department of Economics. University of Vienna, **2009**.
44. Alevizakou E.G.; Pantazis G. A Comparative Evaluation of Various Models for Prediction of Displacements. *Applied Geomatics*. 2017, 9(2), 93–103. Internet access : http://dx.doi.org/10.1007/s12518-017-0189-8
45. Willmott, C.J.; Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim.Res*. **2005**, 30, 79-82.

46. Alevizakou E.G. Forecasting of methodology for the forecasting of changing the position of points in the space by using artificial neural networks. PhD dissertation, NTUA, Athens, **2017**.