A HYBRID APPROACH TO MODEL NONSTATIONARY SPACE-TIME SERIES

T. Cheng [a, *], J.Q. Wang [b], X. Li [b], W. Zhang [b]

[a] Dpartment of Civil, Environmental and Geomatic Engineering, University College London, Gower Street, WC1E 6BT London, United Kingdom {tao.cheng@ucl.ac.uk}
[b]School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, P R China-(wangjiaq@mail2.sysu.edu.cn;lixia@mail.sysu.edu.cn; zhangwen@mail.szbqts.gov.cn)

Commission VI, WG VI/4

KEY WORDS: Space-Time Series; Forecasting; Stationarity; Autocorrelations; Artificial Neural Network; STARMA;

ABSTRACT:

In recent years the Space-Time Autoregressive Moving-Average (STARMA) model family has been proven a useful tool in modelling multiple time series data that correspond to different spatial locations (which are called space-time series). The STARMA model family is a statistical inductive model that can be used to describe stationary (or weak stationary) space-time processes. However, in real applications STARMA model can not be applied directly because of the non-stationary nature of most space-time processes. To overcome this deficiency, a novel approach to model non-stationary space-time series is proposed in this study. It uses artificial neural network (ANN) to develop a non-parametric, robust model to extract the large-scale nonlinear space-time structures, then uses STARMA model to extract the small-scale stochastic space-time variations. The proposed approach has been applied to the forecasting of china annual average temperature at 137 international meteorological stations in China. The experimental results demonstrate that the forecasting using ANN+STARMA method obtains better forecasting accuracy than using conventional pure STARMA method. It proves the mixture of the data-driven ANN and the model-driven STARMA can become a very useful and efficient tool for space-time modelling and prediction of environment data with temporal and spatial dependence.

1. INTRODUCTION

The Space-Time AutoRegressive Moving-Average (STARMA) models have gained widespread popularity in many domains, including imaging, transport, business and economics, and hydrology, etc. For example, Pace et al. (1998) introduced a space-time autocorrelation (STAR) model that predicted house prices by capturing the effect of both spatial and temporal information on real estate prices. Using data on housing prices, they showed that substantial benefits could be obtained by modelling both the data's spatial and temporal dependence. The improved performance of the STAR model was confirmed by comparing it with the traditional indicator-based model. Moreover, Kamarianakis and Prastacos (2005) applied space-time autoregressive integrated moving average (STARIMA) methodology to represent traffic flow patterns. Traffic flow data are in the form of a spatial time series, and are collected at specific locations at constant intervals of time. The experiment, in the centre of the city of Athens, Greece, showed that the STARIMA model can be used for the short-term forecasting of space-time stationary traffic-flow processes, and to assess the impact of traffic-flow changes on other parts of a road network. More recently, Crespo et al. (2007) implemented an image sequence prediction system that offers the most probable image for a given series, using methods based on the space-time autocorrelative (STAR) model. The imaging neighbourhood structure in space and time is obtained from the great number of testing that are made. Comparison with the observed real images shows that the prediction is very successful. All these studies have demonstrated that STARMA can obtain better applications in modelling space-time dynamic processes when the processes can be treated as stationary.

However, in real applications most space-time processes are non-stationary and also are nonlinear. Detrending and differencing are most common approaches to handle non-stationary in spatial data analysis and time series respectively, but they are difficult to make space-time process stationary. Thus, it's vital to subtract space-time patterns out prior to fitting STARMA model.

In spatial data analysis a spatial process z can be decomposed into two parts: large-scale deterministic spatial variation μ plus small-scale stochastic spatial variation e (Haining, 2003; Kanevski and Maignan, 2004), then a space-time process also can be given by:

$$z_i(t) = \mu(i,t) + e_i(t), \tag{1}$$

where $i \in D \subset R^d$ and $t \in T \subset R$; $z_i(t)$ represents multiple time series of spatially location i data; the function $\mu(i,t)$ represents the space-time patterns that explain large-scale nonlinear space-time variations and the residual term $e_i(t)$ is a zero mean space-time correlated error that explains small-scale stochastic space-time variations. The key idea proposed method is to use ANN to develop a non-parametric, robust model for the large-scale nonlinear space-time structures and then to use STARMA model for the analysis of residuals—modeling of small-scale stochastic space-time variations. The objective of the integrated models is two aspects: from one side ANN efficiently solves problems of space-time non-stationary by modeling large-scale nonlinear space-time variation, from

* Corresponding author.

another side spatial weights matrix in the STARMA model is built based on variogram function, which can exactly express spatio-temporal dependence and variance of environmental data.

The paper is organized as follows. Section 2 introduces the procedure and principles of the ANN and the STARMA for modelling non-stationary space-time series, which are continuous in geographic space and discrete in time. Section 3 applied the proposed approach to annual average temperature forecasting, which is compared with the observation data at 137 international meteorological stations in China from 1993 to 2002. Section 4 provides conclusions and directions for further research.

2. PROCEDURE OF MODELING

The procedure of modelling and forecasting non-stationary space-time series can be categorized into four stages: data preparation, data analysis, training and validating. In data preparation stage, outliers in data should be detected and removed from the data sets. In data analysis stage, exploratory space-time analysis should be made to diagnose whether data satisfy modelling conditions such as correlation and stationarity. The data are examined whether spatial and temporal patterns are existent using time series analysis and exploratory spatial data analysis (ESDA) methods. If nonexistent, it means data represent a space-time stationary process. Otherwise ANN model (see Section 2.1) can be performed on the data to capture the non-linear space-time trends. In training stage, ANN model (see Section 2.1) should be applied to discover non-linear space-time trends, then ANN residuals (observation values subtract ANN values) are examined whether correlation is existent using ESTA. If uncorrelated, it means ANN has modelled all space-time structures represented in the raw data. Otherwise candidate STARMA model (see Section 2.2) must be performed on the residuals to capture the correlations. When both ANN and STARMA have been fitted, space-time autocorrelation function of the residuals will be calculated for diagnostic checking whether the residuals are random. If the residuals still obtain obvious stochastic space-time variation structures, candidate STARMA model will again be adjusted till the residuals are approximately white noise. In validating stage, the trained ANN+STARMA model is used to predict non-stationary space-time processes. The space-time forecasting values are obtained by a sum of ANN and the STARMA estimates (see Equation 1). The performance of the modelling is evaluated by the prediction accuracy.

2.1 Artificial neural network Modelling for Space-Time (or Trend) Patterns

Artificial neural network (ANN) models are known to be universal and flexible function approximators, and they have been used to simulate non-linear systems, and to describe all kinds of data (Hagan et al. 1996; Mitchell 2003; Acharya et al. 2006).

In spatial data analysis, ANN is used to discover spatial patterns (Kanevski et al. 1996; Bollivier et al. 1997; Li and Dunham 2002; Kanevski and Maignan 2004). We think that, depending on its architecture, ANN can also capture space-time patterns on different scales, describing both linear and non-linear effects. In this study, an ANN with a back-propagation training algorithm is applied. The back-propagation algorithm is an iterative gradient and supervised learning algorithm that is designed to minimize the error measure between the actual output of the neural network and the desired output. Here, the large-scale nonlinear space-time pattern term $\mu_i(t)$ (in (1)), which uses the ANN with one hidden layer, is modelled as a function in time and space:

$$\hat{\mu}_i(t) = \sum_{k=1}^{n} \beta_k f_k(i,t) + \beta_0 \tag{2}$$

where i, which represents spatial location (which has x and y as two dimensional spatial coordinates) and current time t, is regarded as the input of the neural network; n is the number of the hidden layer nodes; $\hat{\mu}$ represents the forecasted value at spatial location i at current time t, which is the output of the neural network; function f is the non-linear activation function; β_k is conjunctive weight; β_0 is threshold value. The model has very strong processing ability for non-linear spatial trends (or patterns). However, it is weaker in the time aspect because it can reflect only upward or downward temporal trends. The equation will be applied first for the fitting stage and later for the forecasting stage.

2.2 STARMA Modelling for Stationary Space-Time Process

The remaining space-time correlated error term $e_i(t)$ (in (1)) represents small-scale stochastic variations. The STARMA is used to model the space-time correlated error term. The STARMA model class is a linear combination of past observations at location i and their neighbours influence and its basic principle for a space unit forecasting at time t is shown in Figure 1. In this case, the STARMA model consists of autoregressive term and moving average term and it takes the following form (Martin and Oeppen 1975; Pfeifer and Deutsch 1980):

$$z_i(t) = \sum_{k=1}^{p}\sum_{h=0}^{m_k} \phi_{kh} W^{(h)} z_i(t-k) - \sum_{k=1}^{q}\sum_{h=0}^{n_k} \theta_{kh} W^{(h)} \varepsilon_i(t-k) + \varepsilon_i(t) \tag{3}$$

where

p is the autoregressive order,

q is the moving average order,

λ_k is the spatial order of the kth autoregressive term,

m_k is the spatial order of the kth moving average term,

ϕ_{kl} is the autoregressive parameter at temporal lag k and spatial lag l,

θ_{kl} is the moving average parameter at temporal lag k and spatial lag l,

$W^{(l)}$ is the $N \times N$ matrix of weights for spatial order l ($W^{(0)} = I$),

$\varepsilon_i(t)$ is the random normally distributed error vector at time t at location i with conditions.

(a) $E[\varepsilon_i(t)] = 0,$

(b) $E[\varepsilon_i(t)\varepsilon_j(t+s)'] = \begin{cases} \sigma^2 I, i = j, s = 0, \\ 0, i \neq j, s \neq 0, \end{cases}$

(c) $E[e(t)\varepsilon(t+s)'] = 0,\ for\ (s > 0)$

where condition (a) represents that expectation of $\varepsilon_i(t)$ is zero; condition (b) represents the assumptions commonly made with regard to the STARMA model is that the variance-covariance matrix is equal to $\sigma^2 I$ and s represents nonzero temporal lag for the residuals; condition (c) represents that autocovariances at nonzero lags equal to 0. Various tests are available for testing the three conditions to determine whether the model does adequately represent the data such as exploratory spatial data analysis (ESDA), time series analysis and space-time autocorrelation function (Hamilton 1994; Haining 2003; Pfeifer and Deutsch 1980). In this study, space-time autocorrelation function of the residuals will be calculated for diagnostic checking of the residuals. If the residual term $\varepsilon_i(t)$ is approximately white noise, the mean of space-time autocorrelations of the residuals should be closer to zero and the variance should be closer to $[N(T-s)]^{-1}$. If the residual term $\varepsilon_i(t)$ is not random they may follow a pattern that can't be represented by STARMA model (Pfeifer and Deutsch 1980).

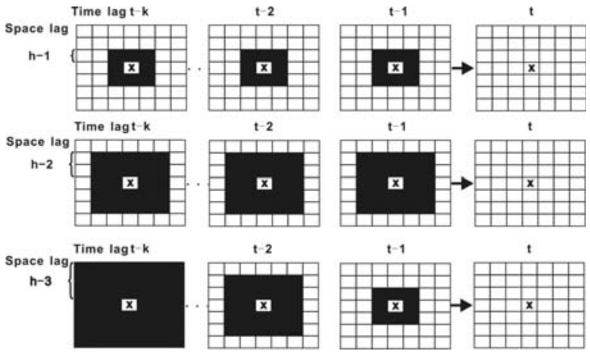

Figure 1. The basic principle of STARMA model for a space unit forecasting at time t.

3. CASE STUDY

3.1 Data Preparation

The proposed framework is tested by forecasting the china annual average temperature (degree/year). The original data are based on annual average temperature at $N = 194$ international meteorological exchanging stations provided by national meteorological centre of P. R. China, which have $T = 52$ year observations from 1951 to 2002. Figure 2 (a) shows a map of International meteorological exchanging stations in China with $N = 194$ monitoring stations under study. Figure 2 (b) shows $T = 52$ sequence plots for stations Beijing, Guangzhou, and Urumchi, whose locations are indicated in Figure 2 (a). In 57 of 194 stations the measurements were of questionable quality after normal distribution checking so the information provided was discarded and the rest 137 stations remained. To train and validate the models, the data sets were be split into two subsets: 80% as sample set to train the model, and 20% as validation set to test and validate the model. Therefore, in our case, the meteorological data between 1951 and 1992 (42 years in total, nearly 80% of 52 years) are chosen as the training dataset for the forecasting between 1993 and 2002 (10 years in total, nearly 20% of 52 years).

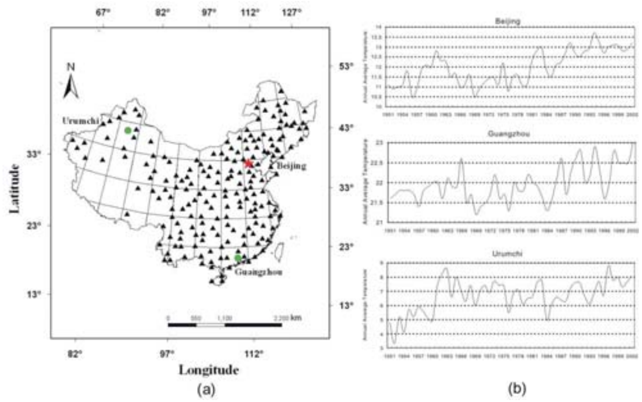

Figure 2. International meteorological exchanging stations in study area: (a) spatial location distribution of the 194 stations; (b) time series of annual average temperature from 1951 to 2002 at the three stations of Beijing, Guangzhou, and Urumchi, which are marked in (a)

3.2 Exploratory space-time analysis

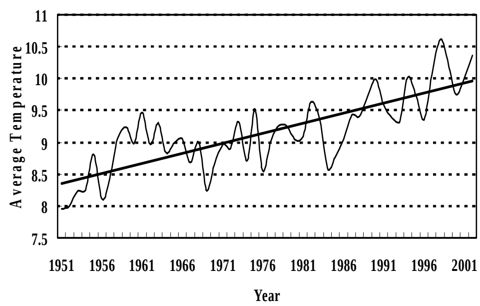

Figure 3. Sequence mean temperature plot for the whole study area

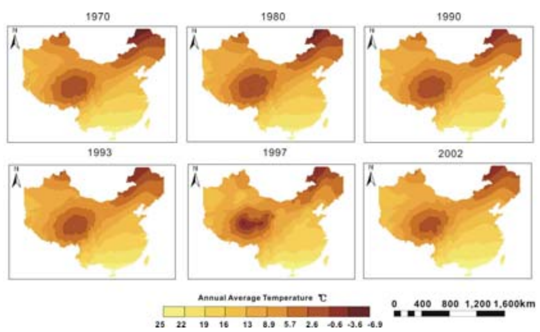

Figure 4. Maps of spatial distribution of annual average temperature for the years 1970, 1980, 1990, 1993, 1997, and 2002

Time series analysis and exploratory spatial data analysis deals with the following steps respectively: statistical analysis, temporal trend analysis, spatial trend analysis. This is an important stage of the study for the ANN and STARMA model.

Means and variances of whole 137 International Meteorological Stations from 1951-2002 are calculated and then the sequence mean temperature plot for the whole study area was drawn (see Figure 3). As is clearly depicted in Figure 3, sample means from 1951-1990 are upward trend and indicate that series are non-stationary. Structure analysis of sample data discovers explicitly spatial trend for use of kriging model (see Figure 4). This conclusion leads to use ANN for space-time trend modeling.

3.3 The ANN model to predict the space-time (or trend) patterns

The ANN model was built to capture non-linear space-time trends (see Section 2.1). The implemented neural network can be seen in Figure 5. The ANN model used had the following parameters: three input neurons with linear activation function of spatial coordinates longitude (x), latitude (y), and time t (year), which were normalized to a specified range [0, 1]; one hidden layer with five processors and a sigmoid activation function; an output neuron with sigmoid activation function, describing annual average temperature at spatial location (x, y) and time t. This choice was based on the analysis of the training and testing errors.

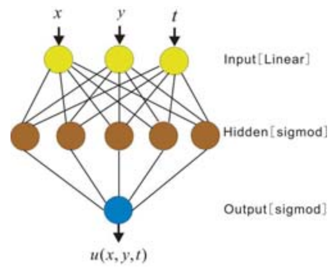

Figure 5. Structure of the implemented ANN model. It should be noted with attention that training data were organized as a sample, the length of which is 137×42, and there are 137 outputs at each year t, which represent annual average temperature forecast at 137 stations

In the ANN model, the training data were organized as a sample in which length is 137×(42) and there are 137 outputs in each year t, which represent fitted annual average temperature at 137 stations. The fitted results in 1970, 1980, and 1990 for large-scale deterministic space-time trends are presented in Figure 6, which shows that the ANN model captured non-linear space-time trends.

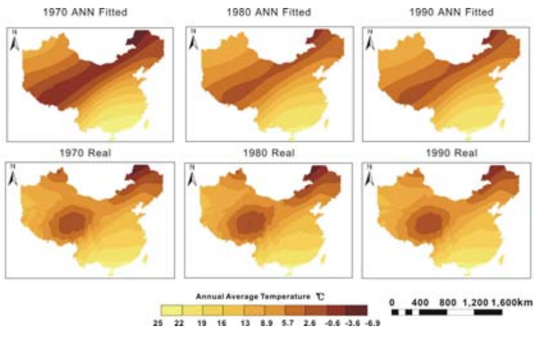

Figure 6. Non-linear space-time trends captured by the ANN model

3.4 The STARMA to model the space-time variances

3.4.1 Define the spatial weight matrix

First, ANN residuals are analyzed. The isotropic semi-variogram model, $\gamma(h)$, with a gaussian function was used to analyze space-time variance structures of ANN residuals. Table 1 shows the parameters of sample ANN residual spatial variance structures at different years.

Year	Range (km)	Partial Sill (C)	Nugget (C_0)	Sill (C_0+C)	C_0/sill (%)
1951	1484.8	9.019	7.826	16.845	46.459
1955	1493.8	6.262	4.090	10.352	39.509
1960	1532.4	5.264	2.540	7.804	32.547
1965	1463.2	6.671	1.312	7.983	36.435
1970	1541.9	16.939	9.833	26.772	36.729
1975	1552.3	4.245	2.510	6.755	37.158
1980	1421.0	4.537	2.567	7.104	36.135
1985	1502.9	4.780	2.841	7.621	37.279
1990	1402.1	3.916	2.501	6.417	38.975

Table 1 Summary of sample ANN residuals isotropic semi-variogram analysis parameters in past several decades

Then, weights were defined according to the Euclidean distance between two points as

$$\begin{cases} w(h) = \left[(C_0 + C_1) - \gamma(h)\right]/(C_0 + C_1) & h \le a \\ w(h) = 0 & h = 0 \text{ or } h > a \end{cases} \quad (4)$$

where

$w(h)$ is a weight function about distance h,
$\gamma(h)$ is the gaussian semi-variogram function value,
a is the spatial correlation distance (or range),
C is partial sill value,
C_0 is nugget value,
$C + C_0$ is the sill or sample variance.

Thus, $w(h)$ tends to decrease as h increases. That is, if values are similar (distance smaller), weight will be close to 1, and if values are dissimilar (distance larger), weight will be close to 0. These weights are expressed as a hierarchical ordering of spatial neighbours. The definition of spatial order represents an ordering in terms of Euclidean distance of all stations surrounding the locations of interest. First order neighbours are those “closest” to the station point of interest. Second order neighbours should be “farther” away than first order neighbours, but “closer” than third order neighbours (Pfeifer and Deutsch 1980). In the study, spatial order is defined as one according to range of spatial autocorrelation.

3.4.2 STARMA Model

To identify spatial lag and temporal lag order of STARMA model, the sample space-time autocorrelation and partial autocorrelation function of ANN residuals is presented in Table 2 and Table 3. The sample residuals’ space-time autocorrelations appear to tail off with both space and time; the sample residuals’ space-time partial autocorrelations seem to cut off at temporal lag second, at spatial lag the zero and the first so that this can be identified as a STARMA (3,0), where STARMA stands for space-time autoregressive moving average process, autoregressive order is 3, and moving average order is 0. The candidate STARMA (3,0) model is defined with the form as follows:

$$e_i(t) = \phi_{10}e_i(t-1) + \phi_{20}e_i(t-2) + \phi_{30}e_i(t-3) + \phi_{11}W^{(1)}e_i(t-1) + \phi_{21}W^{(1)}e_i(t-2) + \phi_{31}W^{(1)}e_i(t-3) + \delta + \varepsilon_i(t) \quad (5)$$

Thus, the ANN residuals taken at a specific point at time t is modelled as a linear combination of the three previous ANN residual values at this point plus a weighted average of the ANN residuals taken from its first order neighbours at time $t-1$, $t-2$, and $t-3$ plus a constant term, and plus a random error term. The least squares estimates of the parameters are performed through a run of Matlab7.0 and the parameter values are depicted in Table 4.

Space lag(l) **Time lag(s)**	0	1
1	0.934	0.059
2	0.890	0.049
3	0.860	0.047
4	0.831	0.049
5	0.799	0.043

Table 2 Sample space-time autocorrelations of the ANN residuals

Space lag(l) **Time lag(s)**	0	1
1	-0.953	-0.713
2	-0.012	-0.248
3	-0.000	-0.027
4	0.005	0.109
5	-0.002	-0.121

Table 3 Sample space-time partial autocorrelations of the ANN residuals

Variable	ϕ_{10}	ϕ_{20}	ϕ_{30}
Coefficient	0.2967	0.4365	0.2632
Std. Error	0.1021	0.0911	0.0086
t-Statistic	2.9043	4.7874	3.0568
Probability	0.0043	0.0000	0.0027
Variable	ϕ_{11}	ϕ_{21}	ϕ_{31}
Coefficient	0.4948	-1.3523	1.0118
Std. Error	0.8855	0.7615	0.4469
t-Statistic	5.5880	-1.7757	2.2643
Probability	0.0504	0.0781	0.0252

Table 4 Parameter estimation for the candidate STARMA model

After the parameters of model (5) were estimated, diagnostic checking of the model of the STARMA residuals was performed through a calculation of the space-time autocorrelations of the STARMA residuals. In the examined STARMA residuals T is equal to 137 and N is equal to 42 so that the standard deviation of the space-time autocorrelations of the STARMA residuals is approximately equal to 0.0132 (see (2)). From table 5, a calculation of mean and variance of the space-time autocorrelations of the STARMA residuals show the results approximately satisfy random normal distribution condition, which mean is close to zero and variance is approximately equal to 0.0132 so that the candidate STARMA model can adequately represent the ANN residual data. That is, the candidate STARMA model captured a majority of small-scale stochastic space-time variances of sample (see Equation 1). The fitted results of the ANN+STARMA model in 1970, 1980, and 1990 for sample data are shown in Figure 7.

Space lag(l) **Time lag(s)**	**0**	**1**
1	0.029	0.021
2	0.014	-0.022
3	-0.011	-0.019
4	0.009	-0.018
5	-0.008	0.017

Table 5 Space-time autocorrelations of the candidate STARMA model residuals

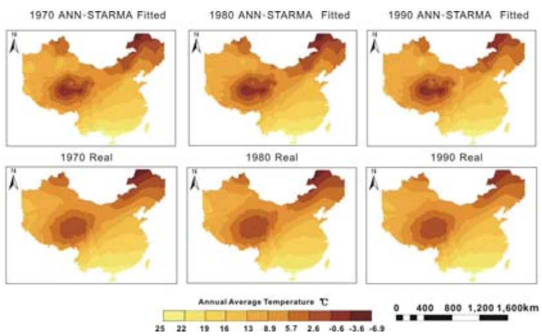

Figure 7. Maps of ANN+STARMA model fitted results for the three years 1970, 1980, and 1990

3.4.3 Validation

The final stage is a validation of trained ANN model and estimated STARMA model. Figure 8 shows a resulting comparison between different models for the forecasted years 1993, 1997, 2002 and performance evaluation is described in Table 6.

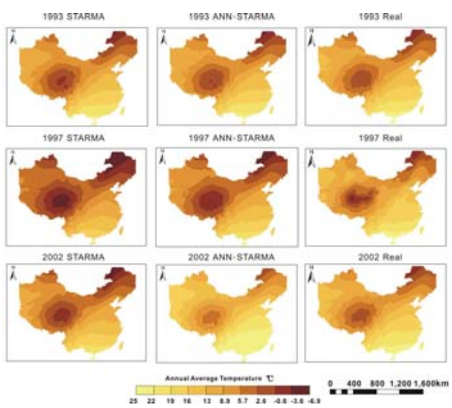

Figure 8. Maps of Pure STARMA and ANN+STARMA model forecast results for the three years 1993, 1997, and 2002

Year	ANN+STARMA		STARMA	
	RMSE	Correlation Coefficient	RMSE	Correlation Coefficient
1993	0.615	0.988	0.649	0.971
1997	3.086	0.982	3.377	0.969
2002	6.703	0.968	7.728	0.957

Table 6 Performance of the different models on test sets

As can be seen in the table 6, the RMSE errors increasingly become bigger and correlation coefficient is smaller than before over time evolution. We find performance of short term space-time forecasting is better than metaphase and long term forecasting for the two models. It also indicates two models can be more suitable for short term space-time forecasting. However, it does show the improvement of the integrated ANN+STARMA model than the pure STARMA model in terms of the RMSE, especially for relatively longer-term prediction.

4. CONCLUSIONS AND DISCUSSION

In the study, we gave a beneficial attempt using Artificial neural network (ANN) to take non-linear space-time trends out from space-time non-stationary process. The ANN+STARMA model is a kind of semi-parametric method, which combines the data-driven ANN and the model-driven STARMA model and it is very useful for data set that is continuous in space and discrete in time. The proposed method has been applied to the forecasting of china annualaverage temperature, which is compared with the observation data at 137 international meteorological stations in China from 1993-2002. The comparison confirms that the forecasting using proposed method can obtain better forecasting accuracy than using the conventional pure STARMA method. It means that proposed model would be able to give useful forecasts for processes with strong non-linear and non-stationary components. In addition, the performance of short term space-time forecasting is better than metaphase and long term forecasting for the two models. It also indicates two models are more suitable for short term space-time forecasting.

Besides, since ANN only is a statistic neural network so in the proposed approach it is not enough to forecast dynamic space-time trend changes, although it can capture current space-time trends. A dynamic recurrent neural network (DRNN) might be a good choice. A dynamic recurrent neural network, which is a neural network with feedback connections, might be more appropriate for this case because in a DRNN the output depends not only on the current input to the network, but also on the previous inputs, outputs, and the state of the network. This feature makes the recurrent neural network particularly suitable for modelling dynamic behaviours, especially, in real time applications that to follow the dynamic changes in space (e.g., forest fires and temperature change). Thus, DRNN should be considered as a possible replacement for ANN in our next work.

REFERENCES

Acharya, C., Mohanty, S., Sukla, L. B., Misra, V. N.**,** 2006. Prediction of sulphur removal with acidithiobacillus using artificial neural networks. *Ecological Modelling*, 190(1), pp. 223-30.

Bollivier, M., Dubois, G., Maignan, M., Kanevsky, M.**,** 1997. Multilayer perceptron with local constraint as an emerging method in spatial data analysis. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389, pp. 226-29.

Cliff, A. D., Ord, J. K., 1975. Space-time modelling with an application to regional forecasting. *Transactions of the Institute of British Geographers*, 64, pp. 119-28.

Crespo, J. L., Zorrilla, M, Bernardos, P., Mora, E., 2007. A new image prediction model based on spatio-temporal techniques. *The Visual Computer,* 23(6), pp. 419-31.

Cressie, N., Majure, J. J., 1997. Spatio-temporal statistical modeling of livestock waste in streams. *Journal of Agricultural, Biological and Environmental Statistics*, 2(1), pp. 24-47.

Deutsch, S. J., Ramos, J. A., 1986. Space-time modelling of vector hydrologic sequences. *Water Resource Bull*, 22, pp. 967-80.

Hagan, M. T., Demuth, H. B., Beale, M. H., 1996. *Neural network design Boston*, PWS Publishing Company.

Haining, R. P., 2003. *Spatial data analysis: theory and practice*. Cambridge, Cambridge University Press.

Hamilton, J. D., 1994. *Time series analysis*. New Jersey, Princeton University Press.

Kamarianakis, Y., Prastacos, P., 2005. Space-time modeling of traffic flow. *Computers & Geosciences*, 31, pp. 119-33.

Kanevski, M., Arutyunyan, R., Bolshov, L., Demianov, V., Maignan, M., 1996. Artificial neural networks and spatial estimation of Chernobyl fallout. *GeoInformatics*, 7(1), pp. 5-11.

Kanevski, M., Maignan, M., 2004. *Analysis and modelling of spatial environmental data Lausanne*, EPFL Press.

Li, Z., Dunham, M. H.**,** 2002. STIFF: A forecasting framework for spatio-temporal data. *International Workshop on Knowledge Discovery in Multimedia and Complex Data (KDMCD 2002),* pp. 183-88.

Martin, R. L., Oeppen, J. E.**,** 1975. The identification of regional forecasting models using space-time correlation functions. *Transactions of the Institute of British Geographers*, 66, pp. 95-118.

Mitchell, T. M., 2003. *Machine learning New York*, McGraw-Hill Press.

Pace, R. K., Barry, R., Clapp, J., Rodriguez, M., 1998. Spatio-temporal autoregressive models of neighbourhood effects. *Journal of Real State Finance and Economics*, 17(1), pp. 15-33.

Pace, R. K., Barry, R., Gilley, O. W., Sirmans, C. F., 2000. A method for spatial-temporal forecasting with an application to real estate prices. *International Journal of Forecasting*,16(2), pp. 229-46.

Pfeifer, P. E., Bodily, S. E., 1990. A test of space-time ARMA modeling and forecasting with an application to real estate prices. *International Journal of Forecasting*, 9(4), pp. 255-72.

Pfeifer, P. E., Deutsch, S. J., 1980. A three-stage iterative procedure for space-time modeling. *Technometrics*, 22(1), pp. 35-47.

ACKNOWLEDGEMENTS

The research is supported by the Major State Basic Research Development Program of China (973 Program, No. 2006CB701305), the Ministry of Education of China (985 Project, No. 105203200400005).