Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 72(3) 2018 239–253

An improved Space-Time Autoregressive Moving Average (STARMA) model for Modelling and Forecasting of Spatio-Temporal time-series data

Santosha Rathod*[1], Bishal Gurung[2], K.N. Singh[2] and Mrinmoy Ray[2]
[1]*ICAR-Indian Institute of Rice Research, Hyderabad*
[2]*ICAR-Indian Agricultural Statistics Research Institute, New Delhi*

Received 31 August 2017; Revised 06 August 2018; Accepted 20 October 2018

SUMMARY

The univariate Box-Jenkins models ended up being extremely helpful in expansive range of time series analysis. Since these models are univariate, they are appropriate just in single series of data and can't manage the factors which are systematically dependent over space. Be that as it may, the greater part of the climatic marvels is subject to dependent of their neighbourhoods. To address these issues, one ought to consider the model which incorporates systematic dependencies in both space and time. On other hand spatio-temporal modelling fuses the spatial correlation between the observations at neighbouring regions over a timeframe. The autoregressive and moving average components of univariate time series slacked in both space and time is alluded as space time autoregressive moving average (STARMA) model. The spatial information on different location is incorporated by considering spatial weight grid. In this article, an attempt has been made to incorporate the second order uniform spatial weight matrix to model and forecast the spatio-temporal time-series data. Efforts also have been made to include the spatial heterogeneity among locations by considering inverse distance weightage derived from Euclidean distance of Riemannian great circle distance using longitude and latitude of the respective locations. The proposed methodology has been implemented in simulated data. As a contextual investigation monthly maximum temperature (°C) of nine districts of North Karnataka has been considered to illustrate STARMA model. As average temperature of North Karnataka is above 30°C and the same is responsible for growing most of the horticultural crops, with these consideration maximum temperature data is considered for this study. The outcomes uncovered that the proposed method of STARMA model outperformed the univariate ARIMA and first spatial order STARMA model for modelling and forecasting, for both simulated as well as in actual data.

Keywords: Spatio-temporal time series, Maximum temperature, ARMA, Spatial weight matrix, STACF, STPACF, STARMA.

1. INTRODUCTION

Spatio-temporal time series are the observations which are recorded over both space and time by considering systematic dependencies across space and time. Spatio-temporal modelling manages the single variables recorded over a timeframe at various locations. The case of spatio-temporal data incorporates; daily or hourly carbon emission data recorded from observatory at many location, daily river flow data recorded from many river basins, hourly daily or weekly record of many weather parameters over different locations, and traffic flow measurements taken from a set of loop detectors in an exceptionally visit premise are cases of spatial time series data. Spatio-temporal modelling is commonly used in numerous areas *viz.,* Geo-statistics, sociology, economics, environmental, ecological and agricultural science Many literatures recommend that incorporation of both spatial and temporal information will enhance the demonstrating effectiveness of phenomenon under consideration (Neuman *et al*. (2010), Schlenker and Roberts (2009) and White *et al.* (2006)). In several cases only temporal approach of modelling has been used (Abdel-Aal and Elhaddy (1994). Baillie and Chung (2002), Kaushik and Singh (2008) and Chattopadhay *et al.* (201) and in some cases only spatial modelling has been used, but in reality the underlying phenomenon depends on both spatial and temporal effects, the univariate temporal or univariate

Corresponding author: Santosha Rathod
E-mail address: santosha.rathod@icar.gov.in

* Paper presented in Dr. G.R. Seth Memorial Young Scientist Award Session on November 26, 2017 during the 71st Annual Conference of the Indian Society of Agricultural Statistics held at ICAR-DRMR, Bharatpur, Rajasthan during November 25-27, 2017.

spatial modelling is incapable of providing information on both spatial and temporal effects (Smith, 2000). In this way, it is sensible to model time and space scales at the same time to catch inherent vulnerability over a time frame over various locations.

Because of computational difficulties and inaccessibility of simultaneous spatial and temporal information, no significant progress is accomplished in spatio-temporal time-series modelling as contrast with univariate time series modelling. Spatio-temporal models are the models which considers concurrent information on both space and time of variables under consideration. In univariate time series we observe autocorrelation between the successive observations over a time-frame. to model these data series, the Box-Jenkins autoregressive moving average (Box and Jenkins (1970) modelis most usually utilized model because of its prominent modelling building process. On other hand, the auto-correlated spatio-temporal time-series phenomenon can be modeled using the space time autoregressive moving average (STARMA) model. The autoregressive and moving average components of univariate time series lagged in both space and time is alluded (rewrite) as space- time autoregressive moving average (STARMA) model.

A classical multivariate time- series model i.e. Vector Auto-Regressive Moving Average (VARMA) model can be used to model the spatio-temporal data, but the number of parameters becomes more (Brynjarsdottir and Berliner (2014) and Gupta and Robinson (2015)). Cliff and Ord (1975) and Martin and Oeppen (1975) were the first to model the relationship between two variables in space and time. Cliff and Ord introduced space time model STARMA model that have less number of parameters compares to VARMA models. The STARMA model was first time introduced in early eighties, from that point forward several techniques have been developed corresponding to different inferential needs and data types. The space-time autoregressive integrated moving average (STARIMA) methodology was initially delineated in a series of papers by Pfeifer and Deutsch (1980, 1980a, 1980b, 1980c, 1981, 1981a, 1981b). Deutsch and Pfeifer (1981) addressed the contemporary correlated errors in STARMA modelling. Tunay (2010) described the process of estimation of STARMA models. Kamarianakis (2003) reviewed the methodologies of space- time auto regressive moving average modelling. Kyriakidis and Journel (1999) proposed on geo-statistical space-time models. Oud *et al.* (2012) addressed spatial dependence in continuous time modelling.

From that point onwards, the STARMA model has been connected to wide assortment of spatial time-series data for example; river flow (Pfeifer and Deutsch (1981a)), spread of disease (Pfeifer and Deutsch (1980a)), real estate prices (Pfeifer and Bodily, (1990)) spatial econometrics (Elhorst (2000), Climatic data (Subba and Antunes (2003)), Disease modelling (write modelling in the entire text) (Lee (2005)),Traffic flow data (Kamarianakis and Prastacos, (2005), Lin *et al.* (2009) and Ding *et al.* (2011)), Timber prices (Zhou and Buongiorno (2006)), Rural watersheds (Dalezios (1995)), Regional employment (Giacinto (2006), Solar radiation (Glasbey and Allcroft (2008)), Damage detection (Hu *et al.* (2011), GDP data (Nurhayati (2012)) and Regional bank deposits (Kurt and Tunay (2015)) *etc.*

In spatio-temporal time series modelling defining of spatial weight matrix is key to model building which thusly decides the model accuracy. In STARMA modelling spatial weight matrix is characterized in light of scope and limitations of the study in different ways. Uniform weight matrix or homogenous spatial weight matrix is most ordinarily utilized as spatial weight matrix in STARMA modelling. The spatial weight matrix again relies on upon number of spatial lag, as number of spatial lag increases, it prompts modelling and computational challenges (Subba and Antunes 2003). However, to improve the performance, second order spatial weight matrix has been considered to build STARMA model in this study. One of the major drawbacks of uniform spatial weight matrix is they do not consider the spatial dynamics and heterogeneity and intern fails to capture the spatial heterogeneity (Cheng *et al.* 2014). Therefore, to overcome this difficulty, the spatial weight matrix has been built using Euclidean distance of Riemannian great circle using longitude and latitude of the locations. The stationary homogenous spatial weight matrix has been constructed by assigning uniform weightage to all the neighbours. The proposed methodology has been implemented in both simulated as well as in actual data. As a contextual investigation, monthly maximum temperature of nine districts of North Karnataka is considered to build STARMA model.

1.1 ARIMA Model

ARIMA is standout amongst the most conventional and broadly utilized model for time- series modelling due to by Box and Jenkins (1970). In contrast to the regression models, the ARIMA model allows to explain by its past, or lagged values and stochastic error terms. These models are often referred to as "mixed models" since they use a combination of autoregressive (AR), integration (I) referring to the reverse process of differencing to produce a stationary series and moving average (MA) operations. An ARIMA model is usually stated as ARIMA (*p d q*).

An autoregressive integrated moving average model is expressed in the following expression

$$\emptyset(B)(1-B)^d Y_t = \theta(B)\varepsilon_t \quad (1)$$

$$Y_t = \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (2)$$

Where, Y_t *is* the $\emptyset_i$ ε_t time series, and θ_j are model parameters, is random error, *p* is number of autoregressive terms, *q* is number of lagged forecast errors and *B* is the backshift operator such that, $BY_t = Y_{t-1}$ (Box and Jenkins (1970)). The ARIMA model building consists of three stages, *viz*. Identification, estimation and diagnostic checking.

1.2 STARMA Model

The space-time models explain the systematic dependencies over both space and time is modelled through the class of STARMA models (Pfeifer and Deutsch,1980b). The autoregressive and moving average form of space time model represented by STARMA model are characterized by single variable $Z_i(t)$, observed at *N* fixed spatial locations ($i = 1, 2, \ldots, N$) on *T* time periods ($t = 1, 2, \ldots, T$). The *N* spatial locations can be a geographical locations, viz. Ccountry, Sstate, *etc*. The spatial dependencies between N times series is incorporated through *N*N* spatial weight matrices. Analogous to univariate time series, ***Z(t)*** is expressed as a linear combination of past observations and errors. The STARMA model (Pfeifer and Deutsch, 1980a), denoted by $STARMA(p_{\lambda_1,\lambda_2,\ldots,\lambda p}, q_{m_1,m_2,\ldots,mq})$ can be represented in the matrix equation as follows;

$$\boldsymbol{Z(t)} = \sum_{K=1}^{p}\sum_{l=0}^{\lambda_k}\phi_{kl}W^l Z(t-k) - \sum_{K=1}^{q}\sum_{l=0}^{m_k}\theta_{kl}W^l \varepsilon(t-k) + \varepsilon(t) \quad (3)$$

Where, $\boldsymbol{z(t)} = [\boldsymbol{z_1}(t), \ldots\ldots, \boldsymbol{z_N}(t)]'$ is a $N \times 1$ vector of observations at time $t = 1,\ldots, T$, *p* is the autoregressive order (AR) with respect to time, *q* is the moving average order (MA) with respect to time, λ_k is the spatial order of the k^{th} AR term, m_k is the spatial order of the k^{th} MA term, ϕ_{kl} is the AR parameter at temporal lag k and spatial lag *l* (scalar), θ_{kl} is the MA parameter at temporal lag k and spatial lag *l* (scalar) and W^l is the *N*N* spatial weight matrix with spatial order *l* with diagonal elements zero and non-diagonal elements is the relation between sites. The spatial weight matrix = I_N *i.e.* Identity matrixand each row of W^l must add up to one. The random error vector is normally distributed at time *t* with $W^{(0)}$

$$\varepsilon(t) = [\varepsilon_1(t), \varepsilon_2(t), \ldots, \varepsilon_N(t)]'$$

$$E[\varepsilon(t)] = 0,$$

$$E[\varepsilon(t)\varepsilon'(t+s)] = \begin{cases} G = \sigma^2 I_N \; is \; s = 0 \\ 0, \; otherwise \end{cases} \text{ and}$$

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

There are two subclasses of the STARMA model, in equation (3) when *q*=0, only autoregressive terms remain and consequently the model progresses toward becoming space-time autoregressive or STAR model which is represented as follows:

$$\boldsymbol{Z(t)} = \sum_{K=1}^{p}\sum_{l=0}^{\lambda_k}\phi_{kl}W^l Z(t-k) + \varepsilon(t) \quad (4)$$

When *p* becomes 0, only moving average terms remains and hence the model becomes spacetime moving average or STMA model which is represented as follows;

$$\boldsymbol{Z(t)} = \varepsilon(t) - \sum_{K=1}^{q}\sum_{l=0}^{m_k}\theta_{kl}W^l \varepsilon(t-k) \quad (5)$$

Similar to Box-Jenkins univariate ARIMA methodology the STARMA model is also build by three stage procedures of model building viz., identification, estimation and diagnostic checking, proposed by (Pfeifer and Deutsch, 1980b). The STARMA model is said to stationary if covariance structure of ***Z(t)*** does not change with time and every ***Z(t)*** lie inside the unit root circle i.e. the STAR model are invertible and STAMA models are stationary. The space time autocorrelation function (STACF) and space time partial autocorrelation function (STPACF) are used to identify the STAR and STMA order. Like univariate ARIMA model, the STAR and STMA model orders are identified in view of significant STAR and STMA spikes. The space time autocorrelation function (STACF) between l[th] and k^{th} order neighbour's time lag apart (s=1,…,k and h=0,1,…, λ) is given underneath;

$$\rho_{lk}(s) = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T-S} W^{(l)} z_i(t) W^{(k)} z_i(t+s)}{\left[\sum_{i=1}^{N}\sum_{t=1}^{T-S}(W^{(l)} z_i(t))^2 \cdot (W^{(k)} z_i(t+s))^2\right]^{1/2}} \quad (6)$$

The space time partial autocorrelation function (STPACF) is expressed in following equation;

$$\rho_{h0}(s) = \sum_{j=1}^{k}\sum_{l=0}^{\lambda} \phi_{jl}\,\rho_{hl}(s-j) \quad (7)$$

After the selection of model orders, the next step is to estimate the parameters of the model. If the residuals are found to be homogenous then conditional least square techniques is use to estimate the model parameters otherwise conditional maximum likelihood estimation method can be used. In light of STACF and STPACF of residuals, independency of residuals can be diagnosed. In the event that residuals are approximately white noise, the sample space-time autocorrelation functions ought to all be viably zero.

1.3 Spatial Weight Matrix

Building of spatial weight matrix plays a key role in STARMA modelling, the hierarchical ordering of neighbours of each site and the selection of an appropriate sequence of weighting matrices is a matter left to the model builder since more complex the weight matrix, more troublesome is to estimate the parameters of STARMA model. In majority of cases, the space pattern is assumed to be equal and regularly spaced to ease the model building. In majority applications, the uniform spatial weight matrix is only a simplifying assumption since typically the sites are irregularly spaced. A weight can be picked in different ways, the least difficult of which is the binary scheme, if two areas shared a common border then we relegate a weight as 1otherwise 0(Griffith (1996) and (2009)). Be that as it may, in spatial weight matrix, row normalization is a common practice i.e. making all rows sum to one is common practice. These weights, in any case, must reflect a hierarchical ordering of spatial neighbours. First order neighbours are those which are closest to the chosen site. Second order neighbours are farther away than first order neighbour's, yet closer than third order neighbour's. The spatial weight matrix again depends on number of spatial lag, as number of spatial lag increases, it prompts modelling and computational challenges (Subba and Antunes, 2003). However, to improve the performance, second order spatial weight matrix has been considered to build STARMA model in this study. The second order spatial weight matrix STARMA model in this paper is denoted as STARMA-II.

1.4 Proposed method to construct the spatial weight matrix based on Riemannian great circle formula

For each location, the Euclidean distance (in kilometres) between the sites is determined using the following expression:

$$d_{ij} = r * acos\{sin\left(\frac{lat_i}{57.2958}\right)sin\left(\frac{lat_j}{57.2958}\right) + cos\left(\frac{lat_i}{57.2958}\right)cos\left(\frac{lat_j}{57.2958}\right)cos\left(\frac{lon_j}{57.2958} - \frac{lon_i}{57.2958}\right)\} \quad (8)$$

Where *r,* is radius of earth is assumed to be 6378.8 kilometres and *i, j*=1,…, 9, lat_i and lon_i are the latitude and longitude of site *i,* respectively. Inverse Euclidean distance is considered to build the first spatial order weight matrix. In this paper this method has been denoted as STARMA-III.

1.5 Derivation of Formulae for Out of Sample Forecasts of STARMA model

In this study, we confine our attention only to deriving out-of-sample of one-step and two-step ahead forecast formulae in respect of the STARMA (*p, q*) models. However, formulae for more than two-step ahead forecasts, though quite complicated, can be derived along similar lines. The optimal predictor, which minimizes the mean one-step ahead squared prediction error is the conditional expectation, obtained from (3).

$$\hat{Z}_{t+i|t+i-1} = \hat{Z}_{t+i|1,2,\ldots,t+i-1}$$
$$= E\left\{\hat{z}_{t+i}\,\middle|\,\chi_{t+i-1,}\Phi\right\} \quad (9)$$

Where, $x_{t-1,}$ is the information contain in $X_{0,}X_{1,\ldots,}X_{t-1}$ and Φ is the parameter space.

Using recursive conditional approach, we get

$$\hat{z}_{t+i|t+i-1} = -\hat{\phi}_{10} z_{(t+i)-1} - \hat{\phi}_{11} w^{(l)} z_{(t+i)-1} + \hat{\theta}_{10}\varepsilon_{(t+i)-1} - \hat{\theta}_{11} w^{(l)}\varepsilon_{(t+i)-1} \quad (10)$$

Formulae for two step ahead out of sample forecasts are derived analytically by recursive use of conditional expectation

$$\hat{z}_{t+i+1|t+i-1} = \hat{z}_{t+i+1|1,2,\ldots,t+i-1}$$

$$= E\left[\left\{z_{t+i+1}\middle|x_{t+i-1,}\Phi\right\}\right]$$

$$= E\left[E[z_{t+i+1}|x_{t+i,}\theta]\middle|x_{t+i-1,}\Phi\right]$$

$$\hat{z}_{t+i+1|t+i-1} = -\hat{\phi}_{10}\hat{z}_{t+i} - \hat{\phi}_{11}w^{(l)}\hat{z}_{t+i} + \hat{\theta}_{10}\hat{\varepsilon}_{t+i} - \hat{\theta}_{11}w^{(l)}\hat{\varepsilon}_{t+i} \quad (11)$$

2. SIMULATION STUDY

The data matrix for the STARMA model has been simulated in STARMA framework. All the simulations reported were performed using normal random numbers. The simulations are designed to show the STARMA model building and forecasting. The data was simulated from a system of nine locations distributed spatially on a regular grid with weighting matrix based on the parameters given in table (Table No.). The data matrix has been generated with the following model specifications;

$$Z_t = 0.81Z_{t-1} + 0.11W^1Z_{t-1} + \varepsilon_t - (-0.11)\varepsilon_{t-1} - (-0.17)\,W^1\varepsilon_{t-2} \quad (11)$$

200 observations were simulated over 9 locations. Among 200 observations 190 were used as training set and 10 were used as testing data. For simulation study the weight matrices has been generated randomly (Table 2 and 3). The Inverse distance spatial weight matrix has been constructed using longitude and latitude of different locations (Table 5). The univariate Box Jenkins ARIMA modelling for simulated data has

Table 1. Spatial weight matrix of order zero.

Location	L1	L2	L3	L4	L5	L6	L7	L8	L9
L1	1	0	0	0	0	0	0	0	0
L2	0	1	0	0	0	0	0	0	0
L3	0	0	1	0	0	0	0	0	0
L4	0	0	0	1	0	0	0	0	0
L5	0	0	0	0	1	0	0	0	0
L6	0	0	0	0	0	1	0	0	0
L7	0	0	0	0	0	0	1	0	0
L8	0	0	0	0	0	0	0	1	0
L9	0	0	0	0	0	0	0	0	1

Table 2. First order spatial weight matrix for simulated data

Location	L1	L2	L3	L4	L5	L6	L7	L8	L9
L1	0	0.50	0	0.50	0	0	0	0	0
L2	0.33	0	0.33	0	0.33	0	0	0	0
L3	0	0.50	0	0	0	0.50	0	0	0
L4	0.33	0	0	0	0.33	0	0.33	0	0
L5	0	0.25	0	0.25	0	0.25	0	0.25	0
L6	0	0	0.33	0	0.33	0	0	0	0.33
L7	0	0	0	0.50	0	0	0	0.50	0
L8	0	0	0	0	0.33	0	0.33	0	0.33
L9	0	0	0	0	0	0.50	0	0.50	0

Table 3. Second order spatial weight matrix for simulated data

Location	L1	L2	L3	L4	L5	L6	L7	L8	L9
L1	0	0	0.33	0	0.33	0	0.33	0	0
L2	0	0	0	0.33	0	0.33	0	0.33	0
L3	0.33	0	0	0	0.33	0	0	0	0.33
L4	0	0.33	0	0	0	0.33	0	0.33	0
L5	0.25	0	0.25	0	0	0	0.25	0	0.25
L6	0	0.33	0	0.33	0	0	0	0.33	0
L7	0.33	0	0	0	0.33	0	0	0	0.33
L8	0	0.33	0	0.33	0	0.33	0	0	0
L9	0	0	0.33	0	0.33	0	0.33	0	0

Table 4. Row normalized Inverse distance spatial weight matrix

Location	L1	L2	L3	L4	L5	L6	L7	L8	L9
L1	0	0.1529	0.1128	0.1093	0.1022	0.1002	0.1071	0.1928	0.1227
L2	0.1172	0	0.2116	0.1169	0.1219	0.1159	0.1146	0.1073	0.0947
L3	0.0773	0.1893	0	0.1176	0.1513	0.1570	0.1312	0.0874	0.0888
L4	0.0831	0.1161	0.1306	0	0.2527	0.1347	0.1063	0.0844	0.0921
L5	0.0642	0.0999	0.1386	0.2086	0	0.2032	0.1250	0.0726	0.0878
L6	0.0599	0.0905	0.1371	0.1059	0.1936	0	0.2315	0.0767	0.1047
L7	0.0683	0.0954	0.1221	0.0892	0.1270	0.2467	0	0.0972	0.1542
L8	0.1628	0.1183	0.1077	0.0936	0.0977	0.1082	0.1287	0	0.1830
L9	0.0966	0.0973	0.1020	0.0953	0.1101	0.1378	0.1903	0.1706	0

Table 5. Longitude and Latitude of locations under consideration

Sl.No.	Location	Latitude	Longitude
1	Gulbarga	17.31320N	76.87497E
2	Bijapur	16.82754N	75.72532E
3	Raichur	16.20082N	77.36228E
4	Bagalkot	16.18170N	75.69580E
5	Belgaum	15.85037N	74.50465E
6	Dharwad	15.46025N	75.01028E
7	Gadag	15.423302N	75.603708E
8	Koppal	15.35070N	76.155434E
9	Bellary	15.13940N	76.92140E

Table 6. Univariate modelling of simulated data

Location	Model	AR Parameters		MA Parameters	
		AR1	AR2	MA1	MA2
L1	ARIMA (101)	0.25 (0.011)	-	0.35 (0.012)	-
L2	ARIMA (100)	0.53 (0.06)	-	-	-
L3	ARIMA (101)	0.47 (0.010)	-	0.29 (0.014)	-
L4	ARIMA (101)	0.15 (0.07)	-	-0.23 (0.02)	-
L5	ARIMA (100)	0.51 (0.06)	-		-
L6	ARIMA (200)	0.66 (0.07)	-0.04 (0.06)		-
L7	ARIMA (102	0.55 (0.06)	-0.07 (0.03)	-0.816 (0.040)	-
L8	ARIMA (100)	0.42 (0.011)	-	-0.65 (0.012)	-0.30 (0.024)
L9	ARIMA (101)	0.59 (0.07)	-	0.31 (0.09)	-

been depicted in Table 6. Finally, the modelling and forecasting performance is depicted in Table 7 to 10. The results clearly indicated that the STARMA model outperformed the ARIMA model in all 9 locations. Further the proposed STARMA models STARMA-II and STARMA-III outperformed over the ARIMA model and STARMA-I (First order uniform spatial weight matrix STARMA model).

Table 7. Model Performance (MAPE) under training data set

Sl. No	Location	ARIMA	STARMA-I	STARMA-II	STARMA-III
1	L1	2.04	1.57	1.54	1.42
2	L2	2.14	1.67	1.63	1.55
3	L3	1.84	1.57	1.49	1.38
4	L4	1.96	1.54	1.43	1.40
5	L5	2.07	1.62	1.58	1.45
6	L6	1.92	1.48	1.45	1.41
7	L7	1.99	1.75	1.72	1.68
8	L8	2.02	1.54	1.50	1.44
9	L9	1.93	1.51	1.48	1.39

3. APPLICATION: A CASE STUDY OF MAXIMUM TEMPERATURE OF NORTH KARNATAKA

In this study monthly mean maximum temperature of nine districts viz., Gulbarga, Bijapur,

Raichur, Bagalkot, Belgaum, Dharwad, Gadag, Koppal and Bellary of North Karnataka state of India are considered to model and forecast using STARMA model. Data on monthly maximum temperature (°C) of north Karnataka districts from January, 2000 to August, 2016 were collected from

Table 8. Model Performance (MAPE) under testing data set (2-steps- ahead forecast)

Obs.	Location 1					Location 2					Location 3				
	Actual	Forecast				Actual	Forecast				Actual	Forecast			
		ARMA	STARMA-I	STARMA-II	STARMA-III		ARMA	STARMA-I	STARMA-II	STARMA-III		ARMA	STARMA-I	STARMA-II	STARMA-III
1	39.86	40.14	40.02	39.93	39.78	42.16	40.54	41.02	41.74	42.74	39.97	40.83	40.72	40.63	40.42
2	41.06	39.93	39.97	40.12	40.29	40.04	40.31	40.78	40.23	40.12	40.00	38.88	39.1	39.24	39.57
3	38.88	39.87	38.81	38.14	38.65	38.47	40.18	40.02	39.92	39.62	40.17	39.00	39.5	39.63	39.87
4	36.92	39.86	37.12	36.18	36.35	39.96	40.11	39.78	40.98	40.17	38.44	37.11	37.64	37.69	37.92
5	37.69	39.08	38.98	38.08	38.03	39.81	40.07	39.89	40.1	39.7	39.96	40.90	40.83	40.78	40.67
6	37.67	39.76	39.23	38.34	38.13	39.45	40.05	40.25	40.56	40.78	39.96	38.92	40.67	40.62	40.47
7	38.58	39.96	39.87	39.84	39.91	39.30	40.04	39.9	38.94	38.94	40.36	39.19	39.23	39.67	39.89
8	38.39	39.86	39.86	39.6	39.13	40.79	40.04	40.47	40.87	40.83	39.12	39.92	40.03	39.82	39.74
9	39.90	39.84	40.4	40.12	40.11	40.58	40.03	40.63	40.59	40.51	39.97	38.76	38.81	39.5	39.78
10	40.09	41.06	42.27	42.23	42.11	41.47	40.03	40.16	40.32	40.5	40.00	42.62	42.54	41.56	41.05
MAPE		3.32	2.52	2.15	1.73	MAPE	2.00	1.68	1.52	1.22	MAPE	3.08	2.6	1.91	1.32

Table 9. Model Performance (MAPE) under testing data set (2-steps- ahead forecast)

Obs.	Location 4					Location 5					Location 6				
	Actual	Forecast				Actual	Forecast				Actual	Forecast			
		ARMA	STARMA-I	STARMA-II	STARMA-III		ARMA	STARMA-I	STARMA-II	STARMA-III		ARMA	STARMA-I	STARMA-II	STARMA-III
1	40.35	39.93	40.03	40.17	40.23	42.31	41.11	41.32	41.25	41.7	41.08	39.21	39.3	40.25	40.48
2	39.56	40.25	40.17	40.13	39.86	41.02	40.53	40.63	40.711	40.92	41.14	39.47	40.2	40.34	40.55
3	38.77	40.40	39.93	39.54	39.05	40.61	40.23	40.72	40.56	40.65	41.37	39.63	39.97	40.12	40.75
4	38.56	40.42	40.11	39.85	39.16	41.45	40.08	41.12	41.2	41.35	41.50	39.72	40.25	40.45	40.75
5	39.17	40.41	40.24	40.13	39.6	41.81	40.00	41.15	41.45	41.7	41.19	39.78	40.03	40.44	40.72
6	38.80	41.08	39.72	39.17	39.01	40.46	41.97	41.5	41.25	41.05	40.46	39.81	39.98	40.117	40.5
7	40.13	41.17	40.98	40.64	40.53	40.61	39.95	40.16	40.55	40.57	41.33	41.83	41.72	41.61	41.52
8	40.60	40.34	40.43	40.32	40.25	40.90	39.94	39.97	40.16	39.97	41.42	39.84	40.25	40.33	40.75
9	40.35	41.04	40.89	40.76	40.5	42.31	38.91	39.18	40.28	40.5	41.72	39.81	40.2	40.34	40.32
10	39.56	42.10	41.86	41.24	41.05	41.02	40.93	40.98	41.08	41.05	40.64	39.85	40.1	40.15	40.3
MAPE		3.22	2.41	1.78	1.09	MAPE	2.82	1.93	1.36	1.04	MAPE	3.70	2.57	2	1.37

Table 10. Model Performance (MAPE) under testing data set (2-steps- ahead forecast)

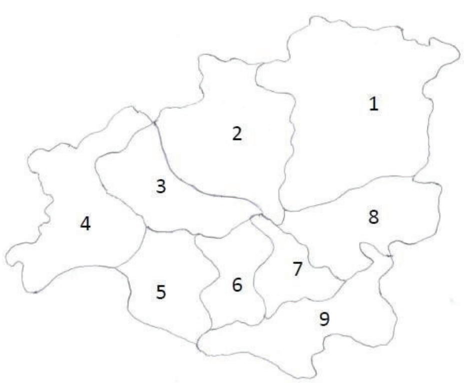
Obs.	Location 7					Location 8					Location 9				
	Actual	Forecast				Actual	Forecast				Actual	Forecast			
		ARMA	STARMA-I	STARMA-II	STARMA-III		ARMA	STARMA-I	STARMA-II	STARMA-III		ARMA	STARMA-I	STARMA-II	STARMA-III
1	40.17	41.55	41.2	41.12	40.7	41.74	40.20	40.32	40.85	41.2	41.28	40.34	39.85	40.2	40.85
2	40.21	42.25	41.5	41.45	41.02	39.66	40.76	40.15	40.06	39.98	40.71	39.80	40.04	40.15	40.3
3	38.77	41.8	40.05	39.95	39.55	39.64	40.99	40.25	40.05	39.9	41.23	40.48	40.89	40.92	41.1115
4	41.35	40.01	40.45	40.7	40.85	41.09	42.64	42.24	41.75	39.97	40.54	40.29	40.7	40.46	40.6
5	39.13	41.25	40.75	40.25	40.68	39.43	40.92	40.56	40.32	40	41.60	40.17	40.48	40.55	40.62
6	37.60	39.24	39.02	38.92	38.2	37.80	41.61	39.5	39.2	39.4	41.32	42.11	41.86	41.45	41.4
7	37.51	39.1	38.115	38.15	37.98	37.99	40.12	39.65	39.06	38.7	41.09	41.97	41.7	41.65	41.5
8	38.89	38.12	38.6	38.55	38.62	38.98	40.98	40.25	39.78	39.93	42.01	40.05	41.1	41.25	41.45
9	38.78	37.89	37.15	37.33	37.5	42.22	39.91	40.75	40.97	40.88	39.31	38.03	38.25	38.3	38.9
10	40.95	38.11	39.25	39.55	39.56	40.68	43.16	42.16	41.8	40.7	39.52	41.02	40.95	40.85	40.78
MAPE		4.47	2.98	2.61	2.07	MAPE	4.98	3.1	2.23	1.86	MAPE	2.62	2.02	1.7	1.16

Table 11. Descriptive Statistics of maximum temperature of North Karnataka

Locations	Mean	S. D	Minimum	Maximum	Skewness	Kurtosis	C.V (%)
Gulbarga	34.06	4.01	28.7	42.78	0.68	-0.83	11.79
Bijapur	32.95	3.79	26.73	41.69	0.67	-0.68	11.49
Raichur	34.47	3.54	29.71	43.71	0.74	-0.66	10.26
Bagalkot	32.6	4.03	26.04	42.5	0.63	-0.65	12.35
Belgaum	31.41	4.53	22.82	42.1	0.08	-0.82	14.42
Dharwad	31.94	4.1	24.97	41.48	0.63	-0.58	12.84
Gadag	31.98	4.13	24.96	41.73	0.67	-0.53	12.92
Koppal	32.87	3.91	27.46	42.31	0.71	-0.63	11.90
Bellary	33.82	3.55	28.61	43.48	0.68	-0.75	10.51

global weather data portal developed by The National Centre for Environmental Prediction (NCEP), Texas A&M University (http://globalweather.tamu.edu). The data from January, 2000 to August, 2015 is used for model building and data from September, 2015 to August, 2016 is used for model validation (Forecasting performance). Regardless of the study, descriptive statistics of all nine districts (Locations) are ascertained to comprehend the nature of data under consideration (Table 11). By observing values

Fig. 1. Geographical map of Karnataka



1. Gulbarga
2. Bijapur
3. Bagalkot
4. Belgaum
5. Dharwad
6. Gadag
7. Koppal
8. Raichur
9. Bellary

Fig. 2. Map of districts/locations considered

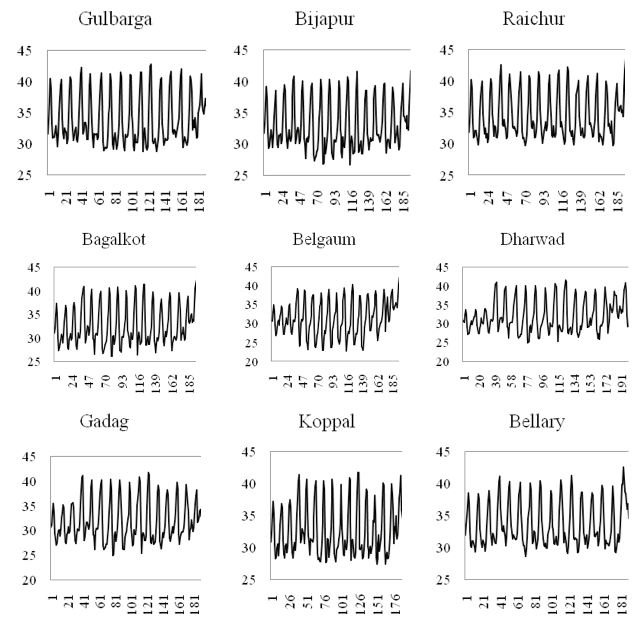


Fig. 3. Original time series of monthly mean maximum temperature of North Karnataka

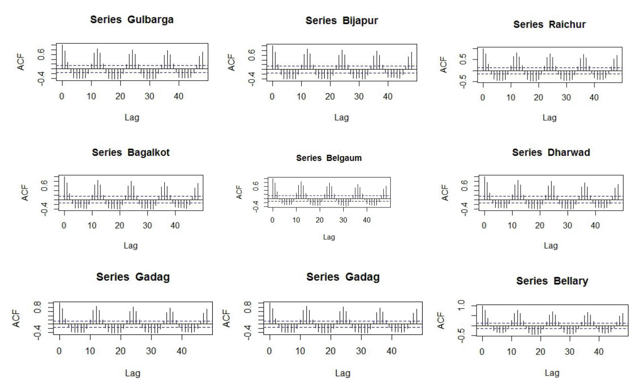


Fig. 4. ACF plots of original monthly mean maximum temperature time series of North Karnataka

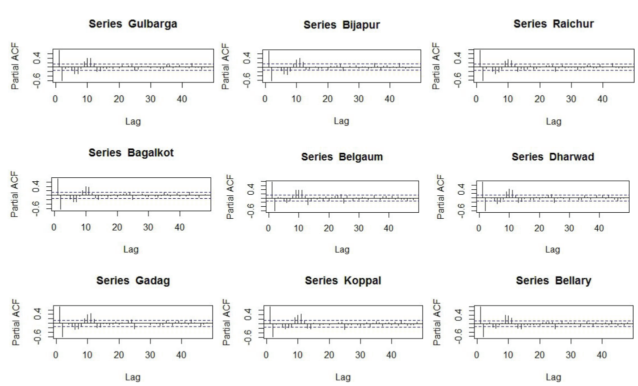


Fig. 5. PACF plots of original monthly mean maximum temperature time series of North Karnataka

of skewness and kurtosis, one can infer that, the data under consideration follows Gaussian distribution. The time series plots of each series plotted in figure 3 indicates presence of seasonality, subsequently the original series are seasonally adjusted.

3.1 Fitting of univariate ARIMA models to maximum temperature series

The univariate ARIMA model has been fitted for all districts separately. The ACF and PACF plots of original series (Figure 4 & 5) indicates the series is having seasonal effect. In this way, for effective modelling and forecasting the time series, data are seasonally adjusted by removing the seasonal effect. The ACF and PACF plots of seasonally adjusted series (Figure 6 & 7) unmistakably indicates, series are free of seasonality and are found to be stationary. Therefore, the series were seasonally adjusted.

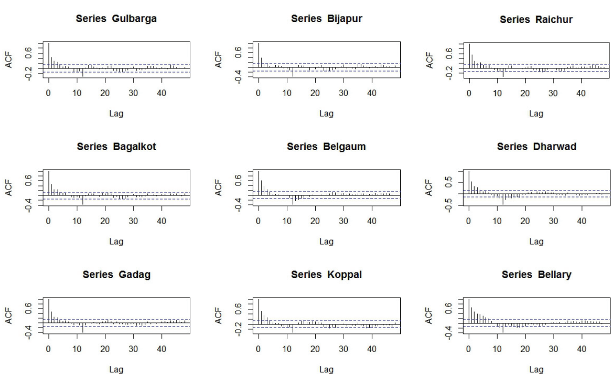


Fig. 6. ACF plots of seasonally adjusted monthly mean maximum temperature time series of North Karnataka

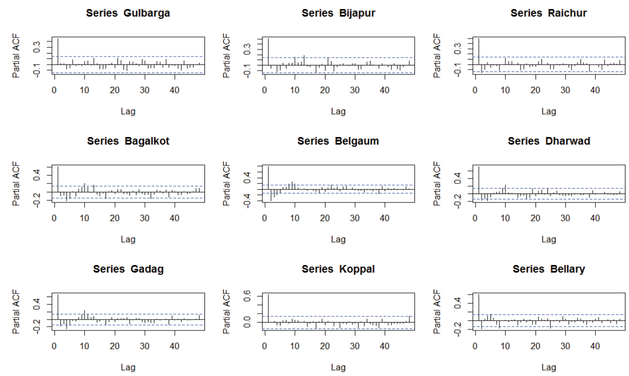


Fig. 7. PACF plots of seasonally adjusted monthly mean maximum temperature time series of North Karnataka (Gulbarga data like that may be mentioned everywhere)

The Box-Jenkins univariate ARIMA model has been fitted for all 9 locations separately. The best candidate ARIMA model has been selected automatically by *auto.arima* function in R package 'forecast'. The autoregressive and moving average parameters of respective models of 9 locations alongside standard errors are mentioned in Table 12. After the candidate model selection and parameters estimation by maximum likelihood estimation method, diagnostic checking of residuals is important. Ljung-Box test statistic has been employed to check the efficacy of the selected model. Chi-square value of Ljung-Box test statistic (Table 13) unmistakably uncovers that residuals of each models are non-noteworthy. Further, performance of selected models under training data set (model building) and testing data set (model validation) using mean absolute percentage error (MAPE) has been reported in table 20 to 23.

Table 12. Univariate ARIMA model fitting for seasonally adjusted monthly maximum temperature of North Karnataka

Location	Model	Intercept	AR Parameters			MA Parameters			
			AR1	AR2	AR3	MA1	MA2	MA3	MA4
Gulbarga	ARIMA (100)	33.80 (0.15)	0.486 (0.066)	-	-	-	-	-	-
Bijapur	ARIMA (100)	32.82 (0.178)	0.542 (0.064)	-	-	-	-	-	-
Raichur	ARIMA (002)	34.52 (0.143)	-	-	-	0.618 (0.072)	0.201 (0.072)	-	-
Bagalkot	ARIMA (203)	32.24 (0.142)	0.28 (0.033)	-0.24 (0.028)	-	-0.134 (0.076)	0.299 (0.102)	0.293 (0.068)	-
Belgaum	ARIMA (302)	31.08 (0.455)	0.579 (0.051)	-0.85 (0.087)	0.6538 (0.049)	-0.726 (0.045)	0.782 (0.036)	-	-
Dharwad	ARIMA (204)	31.67 (0.203)	0.692 (0.030)	-0.64 (0.027)	-	-0.376 (0.080)	0.176 (0.107)	0.236 (0.103)	0.136 (0.070)
Gadag	ARIMA (303)	31.59 (0.159)	0.625 (0.040)	-0.891 (0.0669)	0.239 (0.0374)	-0.316 (0.0407)	0.487 (0.0473)	0.193 (0.0250)	-
Koppal	ARIMA (100)	32.699 (0.273)	0.672 (0.056)	-	-	-	-	-	-
Bellary	ARIMA (200)	33.782 (0.209)	0.786 (0.072)	-0.161 (0.080)	-	-	-	-	-

Table 14. Neighbours of each site for each spatial order

Location	Order	
	1	2
1	2,8	3,6,7
2	1,3	4,8,7
3	2,4,5,6	7,8
4	3,5	2
5	3,4,6	9
6	3,5,7,9	2,8
7	6,8,9	1,2,3
8	1,7,9	2,3,6
9	6,7,8	5

Table 13. Diagnostic checking of ARIMA model of maximum temperature time series of Karnataka

Time-series	Box-Ljung statistic	Probability
Gulbarga	6.23	0.858
Bijapur	12.50	0.328
Raichur	9.10	0.613
Bagalkot	11.45	0.407
Belgaum	8.47	0.671
Dharwad	12.90	0.300
Gadag	9.20	0.608
Koppal	7.09	0.792
Bellary	18.508	0.070

3.2 Construction of spatial weight matrix for maximum temperature series

As explained in methodology section, the spatial weight matrix has been constructed by assigning equal weightage to each neighbours. The map of nine locations under consideration is delineated in Figure 2 and each location are represented by numbers from one to nine. In light of the neighbouring locations, connectivity spatial weight matrices have been considered. For instance, for location 1, location 2 and location 8 are first order neighbours. Again, 3, 6 and 7 are second order neighbours to location one. In a similar manner, first and second order neighbours for all nine locations are reported in table 5. In light of the numbers of neighbours, the spatial weights have been doled out to each location. In uniform spatial weight matrix equal weights are relegate to each neighbours. To make row normalization i.e. making all rows sum to one we divide, one by number of neighbours (correct neighbours in the entire document) *i.e.* $\frac{1}{n}$, here n is number of neighbours. For example, for first location (Gulbarga) there are two first order neighbours, then we divide one by two and assign 0.5 as weight to each locations. As we calculated weight for first location, one can proceed in same manner to calculate weights for all nine locations. In light of this procedure first order spatial weight matrix has been calculated in table 14. As explained in methodology section, in this article attempt has been made to incorporate second order spatial weight matrix in STARMA model. For first location 3, 6 and 7 are second order neighbours, then we divide one by three and assign 0.33 as weight to each locations; in the same manner one can proceed further to calculate weights to all nine locations for second order neighbours. The second order spatial weight matrix for all nine locations are depicted in table 16. To compute STACF and STPACF first order spatial weight matrix (Table 15) is need to be incorporate in the model. In first order spatial weight matrix, since we do not assign weights to any neighbours, diagonal elements end up noticeably equal to one. The proposed inverse distance matrix (Table 4) has been constructed using Euclidean distance of Riemannian great circle using longitude and latitude of the locations under consideration (Table 5).

Table 15. First order spatial weight matrix for Maximum temperature data

Location	Gulbarga	Bijapur	Raichur	Bagalkot	Belgaum	Dharwad	Gadag	Koppal	Bellary
Gulbarga	0	0.5	0	0	0	0	0	0.5	0
Bijapur	0.5	0	0.5	0	0	0	0	0	0
Raichur	0	0.25	0	0.25	0.25	0.25	0	0	0
Bagalkot	0	0	0.5	0	0.5	0	0	0	0
Belgaum	0	0	0.33	0.33	0	0.33	0	0	0
Dharwad	0	0	0.25	0	0.25	0	0.25	0	0.25
Gadag	0	0	0	0	0	0.33	0	0.33	0.33
Koppal	0.33	0	0	0	0	0	0.33	0	0.33
Bellary	0	0	0	0	0	0.33	0.33	0.33	0

3.3 STARMA model fitting

In this article, STARMA model was estimated using the three stage procedure due to Pfeiffer and Deutsch (Pfeiffer and Deutsch, 1980a). As discussed in methodology section, STARMA estimation procedure is extension of Box-Jenkins ARIMA methodology in spatio-temporal set up. As in ARMA, it has three stages of model building *viz.,* model identification, estimation and diagnostic checking. Model identification is most vital stride to utilize the forms of spatio-temporal models viz., STAR, STAMA and STARMA. TACF and PACF plots of original series are depicted in fig. 8 and 9 and which indicates the presence of seasonality, therefore, the data has been adjusted seasonally and STACF and STPACF plot plotted of the seasonally adjusted series are plotted in fig. 10 and 11. Based on the significant spikes in STACF and STPACF plots, the STARMA (1 0 1) model has been selected.

Parameters of the identified models are estimated using maximum likelihood method and are given in Table 17 (First order STARMA), Table 18 (Second order STARMA) and Table 19 (Inverse distance STARMA) alongside their standard errors and probability values. The estimated parameters are then consolidated in the model and predicted values were obtained. For diagnostic checking, multivariate Box-Pierce Non Correlation test is applied and the residuals are observed to be non-correlated. Further, performance of models under consideration for training data set (model building) and testing data set (model validation) utilizing mean absolute percentage error (MAPE) obtained are presented in Table 20 to 23.

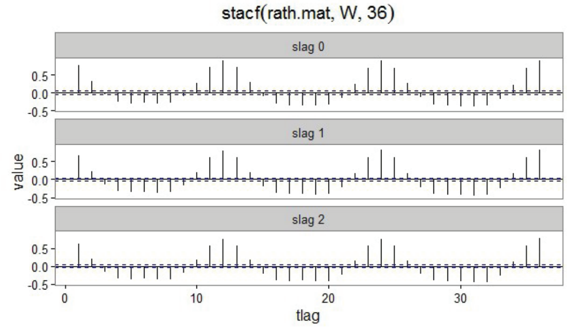


Fig. 8. STACF plots of original monthly mean maximum temperature time series of North Karnataka

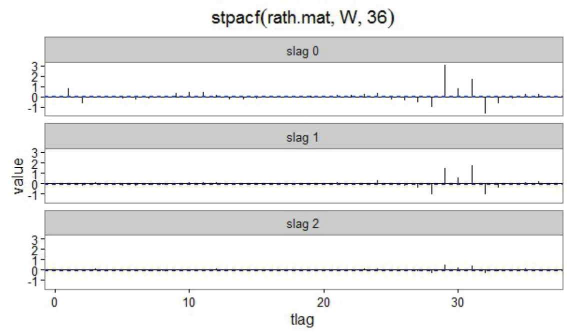


Fig. 9. STPACF plots of original monthly mean maximum temperature time series of North Karnataka

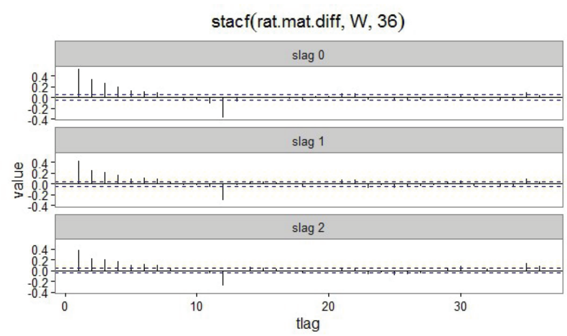


Fig. 10. STPACF plots of seasonally adjusted monthly mean maximum temperature time series of North Karnataka

Table 16. Second order spatial weight matrix

Location	Gulbarga	Bijapur	Raichur	Bagalkot	Belgaum	Dharwad	Gadag	Koppal	Bellary
Gulbarga	0	0	0.33	0	0	0.33	0.33	0	0
Bijapur	0	0	0	0.33	0	0	0.33	0.33	0
Raichur	0	0	0	0	0	0	0.5	0.5	0
Bagalkot	0	1	0	0	0	0	0	0	0
Belgaum	0	0	0	0	0	0	0	0	1
Dharwad	0	0.5	0	0	0	0	0	0.5	0
Gadag	0.33	0.33	0.33	0	0	0	0	0	0
Koppal	0	0.33	0.33	0	0	0.33	0	0	0
Bellary	0	0	0	0	1	0	0	0	0

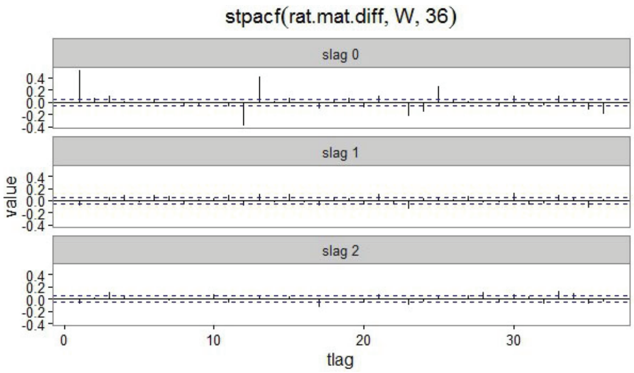


Fig. 11. STPACF plots of seasonally adjusted monthly mean maximum temperature time series of North Karnataka

Table 17. STARMA-I Model parameters

Spatial lag	Slag 0		Slag 1	
	AR	MA	AR	MA
Parameters	0.79	0.23	-0.14	-0.15
(0.019)	(0.04)	(0.02)	(0.04)	
Probability	<0.0011	<0.0011	<0.0011	<0.0011

Multivariate Box-Pierce Non Correlation Test of residuals: Chi-square=62.13 (p=0.21) Values in the parenthesis indicates the standard error

Table 18. STARMA-II Model parameters

Spatial lag	Slag 0		Slag 1		Slag 2	
	AR	MA	AR	MA	AR	MA
Parameters	0.66 (0.023)	0.119 (0.010)	0.171 (0.052)	0.213 (0.0157)	0.79 (0.089)	0.28 (0.116)
Probability	<0.001	<0.001	0.013	0.004	<0.001	0.010

Multivariate Box-Pierce Non Correlation Test of residuals: Chi-square=69.86 (p=0.31) Values in the parenthesis indicates the standard error

Table 19. STARMA-III Model parameters

Spatial lag	Slag 0		Slag 1	
	AR	MA	AR	MA
Parameters	0.81	0.25	-0.18	-0.17
(0.02)	(0.04)	(0.03)	(0.05)	
Probability	<0.001	<0.001	<0.001	<0.001

Multivariate Box-Pierce Non Correlation Test of residuals: Chi-square=65.19 (p=0.25) Values in the parenthesis indicates the standard error

Table 20. Model Performance (MAPE) under training data set

Sl. No	Location	ARIMA	STARMA-I	STARMA-II	STARMA-III
1	Gulbarga	2.54	1.33	1.30	1.25
2	Bijapur	2.73	1.31	1.29	1.19
3	Raichur	2.36	1.25	1.24	1.18
4	Bagalkot	2.80	1.53	1.49	1.42
5	Belgaum	3.42	2.13	2.07	1.89
6	Dharwad	3.31	1.71	1.69	1.58
7	Gadag	2.97	1.61	1.56	1.51
8	Koppal	2.89	1.49	1.41	1.36
9	Bellary	2.45	1.36	1.24	1.18

4. COMPARISON OF FORECASTING PERFORMANCE

The Mean Absolute Percentage Error (MAPE) has been computed to compare the forecasting performances of all the models under considerations for both training and validation data set for all the locations separately. The MAPE values under training and testing data set for simulated data is given in Table 7 to 10 and for maximum temperature series of North Karnataka, MAPE values for training and testing data set is given in Table 20 to Table 23. In view of the lowest MAPE values of the proposed STARMA model for all the locations, it is confirmed that STARMA model performed better than the Box-Jenkins ARIMA model in all the locations for training data set. Further, two- steps- ahead forecasts were calculated to carry out the model validation. In view of results obtained, it is unmistakably indicating that STARMA model likewise performed better than the Box-Jenkins ARMA models for all nine locations for test data set. The second order spatial lag STARMA model performed better than ARIMA and first order STARMA model, it clearly reveals that second order STARMA model tends towards optimality and admissibility over first order STARMA model. Finally, the inverse distance STARMA model (STARMA-III) performed best over remaining models in both simulated as well as actual data set.

5. CONCLUSION

Box-Jenkins univariate ARIMA models are most popularly used in univariate time series cases, whereas their applications are limited when it comes to multivariate spatio-temporal time series analysis.

Table 21. Model Performance (MAPE) under testing data set (2-steps- ahead forecast)

Series	Gulbarga					Bijapur					Raichur				
	Actual	Forecast				Actual	Forecast				Actual	Forecast			
		ARMA	STARMA-I	STARMA-II	STARMA-III		ARMA	STARMA-I	STARMA-II	STARMA-III		ARMA	STARMA-I	STARMA-II	STARMA-III
Sep-15	**36.1**	36.99	36.85	36.78	36.2	33.5	35.16	32.55	32.6	32.9	35.8	33.4	36.9	36.6	35.56
Oct-15	37.2	34.87	37.89	38.65	37.5	34.4	34.09	33.7	33.9	34	36.1	32.7	37.28	37.1	37.02
Nov-15	36.2	34.32	38.28	37.98	37.1	32.5	33.51	33.89	33.7	33.02	34.9	31.4	33.1	35.5	33.98
Dec-15	35.8	34.06	37.6	37.25	37.1	32.8	33.19	33.92	34	32.95	35.4	30.8	34.18	36.2	34.8
Jan-16	34.5	33.93	37.11	36.92	36.7	32.2	33.02	33.5	33.4	32.9	34.3	31.5	35.89	35.6	35.22
Feb-16	38.8	33.87	40.12	39.55	39.3	36.0	32.93	36.95	36.9	36.15	38.9	34	39.2	40.7	38.78
Mar-16	41.4	33.84	43.98	43.56	43.1	38.7	32.88	40.03	40.1	39.25	41.7	37.5	43.27	43	42.3
Apr-16	42.1	33.82	43.75	43.15	43	41.7	32.85	42.98	42.5	41.05	43.7	40.1	44.55	44.4	44.03
May-16	42.8	33.81	43.9	43.2	41.7	40.6	32.84	41.75	41.3	41.25	42.1	40.5	44.75	44.9	43.2
Jun-16	36.4	33.75	37.25	35.9	35.3	33.5	32.83	34.86	34.2	34.18	36.0	35.8	38.25	37.6	36.29
Jul-16	33.4	33.82	31.1	31.25	31.8	31.0	32.83	28.2	29.3	28.9	34.1	32.6	36.75	36.8	35.9
Aug-16	34.1	33.85	32.12	31.87	32.2	31.7	32.82	28.8	29.7	29.1	35.2	32.3	37.85	37.3	37.85
MAPE		**8.45**	**4.45**	**3.89**	**3.08**	**MAPE**	**7.33**	**4.23**	**3.28**	**2.42**	**MAPE**	**7.94**	**4.5**	**3.93**	**2.41**

Table 22. Model Performance (MAPE) under testing data set (2-steps- ahead forecast)

Series	Bagalkot					Belgaum					Dharwad				
	Actual	Forecast				Actual	Forecast				Actual	Forecast			
		ARMA	STARMA-I	STARMA-II	STARMA-III		ARMA	STARMA-I	STARMA-II	STARMA-III		ARMA	STARMA-I	STARMA-II	STARMA-III
Sep-15	34.7	32.9	35.4	35.25	35.02	34.7	36.75	35.75	35.6	35.12	33.3	35.60	34.2	33.83	33.5
Oct-15	35.1	32.7	36.83	36.5	36.03	35.5	35.38	36.35	36.2	35.8	34.6	34.15	36.5	36.2	35.85
Nov-15	33.3	31.3	35.5	35.2	34.89	35.2	33.81	37.16	37	37.12	32.8	34.23	34	33.5	33.95
Dec-15	33.9	30.9	35.6	34.89	34.25	35.1	32.32	35.75	35.9	35.65	33.8	33.63	34.75	34.3	34.62
Jan-16	33.6	31.8	33.9	33.95	33.4	34.2	31.19	36.74	36.4	36.75	33.1	32.54	35.1	34.9	34.12
Feb-16	37.3	34.6	38.8	38.56	37.85	37.4	30.61	38.98	38.7	38	36.5	31.28	38.1	37.9	37.95
Mar-16	40.6	37.7	42.3	41.46	40.95	40.3	30.65	42.3	41.6	41.11	38.9	30.19	41.2	40.9	39.45
Apr-16	42.5	40.1	43.75	42.9	42.95	42.1	31.20	43.87	43.6	43.65	40.8	29.56	41.75	41.6	41.25
May-16	40.9	38.9	43.8	42.19	42.25	40.2	32.06	42.56	42	42.25	38.6	29.54	39.8	37.5	37.5
Jun-16	33.9	32.8	36.8	34.28	34.35	34.2	32.95	35.5	34.1	34.15	32.1	30.08	34.98	34.7	33.3
Jul-16	31.4	29	32.9	32.75	32.25	30.6	33.60	28.75	29.8	29.75	29.3	31.01	27.15	27.8	28.11
Aug-16	32.2	29	31	31.2	31.2	31.7	33.82	30.94	31.4	31.15	29.8	32.07	30.75	30.2	30.25
MAPE		**6.49**	**4.56**	**2.77**	**1.98**	**MAPE**	**11.27**	**4.31**	**3.08**	**2.68**	**MAPE**	**10.18**	**4.66**	**3.63**	**2.65**

Table 23. Model Performance (MAPE) under testing data set (2-steps- ahead forecast)

Series	Gadag					Koppal					Bellary				
	Actual	Forecast				Actual	Forecast				Actual	Forecast			
		ARMA	STARMA-I	STARMA-II	STARMA-III		ARMA	STARMA-I	STARMA-II	STARMA-III		ARMA	STARMA-I	STARMA-II	STARMA-III
Sep-15	34.0	32.41	34.85	34.7	34.11	34.4	34.2	36.21	35.89	35	35.75	38.14	36.30	36.12	35.78
Oct-15	34.3	33.74	35.4	34.8	34.74	35.3	34.0	35.06	36.22	36	35.89	36.06	37.10	36.50	36.12
Nov-15	32.8	33.31	34	33.25	33.02	32.0	32.6	34.29	33.55	32.9	32.33	34.87	36.70	35.20	34.56
Dec-15	33.4	32.64	35.1	34.9	33.99	34.4	32.5	33.77	36.89	36	35.11	34.27	36.40	36.12	36.04
Jan-16	32.3	31.75	33.3	33.5	33.6	33.5	33.5	33.42	35.55	35.6	35	33.99	35.90	35.75	35.12
Feb-16	36.2	30.86	37.9	37.05	37.12	37.2	36.2	33.18	38.91	38.7	38.23	33.87	39.50	39.25	39.22
Mar-16	38.9	30.19	40.4	39.12	39	39.6	39.4	33.02	41.55	41.2	41.25	33.82	40.60	40.80	40.95
Apr-16	41.7	29.92	43.4	42.77	42.33	42.3	42.4	32.92	43.88	43	43.33	33.80	44.20	43.65	43.50
May-16	39.2	30.11	40.8	40.11	40	40.2	41.6	32.85	39.44	39.6	39.15	33.79	40.90	40.65	40.05
Jun-16	32.6	30.69	35.7	34.25	34.1	34.1	35.4	32.80	32.75	32.8	33.25	33.78	35.80	34.45	34.65
Jul-16	31.2	31.49	33.8	33.22	33.06	32.1	32.3	32.77	33.5	33.2	32.5	33.60	33.00	33.20	32.60
Aug-16	33.0	32.30	33.9	33.88	33.75	33.5	32.7	32.74	32.15	32.7	32.96	33.64	34.90	35.10	34.11
MAPE		**9.11**	**4.56**	**2.92**	**2.27**	**MAPE**	**7.6**	**4.38**	**3.19**	**2.54**	**MAPE**	**7.70**	**4.26**	**3.10**	**2.05**

As contrast to univariate ARIMA model, STARMA models have less number of parameters. For the illustrated data set, the STRAMA model has only four parameters for all the nine locations, whereas ARIMA has ARIMA has numerous parameters, under such cases over parameterization may prompt to lower sum of squares of residuals. Based on the results obtained one can infer that STARMA model outperformed the univariate ARIMA models. Again the proposed STARMA model performed better as compared to the univariate ARIMA and classical STARMA models in both training and testing data set respectively. The proposed models are validated by obtaining out of sample forecast for last twelve months' data set. The out performance of STARMA model over univariate ARMA model could be because of the inclusion of spatial information i.e. neighbouring effect in the form of spatial weight matrix (Information matrix). Further, the proposed STARMA model which addresses the optimality, admissibility and spatial heterogeneity outperformed over traditional models

REFERENCES

Abdel-Aal, R.E and Elhaddy, M.A. (1994). Modelling and forecasting of daily maximum temperature using abductive machine learning. Weather and Forecasting, **10**, 310-325.

Baillie, R.T. and Chung, S.K. (2002). Modelling and forecasting from trend-stationary long memory models with applications to climatology. Int. J. Forecasting, **18**, 215–226.

Box, G.E.P. and Jenkins, G. (1970). Time series analysis, Forecasting and control, Holden-Day, San Francisco, CA.

Brynjarsdottir, J. and Berliner, L.M. (2014). Dimension-Reduced Modelling of Spatio-Temporal Processes, J. Amer. Statist. Assoc., 109:508, 1647-1659, DOI: 10.1080/01621459.2014.904232.

Chattopadhyay, S., Jhajharia, D. and Chattopadhyay, G. (2011). Univariate modelling of monthly maximum temperature time series over northeast India: neural network versus Yule–Walker equation based approach. Meteorological applications, **18**, 70-82.

Cheng, T., Wang, J., Haworth, J., Heydecker, B. and Chow, A. (2014). A Dynamic Spatial Weight Matrix and Localized Space–Time Autoregressive Integrated Moving Average for Network Modelling. Geographical Analysis, **46**, 75-97.

Cliff, A.D. and Ord, John K. (1975). Space-time modelling with an application to regional forecasting, Transactions of Institute of British Geographers, **64**, 119-128.

Dalezios, N.R. (1995). Spatio-temporal precipitation modelling in rural watersheds. Hydrological Sciences Bulletin, **40(5)**, 553-568.

Deutsch S.J. and Pfeifer P.E. (1981). Space-time ARIMA modelling with contemporaneously correlated innovations, Technometrics, **23**, 401-409.

Ding, Q., X. Wang, X. Zhang, and Z. Sun. (2011). Forecasting Traffic Volume with Space–Time ARIMA Model. Advanced Materials Research, 156-57, 979-83.

Giacinto, V.D. (2006). A generalized space–time ARMA model with an application to regional unemployment analysis in Italy, International Regional Science Review, **29**, 159-198.

Glasbey, C. and Allcroft, D. (2008). A spatiotemporal auto-regressive moving average model for solar radiation. J. Roy. Statist. Soc. Series C - Applied Statistics, **57**, 343-355.

Griffith, D.A. (1996). Some Guidelines for Specifying the Geographic Weights Matrix Contained in Spatial Statistical Models. Practical Handbook of Spatial Statistics, 82–148, edited by S. L. Arlinghaus, D. A. Griffith, W. D. Drake and J. D. Nystuen. Boca Raton, FL: CRC Press.

Griffith, D.A., and Heuvelink. G.B.M. (2009). Deriving Space–time Variograms from Space– time Autoregressive (STAR) Model Specifications. Paper presented at the StatGIS 2009 conference, Milos, Greece, June 17.

Gupta, A. and Robinson, P.M. (2015). Inference on higher-order spatial autoregressive models with increasingly many parameters, J. Economet., **186**, 19-31.

Hu, X.T., Qin, Z.Y. and Chu, F.L. (2011). Damage detection in plate structures based on spacetime autoregressive moving average processes. In 9th International conference on damage assessment of structures (DAMAS2011). IOP Publishing Journal of Physics: Conference series, 305, 1.

Kamarianakis, Y. and Prastacos, P. Space–time modelling of traffic flow. Computers & Geosciences, 31 (2005), 119-133.

Kamarianakis, Y. (2003). Spatial Time Series Modelling: A Review of the Proposed Methodologies. Working paper, Regional Economics Applications Laboratory, Urbana, IL.

Kaushik, I. and Singh, S.M. (2008). Seasonal ARIMA model for forecasting of monthly rainfall and temperature, Journal of Environmental Research and Development, **3(2)**, 506-514.

Kyriakidis, P.C. and Journel, A.G. (1999). Geostatistical space-time models. Mathematical Geology. **31(6)**, 651-684.

Lee, C.Y. (2005). Space-Time Modelling and Application to Emerging Infectious Diseases, Doctorate Dissertation, Supervisor: B.K. Epperson, Michigan State University, Department of Forestry.

Lin, S.L., Huang, H.Q., Zhu, D.Q. and Wang, T.Z. (2009). The application of space-time

ARIMA model on traffic flow forecasting. IEEE Proceedings of the Eighth International

Conference on Machine Learning and Cybernetics, Baoding, 12-15 July 2009. 34083412.

Martin, R.L. and Oeppen, J.E. (1975). The identification of regional forecasting models using space-time correlation functions. Trans. Inst. British Geographers, **66**, 95-118.

Neumann, K., Verburg, P.H., Stehfest, E. and Muller. C. (2010). The yield gap of global grain production: A spatial analysis. Agricultural Systems, **103(5)**: 316-326.

Nunung, N, Udjianna S.P. and Oki, N (2012). Application of Generalized Space-Time Autoregressive Model on GDP Data in West European Countries, *J. Probability Statist.*, 2012: 1-16

Oud, J.H.L., H. Folmer, R. Patuelli, and P. Nijkamp. (2012). Continuous-Time Modelling with Spatial Dependence. Geographical Analysis, **44**, 29-46.

Pfeifer, P.E., and Bodily, S.E. (1990). A test of space-time ARMA modelling and forecasting with an application to real estate prices, *Int. J. Forecasting*, **16**, 255-272.

Pfeifer, P.E., and Deutsch, S.J. (1980). A Comparison of Estimation Procedures for the Parameters of the STAR Model. Communication in Statistics, simulation and Comput., **B9(3)**, 255-270.

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

Pfeifer, P.E., and Deutsch, S.J. (1980b). Identification and Interpretation of First-Order SpaceTime ARMA Models. Technometrics, **22(3)**, 397-403.

Pfeifer, P.E., and Deutsch, S.J. (1980c). Independence and Sphericity Tests for the residuals of Space Time ARIMA Models, Communication in Statistics, simulation and Comput., **B9(5)**, 533-549.

Pfeifer, P.E., and Deutsch, S.J. (1981). Variance of the Sample-Time Autocorrelation Function of Contemporaneously Correlated Variables. SIAM Journal of Applied Mathematics, Series A, **40(1)**, 133-136.

Pfeifer, P.E., and Deutsch, S.J. (1981a). Seasonal Space-Time ARIMA modelling. Geographical Analysis, **13(2)**, 117-133.

Pfeifer, P.E., and Deutsch, S.J. (1981b). Space-Time ARMA Modelling with contemporaneously correlated innovations. Technometrics, **23(4)**, 410-409.

Schlenker, W. and Roberts, M.J. (2009). Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. Proceedings of the National Academy of Sciences, **106(37)**, 15594-15598. doi:10.1073/ pnas.0906865106.

Smith, R.L., (2000). Spatial Statistics in Environmental Science. In Nonlinear and Nonstationary Signal Processing, W.J. Fitzgerald, RL. Smith, A. Warden, and P.C. Young, eds.

Subba, R.T. and Antunes, A. (2003) Spatio-temporal modelling of temperature time series: a comparative study. In: Schonberg, F., Brillinger, D.R., Robinson, E. (Eds.), Time Series Analysis and Applications to Geophysical Systems. IMA Publications/ Springer-Verlag, New York, pp. 123-150.

Tunay, K.B. (2010). Space-Time Autoregressive Moving Average (STARMA) Models and Estimation Process. Journal of Financial Researches and Studies, **1(2)**, 47-66 (in Turkish).

White, M.A., Diffenbaugh, N.S., Jones, G.V., Pal, J.S. and Giorgi, F. (2006). Extreme heat reduces and shifts United States premium wine production in the 21st century. Proceedings of the National Academy of Sciences **103(3)**, 11217-11222. doi:10.1073/ pnas.0603230103.

Zhou, M., Buongiorno, J. (2006). Space-time modelling of timber prices. Journal of Agricultural and Resource Economics **31**, 40-56.