ISSN 2229-5518

Model Based Clustering In Time Series With Bivariate AR(P) Model

Vedavathi. K, Srinivasa Rao. K and Nirupama Devi. K

Abstract — Time series data is ubiquitous and is available in many fields including medicine, metrological studies, finance, oil and gas industries and other business domains. Mining of time series data has been the subject of research for decades. The clustering and classification together is needed for efficient and effective mining applications. In clustering, the first step is to obtain the number of classes and then cluster the whole data into those classes. The clustering algorithm is developed by utilizing model building techniques for proper identification of the data set to its associated cluster. The approaches to clustering time series can be classified as model based or model free. The model based approaches assume some form of underline generating process, estimating the model parameters and then perform clustering based on the sample information. In this paper, we develop an algorithm for clustering bivariate time series using autoregessive model of order (p). The initial clusters are found using K-means algorithm *and* the model parameters are estimated using the EM algorithm. The clustering algorithm is developed by utilizing component maximum likelihood. The performance of the developed algorithm is evaluated using real time data collected from a materialogical station. A comparative study of the proposed algorithm is made with the existing data mining algorithm that uses univariate autoregressive process.

Key words — Clustering, Time Series, EM Algorithm, AR Process.

1. INTRODUCTION

Most clustering algorithms adopt the assumption that the data available for grouping is static in nature. The evolutionary nature of time series is very important property in clustering because time series observations are not static in nature. The clustering algorithms usually consider the time series as the points in a n-dimensional space but the property of dynamic behavior of time series over time is neglected. Min Ji et.al., (2013) proposed a dynamic fuzzy clustering algorithm which works by determining those time series whose class labels are vague and further partition them into different clusters over time. Xu-Hong Lin et al., (2008) proposed a dynamic fuzzy clustering algorithm which combines auto regression model and conventional fuzzy clustering algorithm.

Saeed Aghabozorgi et.al., (2015) exposed four main components of time series clustering and aimed to represent an updated investigation on the trend of improvements in efficiency, quality and complexity of clustering time series approaches during last decade. Luis E. Nieto-Barajas and Alberto Contreras-Cristan (2014) proposed a model-based clustering method for time series. The model uses Bayesian semi parametric mixture model centered in a state-space model. The model allows for selecting different features of the series for clustering purposes. Shima Ghassempour et.al., (2014) described an algorithm for clustering multivariate time series with variables taking both categorical and continuous values. The proposed approach is based on Hidden Markov Models (HMMs), where each trajectory is mapped in to an HMM, then defined a suitable distance between HMMs and finally proceed to cluster the HMMs with a method based on a distance matrix.

Saeed Aghabozorgi and Teh Ying Wah (2014) employed a two-level fuzzy clustering strategy in order to achieve the objective of clustering. In the first level, upon dimensionality reduction by a symbolic representation, time series data are clustered in a high-level phase using the longest common subsequence as similarity measurement. Then, by utilizing an efficient method, prototypes are made based on constructed clusters and passed to the next level to be reused as initial centroids. Afterwards, a fuzzy clustering approach is utilized to justify the clusters precisely.

Jose A. Vilar and Juan M. Vilar (2013) proposed a time series clustering method based on comparing forecast densities for a sequence of $k > 1$ consecutive horizons. The unknown k-dimensional forecast densities can be non-parametrically approximated by using bootstrap procedures that mimic the generating processes without parametric restrictions. The authors evaluated the clustering procedure via simulation and are applied to a real dataset involving electricity prices series. Warren Liao,T., (2005) surveyed and summarized previous works that investigated the clustering of time series data in various application domains. The past research are organized into three groups depending upon whether the clustering process work directly with the raw data either in the time or frequency domain, indirectly with features extracted from the raw data, or indirectly with models built from the raw data.

Bandyopadhyay,S., et.al., (2010) used the Xie-Beni index and the C-means functional as objective functions to evaluate the cluster validity in a multi objective optimization framework. Elizabeth E. Holmes (2012) developed MARSS package for fitting multivariate autoregressive state-space models to time-series data. It is an R package for fitting linear multivariate autoregressive state-space (MARSS) models with Gaussian errors to time-series data. The MARSS package implements state-space models in a maximum likelihood framework. The core functionality of MARSS is based on likelihood

- *Vedavathi. K, Professor, E-mail: vedavathi_k@yahoo.com Department of Computer Science, GITAM University.india*
- *Srinivasa Rao. K, Professor, E-mail: ksraoau@yahoo.co.in Dept.of Statistics, Andhra University, Visakhapatnam, India*
- *Nirupama Devi. K, Professor, E-mail: knirupamadevi@gmail.com Dept.of Statistics, Andhra University, Visakhapatnam, India.*

maximization using the Kalman filter/smoother, combined with an EM algorithm.

Shusong Jin and Wai Keung Li (2006) considered a mixture autoregressive panel (MARP) model which can capture the multi-modal phenomenon in some data sets. EM algorithm is used to estimate the MARP model. Christoph Pamminger and Sylvia Frauhwirth-Schnatter (2010) discussed model-based clustering of categorical time series based on time-homogeneous first-order Markov chains with unknown transition matrices. In the Markov chain clustering approach the individual transition probabilities are fixed to a group specific transition matrix. The approach Dirichlet multinomial clustering assumed that within each group unobserved heterogeneity is still existent and is captured by allowing the individual transition matrices to deviate from the group means by describing this variation for each row through a Dirichlet distribution with unknown hyper parameters.

Pablo Montero (2014) described the use of R package TSclust to integrate different time series dissimilarity criteria in a single software package in order to check and compare their behavior in clustering. Luciana Alvim S. Romani et.al. (2014), presented an unsupervised algorithm called CLimate and rEmote sensing Association patterns Miner, for mining association patterns on heterogeneous time series from climate and remote sensing data integrated in a remote sensing information system developed to improve the monitoring of sugar cane fields. The proposed method can be used by agro meteorologists to mine and discover knowledge from their long time series of past and forecasting data, being a valuable tool to support their decision-making process.

Emrah Bulut et.al. (2012), ended a fuzzy integrated logical forecasting method (FILF) for multi-variate systems by using a vector autoregressive model. The forecasting results of the VAR-FILF approach are compared with mostly used FTS methods and traditional time series analysis. Marcella Corduas and Domenico Piccolo (2008) discussed the statistical properties of the AR distance by deriving the asymptotic distribution and an adequate approximation which is easily computable. Kavitha,V. and Punithavalli,M. (2010) presented a survey on various clustering algorithms available for time series datasets.

Eamonn Keogh et.al., (2006) defined time series discords, a new primitive for time series data mining and introduced an algorithm to efficiently find discords and also demonstrated their utility on a host of domains. Xiaozhe Wang et.al (2006) proposed a method for clustering of time series based on their structural characteristics. The proposed method does not cluster point values using a distance metric, rather it clusters based on global features extracted from the time series. The feature measures are obtained from each individual series are fed into arbitrary clustering algorithms, including an unsupervised neural network algorithm, self-organizing map, or hierarchal clustering algorithm.

In atmospheric science the statistical problems are frequently characterized by large spatio-temporal data sets and pose difficult computational challenges in classification and pattern recognition. The present work analyzes the time series data of the regions with the feature vector consisting of Temperature and Humidity. For this, the whole data can be characterized by considering mixture of bivariate AR(p) process. But the number of components in this mixture is unknown as a priori and it requires the regions are to be divided based on the data set of the bivariate feature vector. Hence in this paper, an unsupervised learning algorithm is developed and analyzed using bivariate AR(p) model and K-means algorithm. K-means algorithm is used to identify the number of regions according to feature vector. $\begin{bmatrix} Temperature \\ Humidity \end{bmatrix}$. The model parameters are estimated using the EM algorithm. The initial estimates of the parameters are obtained with K-means algorithm and least squares estimates. The unsupervised learning algorithm is developed by utilizing the conditional maximum likelihood of each class. The performance of the developed algorithm is studied through computing the performance measures like sensitivity, specificity, precision, recall, F-measure and misclassification rate. A comparative study of the developed algorithm with existing univariate autoregressive process of order p algorithm is also discussed.

2. FINITE MIXTURE OF BIVARIATE AR(P)

A bivariate process $\left\{\begin{pmatrix} X_t \\ Y_t \end{pmatrix}, t \in T\right\}$ is said to follow a bivariate autoregressive process of order p if it can be expressed as

$$Z_t = \mu + \Phi_{(1)}(Z_{t-1} - \mu) + \Phi_{(2)}(Z_{t-2} - \mu) + \ \dots + \Phi_{(p)}\left(Z_{t-p} - \mu\right) + \ \epsilon_t$$

where, $Z_t = \begin{pmatrix} X_t \\ Y_t \end{pmatrix}, \mu = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}, \Phi_{(j)} = \begin{pmatrix} \phi_{11}^{(j)} & \phi_{12}^{(j)} \\ \phi_{21}^{(j)} & \phi_{22}^{(j)} \end{pmatrix};$

$j = 1, 2, 3, \dots, p$ and $\epsilon_t = \begin{pmatrix} e_{X_t} \\ e_{Y_t} \end{pmatrix}$ and ϵ_t of Z_t , t = 1, 2, 3, . . . , N follows a bivariate Gaussian distribution with mean vector as null vector and variance – Covariance matrix as

$$\Sigma \ = \ \begin{pmatrix} \sigma_{e_{X_t}}^2 & \rho\sigma_{e_{X_t}}\sigma_{e_{Y_t}} \\ \rho\sigma_{e_{X_t}}\sigma_{e_{Y_t}} & \sigma_{e_{Y_t}}^2 \end{pmatrix}$$

The probability density function of ϵ_t is

$$f\left(e_{x_t}, e_{y_t}\right) = \left(\frac{1}{2\pi\,\sigma_{e_{X_t}}\,\sigma_{e_{Y_t}}\sqrt{1-\rho^2}}\right) \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{e_{X_t}^2}{\sigma_{e_{X_t}}{}^2} - \frac{2\rho e_{X_t} e_{Y_t}}{\sigma_{e_{X_t}}\sigma_{e_{Y_t}}} + \frac{e_{Y_t}^2}{\sigma_{e_{Y_t}}{}^2}\right)\right\},$$

$$-1 < \rho < 1, \ \sigma_{e_X} > 0 \ and \ \sigma_{e_Y} > 0 \qquad (2.1)$$

This process reduces to univariate AR(p) process if $\phi_{12}^{(j)}, \phi_{21}^{(j)}$ and $\phi_{22}^{(j)}$ for $j = 1, 2, 3, \dots, p$ are 0. This process also includes bivariate AR(1) process when $p = 1$.

The conditional likelihood function of the time series can be expressed as

$$P(Z_t, \Phi) = \prod_{u=p+1}^{n} f\left(\left(e_{X_{t,u}}, e_{Y_{t,u}}\right), \Phi\right) \qquad (2.2)$$

Since the residual terms in sample realization of time series starts from $u = p+1$ in bivariate AR(p) process , i.e.

$$e_{X_{t,u}} = X_{t,u} - \phi_{11}^{(1)} X_{t,u-p} - \phi_{12}^{(1)} Y_{t,u-p} - \phi_{11}^{(2)} X_{t,u-(p-1)}$$

$$-\phi_{12}^{(2)} Y_{t,u-(p-1)} - \ . \ . \ . \ - \phi_{11}^{(p)} X_{t,u-1} - \phi_{12}^{(p)} Y_{t,u-1}$$

$$= X_{t,u} - \sum_{j=1}^{p} \phi_{11}^{(j)} X_{t,u-(p-j+1)} - \sum_{j=1}^{p} \phi_{12}^{(j)} Y_{t,u-(p-j+1)},$$

$$u = p+1, p+2, \ldots, n$$

and

$$e_{Y_{t,u}} = Y_{t,u} - \phi_{21}^{(1)} X_{t,u-p} - \phi_{22}^{(1)} Y_{t,u-p} - \phi_{21}^{(2)} X_{t,u-(p-1)}$$

$$-\phi_{22}^{(2)} Y_{t,u-(p-1)} - \ . \ . \ . \ - \phi_{21}^{(p)} X_{t,u-1} - \phi_{22}^{(p)} Y_{t,u-1}$$

$$= Y_{t,u} - \sum_{j=1}^{p} \phi_{21}^{(j)} X_{t,u-(p-j+1)} - \sum_{j=1}^{p} \phi_{22}^{(j)} Y_{t,u-(p-j+1)}$$

$$u = p+1, p+2, \ldots, n$$

where, $\Phi = \left(\sigma_{e_X}^2, \ \sigma_{e_Y}^2, \rho, \phi_{11}^{(j)}, \phi_{12}^{(j)}, \phi_{21}^{(j)}, \phi_{22}^{(j)}, \ j = 1,2, \ldots, p\right)$ are the set of model parameters and n is the number of observations in a given time series realization D.

Substituting the value of $f(e_{X_t}, e_{Y_t})$ given in equation (2.1) in (2.2) one can obtain likelihood function of the time series as

$$P(Z_t, \Phi) = \left(\frac{1}{2\pi\sigma_{e_{X_{t,u}}} \sigma_{e_{Y_{t,u}}} \sqrt{1-\rho^2}}\right)^{n-p} exp\left\{-\frac{1}{2(1-\rho^2)} \sum_{u=p+1}^{n} \left(\frac{e_{X_{t,u}}^2}{\sigma_{e_X}^2} - \frac{2\rho e_{X_{t,u}} e_{Y_{t,u}}}{\sigma_{e_X}\sigma_{e_Y}} + \frac{e_{Y_{t,u}}^2}{\sigma_{e_Y}^2}\right)\right\} \quad (2.3)$$

The logarithm of the likelihood function is

$$P(Z_t, \Phi) = -(n-p) \ \ln\left(2\pi \ \sigma_{e_X} \ \sigma_{e_Y} \sqrt{1-\rho^2}\right) - \frac{1}{2(1-\rho^2)} \sum_{u=p+1}^{n} \left(\frac{e_{X_{t,u}}^2}{\sigma_{e_X}^2} - \frac{2\rho e_{X_{t,u}} e_{Y_{t,u}}}{\sigma_{e_X}\sigma_{e_Y}} + \frac{e_{Y_{t,u}}^2}{\sigma_{e_Y}^2}\right) \quad (2.4)$$

Now, let the whole time series data are generated by M different bivariate autoregressive process of order p which corresponds to M class labels of interest with weights $P(\omega_1), P(\omega_2), \ldots, P(\omega_M)$.

Let $P(Z_t|\omega_k, \Phi_k)$ denote the conditional likelihood function of k^{th} class with Φ_k as the set of parameters for the model. Then the conditional likelihood function of the mixture of finite bivariate AR process of order p can be expressed in the form of

$$P(Z_t| \Theta) = \sum_{k=1}^{M} P(Z_t|\omega_k, \Phi_k) P(\omega_k) \quad (2.5)$$

where, $\Theta = \{\Phi_1, \Phi_2, \ldots, \Phi_M, P(\omega_1), P(\omega_2), \ldots, P(\omega_M)\}$ represents the set of model parameters for the mixture model and $P(Z_t|\omega_k, \Phi_k)$ is as given in equation (2.3).

A bivariate time series Z_t is assigned to cluster ω_k with posterior probability

$P(\omega_k|Z_t)$ such that $\sum_{k=1}^{M} P(\omega_k) = 1$.

This model includes the mixture of univariate autoregressive process of order p when the model parameters $\phi_{21}{}^{(j)}, \phi_{12}{}^{(j)}$ and $\phi_{22}{}^{(j)}$; $j = 1,2,3, \ldots, p$ are zero for $k = 1,2, \ldots, M$. If $M = 1$, then this process becomes bivariate AR(p) process.

3. K- MEANS ALGORITHM

The K- means algorithm is a faster method to perform clustering of time series data. In K-means algorithm, continuous reassignments of objects into different clusters are done so that the distance within cluster is minimized. Let the time series are generated by M different bivariate AR(p) models which corresponds to M clusters of interest with weights $P(\omega_1), P(\omega_2), \ldots, P(\omega_M)$. Let $P(Z_t|\omega_k, \Phi_k)$ denote the conditional likelihood function of k^{th} cluster with the set of parameters as Φ_k for the model. Then the conditional likelihood function of mixture of finite bivariate AR process of order p can be expressed in the form

$$P(Z_t| \Theta) = \sum_{k=1}^{M} P(Z_t|\omega_k, \Phi_k) P(\omega_k) \quad (3.1)$$

where, $\Theta = \{\Phi_1, \Phi_2, \ldots, \Phi_M, P(\omega_1), (\omega_2), \ldots P(\omega_M)\}$ represents the set of model parameters for mixture model.

A bivariate time series Z_t is assigned to cluster ω_k with posterior probability $P(\omega_k|Z_t)$ such that $\sum_{k=1}^{M} P(\omega_k) = 1$. This model includes mixture of univariate autoregressive process of order p model when the model parameters $\phi_{12}^{(j)}, \phi_{21}^{(j)}$ and $\phi_{22}^{(j)}$ are zero for $k = 1,2, \ldots, M$. If $M = 1$, this process becomes Bivariate AR(p) process. The feature vector representing $\Phi = \left(\phi_{11}^{(j)}, \ \phi_{12}^{(j)}, \phi_{21}^{(j)}, \phi_{22}^{(j)}\right)$; $j = 1,2, \ldots, p$ for each bivariate time series is obtained through using least square method of estimation (Douglas C. Montgomery et.al., 1990).

For utilizing the K-means algorithm, each time series is reduced to the object vector $\Phi_t = \{\left(\phi_{11t}^{(j)}, \ \phi_{12t}^{(j)}, \phi_{21t}^{(j)}, \phi_{22t}^{(j)}\right); j = 1,2, \ldots, p$ of the parameters of bivariate AR(p) which are obtained using ordinary least square estimates of the model. If C_k is the center of the k^{th} cluster then K-means algorithm attempts to minimize the objective function $F = \sum_{t=1}^{N} \sum_{k=1}^{M} (\Phi_{t\,k} - C_k)^2$. The K-means algorithm also requires initial number of clusters, which can be obtained by plotting a bivariate scatter surface for all time series of training data. Based on the number of surfaces visualized, the initial value of M is obtained.

The K-means algorithm for obtaining the number of clusters is as follows:

Step 1: Identify the value of K.
Step 2: Initialize K cluster centers.
Step 3: Decide the class memberships of N objects by assigning them to nearest Cluster center.
Step 4: Re estimate the K cluster center by assuming the membership found above as correct.
Step 5: If none of the N objects changed membership in last iteration, then exit otherwise go to step 3.

For grouping the training time series data into M clusters the residual variances of each class, correlation coefficient between the attributes in each class and model parameters $\phi_{11}^{(j)}, \phi_{12}^{(j)}, \phi_{21}^{(j)}, \phi_{22}^{(j)}$; $j = 1,2, \ldots, p$ within each class are obtained using clustered time series of that class and taking average over the number of time series in that class.

4. ESTIMATION OF PARAMETERS USING EM ALGORITHM

This section presents the estimation of parameters using Expectation - Maximization algorithm. Here, without loss of generality, it is assumed that the time series values are standardized by taking each individual observation from their mean value and hence the mean function of each time series is considered as zero. Therefore the problem of estimation is to obtain the estimates of the parameters $\sigma^2_{e_X}, \sigma^2_{e_Y}, \rho, \phi^{(j)}_{11}, \phi^{(j)}_{12}, \phi^{(j)}_{21}$ and $\phi^{(j)}_{22}$; $j = 1, 2, 3, \ldots, p$. For estimating the parameters through EM algorithm, one can use likelihood function of the sample.

Let $D = \{Z_1, Z_2, \ldots, Z_N\}$ are N time series and assuming that these are conditionally independent under the given model, the likelihood function of realization is

$$P(D|\Theta) = \prod_{t=1}^{N} P(Z_t|\Theta) = \prod_{t=1}^{N}\sum_{k=1}^{M} P(Z_t|\omega_k, \Phi_k)P(\omega_k)$$

The log likelihood function of the realization D is

$$\ln P(D|\Theta) = \sum_{t=1}^{N} \ln\left[\sum_{k=1}^{M} P(Z_t|\omega_k, \Phi_k)\, P(\omega_k)\right]$$

In the first step of EM algorithm the expectation of log likelihood function is obtained. Given the observed data set D and current parameter estimate $\Theta(r)$, the expected value of log likelihood for M labels of time series each having n points of observations can be expressed as

$$Q(\Theta|\Theta(r)) = E\big(\ln P(D|\Theta)\big) = \sum_{t=1}^{N}\sum_{k=1}^{M} P(\omega_k|Z_t, \Theta(r)) \ln P(Z_t|\omega_k, \Phi_k) + \sum_{t=1}^{N}\sum_{k=1}^{M} P(\omega_k|Z_t, \Theta(r)) \ln P(\omega_k) \tag{4.1}$$

where, the posterior probabilities $P(\omega_k|Z_t, \Theta)$ are computed using the Bayes rule as

$$P(\omega_k|Z_t, \Theta) = \frac{P(Z_t|\omega_k, \Phi_k)P(\omega_k)}{\sum_{q=1}^{M} P(Z_t|\omega_q, \Phi_q)P(\omega_q)},$$

$$t = 1, 2, \ldots, N \text{ and } k\ 1, 2, \ldots, M. \tag{4.2}$$

This implies, $Q(\Theta|\Theta(r)) = \sum_{t=1}^{N}\sum_{k=1}^{M} P\big(\omega_k|Z_t, \Theta(r)\big)$

$$\left\{-(n-p)\ln\left(2\pi\,\sigma_{e_{X_k}}\sigma_{e_{Y_k}}\sqrt{1-\rho_k^2}\right) - \frac{1}{2(1-\rho_k^2)}\sum_{u=p+1}^{n}\left(\frac{e^2_{X_{t,u}}}{\sigma^2_{e_{X_k}}} - \frac{2\rho_k e_{X_{t,u}} e_{Y_{t,u}}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} + \frac{e^2_{Y_{t,u}}}{\sigma^2_{e_{Y_k}}}\right)\right\}$$
$$+ \sum_{t=1}^{N}\sum_{k=1}^{M} P\big(\omega_k|Z_t, \Theta(r)\big)\ \ln\left\{\frac{P(Z_t|\omega_k, \Phi_k)P(\omega_k)}{\sum_{q=1}^{M} P\big(Z_t|\omega_q, \Phi_q\big)P\big(\omega_q\big)}\right\} \tag{4.3}$$

In the M- step of EM algorithm, one has to find estimates of parameters by maximizing $Q(\Theta|\Theta(r))$.

The updated equation $P(\hat{\omega}_k)$ of the model is

$$P(\hat{\omega}_k) = \frac{1}{N}\sum_{t=1}^{N} P\big(\omega_k|Z_t, \Theta(r)\big) \qquad k = 1, 2, \ldots, M \tag{4.4}$$

For estimating $\sigma^2_{e_{X_k}}, \sigma^2_{e_{Y_k}}, \phi^{(j)}_{11k}, \phi^{(j)}_{12k}, \phi^{(j)}_{21k}, \phi^{(j)}_{22k}$; $j = 1, 2, \ldots, p$ and ρ_k the updated equations respectively are

$$\sum_{t=1}^{N} P\big(\omega_k|\, Z_t, \Theta(r)\big)\Bigg[(n-p) - \frac{1}{(1-\rho_k^2)}\sum_{u=p+1}^{n}\Bigg(\frac{\left(X_{t,u} - \sum_{j=1}^{p}\phi^{(j)}_{11k}X_{t,u-(p-j+1)} - \sum_{j=1}^{p}\phi^{(j)}_{12k}Y_{t,u-(p-j+1)}\right)^2}{\sigma^2_{e_{X_k}}}$$
$$- \frac{\rho_k}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}}\left(X_{t,u} - \sum_{j=1}^{p}\phi^{(j)}_{11k}X_{t,u-(p-j+1)} - \sum_{j=1}^{p}\phi^{(j)}_{12k}Y_{t,u-(p-j+1)}\right)$$
$$\left(Y_{t,u} - \sum_{j=1}^{p}\phi^{(j)}_{21k}X_{t,u-(p-j+1)} - \sum_{j=1}^{p}\phi^{(j)}_{22k}Y_{t,u-(p-j+1)}\right)\Bigg)\Bigg] = 0 \tag{4.5}$$

$$\sum_{t=1}^{N} P\big(\omega_k|Z_t, \Theta(r)\big)\Bigg[(n-p) - \frac{1}{(1-\rho_k^2)}\sum_{u=p+1}^{n}\Bigg(\frac{\left(Y_{t,u} - \sum_{j=1}^{p}\phi^{(j)}_{21k}X_{t,u-(p-j+1)} - \sum_{j=1}^{p}\phi^{(j)}_{22k}Y_{t,u-(p-j+1)}\right)^2}{\sigma^2_{e_{Y_k}}}$$
$$- \frac{\rho_k}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}}\left(X_{t,u} - \sum_{j=1}^{p}\phi^{(j)}_{11k}X_{t,u-(p-j+1)} - \sum_{j=1}^{p}\phi^{(j)}_{12k}Y_{t,u-(p-j+1)}\right)$$
$$\left(Y_{t,u} - \sum_{j=1}^{p}\phi^{(j)}_{21k}X_{t,u-(p-j+1)} - \sum_{j=1}^{p}\phi^{(j)}_{22k}Y_{t,u-(p-j+1)}\right)\Bigg)\Bigg] = 0 \tag{4.6}$$

$$\sum_{t=1}^{N} P\big(\omega_k|Z_t, \Theta(r)\big)$$
$$\sum_{u=p+1}^{n}\Bigg\{\frac{1}{\sigma_{e_{X_k}}}X_{t,u-(p-j+1)}X_{t,u} - \frac{\rho_k}{\sigma_{e_{Y_k}}}X_{t,u-(p-j+1)}Y_{t,u}$$
$$- \left[\sum_{j=1}^{p}\left(\frac{\phi^{(j)}_{11k}}{\sigma_{e_{X_k}}} - \frac{\rho_k\phi^{(j)}_{21k}}{\sigma_{e_{Y_k}}}\right)X_{t,u-(p-j+1)}\right]X_{t,u-(p-j+1)}$$
$$- \left[\sum_{j=1}^{p}\left(\frac{\phi^{(j)}_{12k}}{\sigma_{e_{X_k}}} - \frac{\rho_k\phi^{(j)}_{22k}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-(p-j+1)}\right]X_{t,u-(p-j+1)}\Bigg\} = \tag{4.7}$$
$$\sum_{t=1}^{N} P\big(\omega_k|Z_t, \Theta(r)\big)\sum_{u=p+1}^{n}\Bigg\{\frac{1}{\sigma_{e_{X_k}}}X_{t,u}Y_{t,u-(p-j+1)} - \frac{\rho_k}{\sigma_{e_{Y_k}}}Y_{t,u-(p-j+1)}Y_{t,u}$$

$$- \left[\sum_{j=1}^{p} \left(\frac{\phi_{11k}^{(j)}}{\sigma_{e_{X_k}}} - \frac{\rho_k \phi_{21k}^{(j)}}{\sigma_{e_{Y_k}}} \right) X_{t,u-(p-j+1)} \right] Y_{t,u-(p-j+1)}$$

$$- \left[\sum_{j=1}^{p} \left(\frac{\phi_{12k}^{(j)}}{\sigma_{e_{X_k}}} - \frac{\rho_k \phi_{22k}^{(j)}}{\sigma_{e_{Y_k}}} \right) Y_{t,u-(p-j+1)} \right] Y_{t,u-(p-j+1)} \Bigg\} = 0 \qquad (4.8)$$

$$\sum_{t=1}^{N} P(\omega_k|Z_t,\Theta(r))$$

$$\sum_{u=p+1}^{n} \Bigg\{ \frac{\rho_k}{\sigma_{e_{X_k}}} X_{t,u-(p-j+1)} X_{t,u} - \frac{1}{\sigma_{e_{Y_k}}} X_{t,u-(p-j+1)} Y_{t,u}$$

$$- \left[\sum_{j=1}^{p} \left(\frac{\rho_k \phi_{11k}^{(j)}}{\sigma_{e_{X_k}}} - \frac{\phi_{21k}^{(j)}}{\sigma_{e_{Y_k}}} \right) X_{t,u-(p-j+1)} \right] X_{t,u-(p-j+1)}$$

$$- \left[\sum_{j=1}^{p} \left(\frac{\rho_k \phi_{12k}^{(j)}}{\sigma_{e_{X_k}}} - \frac{\phi_{22k}^{(j)}}{\sigma_{e_{Y_k}}} \right) Y_{t,u-(p-j+1)} \right] X_{t,u-(p-j+1)} \Bigg\} = 0 \qquad (4.9)$$

$$\sum_{t=1}^{N} P(\omega_k|Z_t,\Theta(r)) \sum_{u=p+1}^{n} \Bigg\{ \frac{\rho_k}{\sigma_{e_{X_k}}} X_{t,u} Y_{t,u-(p-j+1)}$$

$$- \frac{1}{\sigma_{e_{Y_k}}} Y_{t,u-(p-j+1)} Y_{t,u}$$

$$- \left[\sum_{j=1}^{p} \left(\frac{\rho_k \phi_{11k}^{(j)}}{\sigma_{e_{X_k}}} - \frac{\phi_{21k}^{(j)}}{\sigma_{e_{Y_k}}} \right) X_{t,u-(p-j+1)} \right] Y_{t,u-(p-j+1)}$$

$$- \left[\sum_{j=1}^{p} \left(\frac{\rho_k \phi_{12k}^{(j)}}{\sigma_{e_{X_k}}} - \frac{\phi_{22k}^{(j)}}{\sigma_{e_{Y_k}}} \right) Y_{t,u-(p-j+1)} \right] Y_{t,u-(p-j+1)} \Bigg\} = 0 \qquad (4.10)$$

$$\sum_{t=1}^{N} P(\omega_k|Z_t,\Theta(r)) \Bigg[(n-p)\rho_k -$$

$$\frac{\rho_k}{(1-\rho_k^2)} \sum_{u=p+1}^{n} \left(\frac{\left(X_{t,u} - \sum_{j=1}^{p} \phi_{11k}^{(j)} X_{t,u-(p-j+1)} - \sum_{j=1}^{p} \phi_{12k}^{(j)} Y_{t,u-(p-j+1)} \right)^2}{\sigma_{e_{X_k}}^2} \right.$$

$$\left. \left(Y_{t,u} - \sum_{j=1}^{p} \phi_{21k}^{(j)} X_{t,u-(p-j+1)} - \sum_{j=1}^{p} \phi_{22k}^{(j)} Y_{t,u-(p-j+1)} \right)^2 \right) \Bigg]$$

$$+ \left(\frac{1+\rho_k^2}{1-\rho_k^2} \right) \sum_{u=p+1}^{n} \frac{\left[X_{t,u} - \sum_{j=1}^{p} \phi_{11k}^{(j)} X_{t,u-(p-j+1)} - \sum_{j=1}^{p} \phi_{12k}^{(j)} Y_{t,u-(p-j+1)} \right]}{\sigma_{e_{X_k}}}$$

$$\left. \frac{\left[Y_{t,u} - \sum_{j=1}^{p} \phi_{21k}^{(j)} X_{t,u-(p-j+1)} - \sum_{j=1}^{p} \phi_{22k}^{(j)} Y_{t,u-(p-j+1)} \right]}{\sigma_{e_{Y_k}}} \right\} = 0 \qquad (4.11)$$

Solving the equations (4.4), (4.5), (4.6), (4.7), (4.8), (4.9), (4.10) and (4.11) simultaneously and iteratively the refined estimates of model parameters $P(\omega_k), \sigma^2_{e_{X_k}}, \sigma^2_{e_{Y_k}}, \phi^{(j)}_{11k}, \phi^{(j)}_{12k}, \phi^{(j)}_{21k}$ and $\phi^{(j)}_{22k}$, $j = 1,2,3,\ldots,p$ and ρ_k can be obtained.

Expectation Maximization Algorithm

Step 1: Find the initial parameters using equations (5.1) to (5.3) given in section 5.

Step 2: Obtain revised estimates of the parameters $P(\omega_k), \sigma^2_{e_{X_k}}, \sigma^2_{e_{Y_k}}, \phi^{(j)}_{11k}, \phi^{(j)}_{12k}, \phi^{(j)}_{21k}, \phi^{(j)}_{22k}$; $j = 1,2,3,\ldots,p$ and ρ_k using equations (4.4), (4.5), (4.6), (4.7), (4.8), (4.9), (4.10) and (4.11).

Step 3: Repeat the process until the parameters do not change or the difference in successive computations is within given threshold value.

Step 4: Write final estimates of parameters $P(\omega_k), \sigma^2_{e_{X_k}}, \sigma^2_{e_{Y_k}}, \phi^{(j)}_{11k}, \phi^{(j)}_{12k}, \phi^{(j)}_{21k}, \phi^{(j)}_{22k}$; $j = 1,2,3,\ldots,p$ and ρ_k.

5. INITIALISATION OF PARAMETERS

To initialize the parameters $\sigma^2_{e_{X_k}}, \sigma^2_{e_{Y_k}}$ and ρ_k using K-means algorithm one can obtain the number of clusters of the training data set. After classifying the time series data into M clusters, the residual variance of each attribute for each class, the correlation coefficient between the attributes in each class and the model parameters with in the class are obtained using the clustered time series of that class.

Thus the initial estimates of $\sigma^2_{e_{X_k}}$ and $\sigma^2_{e_{Y_k}}$ can be taken as

$$\hat{\sigma}^2_{e_{X_k}} = \frac{1}{(n-p)\,G_k} \sum_{g=1}^{G_k} \sum_{u=2}^{n} e^{2(g)}_{X_{t,u}}$$

$$\text{and } \hat{\sigma}^2_{e_{Y_k}} = \frac{1}{(n-p)\,G_k} \sum_{g=1}^{G_k} \sum_{u=2}^{n} e^{2(g)}_{Y_{t,u}}, \quad \text{for } k = 1,2,\ldots,M \qquad (5.1)$$

where, G_k is the number of time series in the k^{th} cluster.

The initial estimate of ρ_k, the correlation coefficient between X_t and Y_t is given by

$$\hat{\rho}_k = \frac{1}{G_k} \sum_{g=1}^{G_k} \frac{Cov^{(g)}(X_t,Y_t)}{\sqrt{Var^{(g)}(X_t)Var^{(g)}(Y_t)}}, \quad k = 1,2,\ldots,M \qquad (5.2)$$

where, G_k is the number of time series in the k^{th} cluster.

For initializing parameters $\phi^{(j)}_{11}$, $\phi^{(j)}_{12}$, $\phi^{(j)}_{21}$ and $\phi^{(j)}_{22}$; $j = 1,2,3,\ldots,p$, ordinary least squares estimates of the bivariate auto-regressive process of order p are obtained for each time series data in each class and averaged over the number of time series in that class. The sample realization of bivariate autoregressive process of order p can be written in the following format:

ISSN 2229-5518

$$\begin{bmatrix} X_{t,p+1} & Y_{t,p+1} \\ X_{t,p+2} & Y_{t,p+2} \\ . & . \\ . & . \\ . & . \\ X_{t,n} & Y_{t,n} \end{bmatrix}$$

$$= \begin{bmatrix} X_{t,1} & Y_{t,1} & X_{t,2} & Y_{t,2} & . & . & . & X_{t,p} & Y_{t,p} \\ X_{t,2} & Y_{t,2} & X_{t,3} & Y_{t,3} & . & . & . & X_{t,p+1} & Y_{t,p+1} \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ X_{t,n-p} & Y_{t,n-p} & X_{t,n-(p-1)} & Y_{t,n-(p-1)} & . & . & . & X_{t,n-1} & Y_{t,n-1} \end{bmatrix}.$$

$$\begin{bmatrix} \phi_{11t}^{(1)} & \phi_{21t}^{(1)} \\ \phi_{12t}^{(1)} & \phi_{22t}^{(1)} \\ \phi_{11t}^{(2)} & \phi_{21t}^{(2)} \\ \phi_{12t}^{(2)} & \phi_{22t}^{(2)} \\ . & . \\ . & . \\ . & . \\ \phi_{11t}^{(p)} & \phi_{21t}^{(p)} \\ \phi_{12t}^{(p)} & \phi_{22t}^{(p)} \end{bmatrix} + \begin{bmatrix} e_{X_{t,p+1}} & e_{Y_{t,p+1}} \\ e_{X_{t,p+2}} & e_{Y_{t,p+2}} \\ . & . \\ . & . \\ . & . \\ e_{X_{t,n}} & e_{Y_{t,n}} \end{bmatrix}$$

which can be represented as

$$\Upsilon_t = X_t \Phi_t + \xi_t$$

where,

$$\Upsilon_t = \begin{bmatrix} X_{t,p+1} & Y_{t,p+1} \\ X_{t,p+2} & Y_{t,p+2} \\ . & . \\ . & . \\ . & . \\ X_{t,n} & Y_{t,n} \end{bmatrix}, \quad X_t =$$

$$\begin{bmatrix} X_{t,1} & Y_{t,1} & X_{t,2} & Y_{t,2} & . & . & . & X_{t,p} & Y_{t,p} \\ X_{t,2} & Y_{t,2} & X_{t,3} & Y_{t,3} & . & . & . & X_{t,p+1} & Y_{t,p+1} \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ X_{t,n-p} & Y_{t,n-p} & X_{t,n-(p-1)} & Y_{t,n-(p-1)} & . & . & . & X_{t,n-1} & Y_{t,n-1} \end{bmatrix}.$$

$$\Phi_t = \begin{bmatrix} \phi_{11t}^{(1)} & \phi_{21t}^{(1)} \\ \phi_{12t}^{(1)} & \phi_{22t}^{(1)} \\ \phi_{11t}^{(2)} & \phi_{21t}^{(2)} \\ \phi_{12t}^{(2)} & \phi_{22t}^{(2)} \\ . & . \\ . & . \\ . & . \\ \phi_{11t}^{(p)} & \phi_{21t}^{(p)} \\ \phi_{12t}^{(p)} & \phi_{22t}^{(p)} \end{bmatrix} \text{ and } \xi_t = \begin{bmatrix} e_{X_{t,p+1}} & e_{Y_{t,p+1}} \\ e_{X_{t,p+2}} & e_{Y_{t,p+2}} \\ . & . \\ . & . \\ . & . \\ e_{X_{t,n}} & e_{Y_{t,n}} \end{bmatrix}$$

Then the ordinary least square estimate of Φ_t is

$$\widehat{\Phi}_t = (X_t^T X_t)^{-1} X_t^T \Upsilon_t \quad \text{for } t = 1, 2, \; . . . \; , N$$

6. MODEL BASED CLUSTERING ALGORITHM

Here, the clustering algorithm is presented for identifying the new time series data with one of the available clusters. The steps involved in this algorithm are as follows:

Step 1: Draw scatter surface diagram for training data set in order to obtain the initial number of clusters for using in the K-means algorithm.

Step 2: Through K means algorithm obtain the refined num ber of classes and elements in each cluster.

Step 3: Obtain the initial estimates of model parameters using bivariate AR(p) process and the least square method of estimation given in section 5.

Step 4: Obtain the refined estimates of model parameters using updated equations of EM algorithm for bivariate AR(p) process given in section 4.

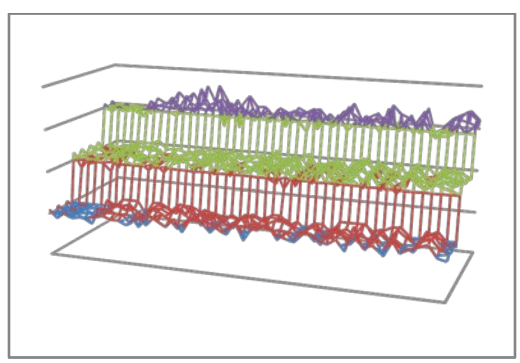
Step 5: For new Time series Bivariate dataset, compute the conditional Liklihood with model parameters of i^{th} class derived from step 4 over 1 to k and assign it to a class for which the component conditional Likelihood is maximum.

7. EXPERIMENTAL RESULTS AND PERFOANCE EVALUATION

In this section, the utility of developed algorithm for clustering regions is demonstrated. After discussion with the people in Cyclone Warning Centre at Visakhapatnam, it is understood that two variables Temperature and Humidity levels are most important for taking control measures of weather forecast of a region. To understand the dynamics of weather in the regions, it is required to segment the total area into various regions based on the vector $\begin{bmatrix}\text{Temperature}\\ \text{Humidity}\end{bmatrix}$. The number of regions in an area is not known as a priori and requires clustering methods to identify a sample time series with the region. Hence a study is carried out by collecting the bivariate variable $\begin{bmatrix}\text{Temperature}\\ \text{Humidity}\end{bmatrix}$ over a period with a sample collected over 100 locations covering the whole area.

Using the K-means algorithm, the number of regions is determined. For implementing K-means algorithm, the initial number of regions is required. Using training data the bivariate time series are plotted in scatter responses through a 3-dimensional graph shown in figure 7.1.

Figure 7.1



Scatter plot for Temperature and Humidity against time

From figure 7.1, it is observed that the time series are

ISSN 2229-5518

stationary after obtaining deviation from mean values and it is also observed that there are three surfaces which represent three regions. Taking the initial value of number of clusters as three, the K-means algorithm is implemented on the training dataset of 100 time series and found that the number of regions in that area also represent three levels. The whole training data is classified into three segments representing three regions

For developing BAR(2) model for each region, one can obtain the updated equations of the model parameters for EM algorithm. Taking p = 2 in equations (4.11) to (4.20) one can obtain the updated equations of the model parameters of $\sigma^2_{e_{X_k}}, \sigma^2_{e_{Y_k}}$, $\phi^{(1)}_{11k}, \phi^{(1)}_{12k}, \phi^{(1)}_{21k}, \phi^{(1)}_{22k}$, $\phi^{(2)}_{11k}, \phi^{(2)}_{12k}, \phi^{(2)}_{21k}, \phi^{(2)}_{22k}$ and ρ_k respectively as

$$\sum_{t=1}^{N} P\big(\omega_k | Z_t, \Theta(r)\big) \Big\{ (n-2)$$

$$- \frac{1}{(1-\rho_k^2)} \sum_{u=3}^{n} \left[\frac{X_{t,u}^2}{\sigma^2_{e_{X_k}}} + \left(\frac{\left(\phi^{(1)}_{11k}\right)^2}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(1)}_{11k}\phi^{(1)}_{21k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X^2_{t,u-1} \right.$$

$$+ \left(\frac{\left(\phi^{(2)}_{11k}\right)^2}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(2)}_{11k}\phi^{(2)}_{21k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X^2_{t,u-2}$$

$$+ \left(\frac{\left(\phi^{(1)}_{12k}\right)^2}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(1)}_{12k}\phi^{(1)}_{22k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) Y^2_{t,u-1}$$

$$+ \left(\frac{\left(\phi^{(2)}_{12k}\right)^2}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(2)}_{12k}\phi^{(2)}_{22k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) Y^2_{t,u-2} - \left(\frac{2\phi^{(1)}_{11k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(1)}_{21k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-1}X_{t,u}$$

$$- \left(\frac{2\phi^{(2)}_{11k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(2)}_{21k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-2}X_{t,u}$$

$$+ \left(\frac{2\phi^{(1)}_{11k}\phi^{(2)}_{11k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \left(\phi^{(1)}_{11k}\phi^{(2)}_{21k} + \phi^{(2)}_{11k}\phi^{(1)}_{21k}\right)}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-1}X_{t,u-2}$$

$$- \left(\frac{2\phi^{(1)}_{12k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(1)}_{22k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u}Y_{t,u-1} - \left(\frac{2\phi^{(2)}_{12k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(2)}_{22k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u}Y_{t,u-2}$$

$$+ \left(\frac{2\phi^{(1)}_{11k}\phi^{(1)}_{12k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \left(\phi^{(1)}_{11k}\phi^{(1)}_{22k} + \phi^{(1)}_{12k}\phi^{(1)}_{21k}\right)}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-1}Y_{t,u-1}$$

$$+ \left(\frac{2\phi^{(1)}_{11k}\phi^{(2)}_{12k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \left(\phi^{(1)}_{11k}\phi^{(2)}_{22k} + \phi^{(2)}_{12k}\phi^{(1)}_{21k}\right)}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-1}Y_{t,u-2}$$

$$+ \left(\frac{2\phi^{(1)}_{12k}\phi^{(2)}_{11k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \left(\phi^{(1)}_{12k}\phi^{(2)}_{21k} + \phi^{(2)}_{11k}\phi^{(1)}_{22k}\right)}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-2}Y_{t,u-1}$$

$$+ \left(\frac{2\phi^{(2)}_{11k}\phi^{(2)}_{12k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \left(\phi^{(2)}_{11k}\phi^{(2)}_{22k} + \phi^{(2)}_{12k}\phi^{(2)}_{21k}\right)}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-2}Y_{t,u-2}$$

$$+ \left(\frac{2\phi^{(1)}_{12k}\phi^{(2)}_{12k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \left(\phi^{(1)}_{12k}\phi^{(2)}_{22k} + \phi^{(2)}_{12k}\phi^{(1)}_{22k}\right)}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) Y_{t,u-1}Y_{t,u-2}$$

$$- \frac{\rho_k}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \Big(X_{t,u} - \phi^{(1)}_{11k}X_{t,u-1} - \phi^{(2)}_{11k}X_{t,u-2} - \phi^{(1)}_{12k}Y_{t,u-1}$$

$$- \phi^{(2)}_{12k}Y_{t,u-2}\Big)Y_{t,u} \Big\} \qquad (7.1)$$

$$\sum_{t=1}^{N} P\big(\omega_k | Z_t, \Theta(r)\big) \Big\{ (n-2)$$

$$- \frac{1}{(1-\rho_k^2)} \sum_{u=3}^{n} \left[\frac{Y_{t,u}^2}{\sigma^2_{e_{X_k}}} + \left(\frac{\left(\phi^{(1)}_{21k}\right)^2}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(1)}_{11k}\phi^{(1)}_{21k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X^2_{t,u-1} \right.$$

$$+ \left(\frac{\left(\phi^{(2)}_{21k}\right)^2}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(2)}_{11k}\phi^{(2)}_{21k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X^2_{t,u-2}$$

$$+ \left(\frac{\left(\phi^{(1)}_{22k}\right)^2}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(1)}_{12k}\phi^{(1)}_{22k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) Y^2_{t,u-1}$$

$$+ \left(\frac{\left(\phi^{(2)}_{22k}\right)^2}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(2)}_{12k}\phi^{(2)}_{22k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) Y^2_{t,u-2}$$

$$- \left(\frac{2\phi^{(1)}_{21k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(1)}_{11k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-1}Y_{t,u}$$

$$- \left(\frac{2\phi^{(2)}_{21k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(2)}_{11k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-2}Y_{t,u}$$

$$+ \left(\frac{2\phi^{(1)}_{21k}\phi^{(2)}_{21k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \left(\phi^{(1)}_{11k}\phi^{(2)}_{21k} + \phi^{(2)}_{11k}\phi^{(1)}_{21k}\right)}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-1}X_{t,u-2}$$

$$- \left(\frac{2\phi^{(1)}_{22k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(1)}_{12k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) Y_{t,u}Y_{t,u-1} - \left(\frac{2\phi^{(2)}_{22k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \phi^{(2)}_{12k}}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) Y_{t,u}Y_{t,u-2}$$

$$+ \left(\frac{2\phi^{(1)}_{21k}\phi^{(1)}_{22k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \left(\phi^{(1)}_{11k}\phi^{(1)}_{22k} + \phi^{(1)}_{12k}\phi^{(1)}_{21k}\right)}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-1}Y_{t,u-1}$$

$$+ \left(\frac{2\phi^{(1)}_{21k}\phi^{(2)}_{22k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \left(\phi^{(1)}_{11k}\phi^{(2)}_{22k} + \phi^{(2)}_{12k}\phi^{(1)}_{21k}\right)}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-1}Y_{t,u-2}$$

$$+ \left(\frac{2\phi^{(2)}_{21k}\phi^{(1)}_{22k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \left(\phi^{(2)}_{11k}\phi^{(1)}_{22k} + \phi^{(1)}_{12k}\phi^{(2)}_{21k}\right)}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-2}Y_{t,u-1}$$

$$+ \left(\frac{2\phi^{(2)}_{21k}\phi^{(2)}_{22k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \left(\phi^{(2)}_{11k}\phi^{(2)}_{22k} + \phi^{(2)}_{12k}\phi^{(2)}_{21k}\right)}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) X_{t,u-2}Y_{t,u-2}$$

$$+ \left(\frac{2\phi^{(1)}_{22k}\phi^{(2)}_{22k}}{\sigma^2_{e_{X_k}}} - \frac{\rho_k \left(\phi^{(1)}_{12k}\phi^{(2)}_{22k} + \phi^{(2)}_{12k}\phi^{(1)}_{22k}\right)}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \right) Y_{t,u-1}Y_{t,u-2}$$

$$- \frac{\rho_k}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}} \Big(Y_{t,u} - \phi^{(1)}_{21k}X_{t,u-1} - \phi^{(2)}_{21k}X_{t,u-2} - \phi^{(1)}_{22k}Y_{t,u-1}$$

$$\left. - \phi^{(2)}_{22k}Y_{t,u-2}\Big)X_{t,u} \right] \Big\} \qquad (7.2)$$

$$\sum_{t=1}^{N} P\big(\omega_k|Z_t,\Theta(r)\big)\sum_{u=3}^{n}\Bigg[\left(\frac{X_{t,u}}{\sigma_{e_{X_k}}}-\frac{\rho_k Y_{t,u}}{\sigma_{e_{Y_k}}}\right)X_{t,u-1}$$
$$-\left\{X_{t,u-1}\left(\frac{\phi_{11k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{21k}^{(1)}}{\sigma_{e_{Y_k}}}\right)+X_{t,u-2}\left(\frac{\phi_{11k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{21k}^{(2)}}{\sigma_{e_{Y_k}}}\right)\right\}X_{t,u-1}$$
$$-\left\{\left(\frac{\phi_{12k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{22k}^{(1)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-1}+\left(\frac{\phi_{12k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{22k}^{(2)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-2}\right\}X_{t,u-1}\Bigg]=0 \tag{7.3}$$

$$\sum_{t=1}^{N} P\big(\omega_k|Z_t,\Theta(r)\big)\sum_{u=3}^{n}\Bigg[\left(\frac{X_{t,u}}{\sigma_{e_{X_k}}}-\frac{\rho_k}{\sigma_{e_{Y_k}}}Y_{t,u}\right)Y_{t,u-1}$$
$$-\left\{X_{t,u-1}\left(\frac{\phi_{11k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{21k}^{(1)}}{\sigma_{e_{Y_k}}}\right)+X_{t,u-2}\left(\frac{\phi_{11k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{21k}^{(2)}}{\sigma_{e_{Y_k}}}\right)\right\}Y_{t,u-1}$$
$$-\left\{\left(\frac{\phi_{12k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{22k}^{(1)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-1}+\left(\frac{\phi_{12k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{22k}^{(2)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-2}\right\}Y_{t,u-1}\Bigg]=0 \tag{7.4}$$

$$\sum_{t=1}^{N} P\big(\omega_k|Z_t,\Theta(r)\big)\sum_{u=3}^{n}\Bigg[\left(\frac{\rho_k}{\sigma_{e_{X_k}}}X_{t,u}-\frac{Y_{t,u}}{\sigma_{e_{Y_k}}}\right)X_{t,u-1}$$
$$-\left\{X_{t,u-1}\left(\frac{\rho_k\phi_{11k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\phi_{21k}^{(1)}}{\sigma_{e_{Y_k}}}\right)+X_{t,u-2}\left(\frac{\rho_k\phi_{11k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\phi_{21k}^{(2)}}{\sigma_{e_{Y_k}}}\right)\right\}X_{t,u-1}$$
$$-\left\{\left(\frac{\rho_k\phi_{12k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\phi_{22k}^{(1)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-1}+\left(\frac{\rho_k\phi_{12k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\phi_{22k}^{(2)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-2}\right\}X_{t,u-1}\Bigg]=0 \tag{7.5}$$

$$\sum_{t=1}^{N} P\big(\omega_k|Z_t,\Theta(r)\big)\sum_{u=3}^{n}\Bigg[\left(\frac{\rho_k}{\sigma_{e_{X_k}}}X_{t,u}-\frac{Y_{t,u}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-1}$$
$$-\left\{X_{t,u-1}\left(\frac{\rho_k\phi_{11k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\phi_{21k}^{(1)}}{\sigma_{e_{Y_k}}}\right)+X_{t,u-2}\left(\frac{\rho_k\phi_{11k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\phi_{21k}^{(2)}}{\sigma_{e_{Y_k}}}\right)\right\}Y_{t,u-1}$$
$$-\left\{\left(\frac{\rho_k\phi_{12k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\phi_{22k}^{(1)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-1}+\left(\frac{\rho_k\phi_{12k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\phi_{22k}^{(2)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-2}\right\}Y_{t,u-1}\Bigg]=0 \tag{7.6}$$

$$\sum_{t=1}^{N} P\big(\omega_k|Z_t,\Theta(r)\big)\sum_{u=3}^{n}\Bigg[\left(\frac{X_{t,u}}{\sigma_{e_{X_k}}}-\frac{\rho_k}{\sigma_{e_{Y_k}}}Y_{t,u}\right)X_{t,u-2}$$
$$\sum_{t=1}^{N} P\big(\omega_k|Z_t,\Theta(r)\big)\sum_{u=3}^{n}\Bigg[\left(\frac{X_{t,u}}{\sigma_{e_{X_k}}}-\frac{\rho_k}{\sigma_{e_{Y_k}}}Y_{t,u}\right)X_{t,u-2}$$
$$-\left\{\left(\frac{\phi_{12k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{22k}^{(1)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-1}+\left(\frac{\phi_{12k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{22k}^{(2)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-2}\right\}X_{t,u-2}\Bigg]=0 \tag{7.7}$$

$$\sum_{t=1}^{N} P\big(\omega_k|Z_t,\Theta(r)\big)\sum_{u=3}^{n}\Bigg[\left(\frac{X_{t,u}}{\sigma_{e_{X_k}}}-\frac{\rho_k}{\sigma_{e_{Y_k}}}Y_{t,u}\right)Y_{t,u-2}$$
$$-\left\{X_{t,u-1}\left(\frac{\phi_{11k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{21k}^{(1)}}{\sigma_{e_{Y_k}}}\right)+X_{t,u-2}\left(\frac{\phi_{11k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{21k}^{(2)}}{\sigma_{e_{Y_k}}}\right)\right\}Y_{t,u-2}$$
$$-\left\{\left(\frac{\phi_{12k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{22k}^{(1)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-1}+\left(\frac{\phi_{12k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\rho_k\phi_{22k}^{(2)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-2}\right\}Y_{t,u-2}\Bigg]=0 \tag{7.8}$$

$$\sum_{t=1}^{N} P\big(\omega_k|Z_t,\Theta(r)\big)\sum_{u=3}^{n}\Bigg[\left(\frac{\rho_k}{\sigma_{e_{X_k}}}X_{t,u}-\frac{Y_{t,u}}{\sigma_{e_{Y_k}}}\right)X_{t,u-2}$$
$$-\left\{X_{t,u-1}\left(\frac{\rho_k\phi_{11k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\phi_{21k}^{(1)}}{\sigma_{e_{Y_k}}}\right)+X_{t,u-2}\left(\frac{\rho_k\phi_{11k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\phi_{21k}^{(2)}}{\sigma_{e_{Y_k}}}\right)\right\}X_{t,u-2}$$
$$-\left\{\left(\frac{\rho_k\phi_{12k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\phi_{22k}^{(1)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-1}+\left(\frac{\rho_k\phi_{12k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\phi_{22k}^{(2)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-2}\right\}X_{t,u-2}\Bigg]=0 \tag{7.9}$$

$$\sum_{t=1}^{N} P\big(\omega_k|Z_t,\Theta(r)\big)\sum_{u=3}^{n}\Bigg[\left(\frac{\rho_k}{\sigma_{e_{X_k}}}X_{t,u}-\frac{Y_{t,u}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-1}$$
$$-\left\{X_{t,u-1}\left(\frac{\rho_k\phi_{11k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\phi_{21k}^{(1)}}{\sigma_{e_{Y_k}}}\right)+X_{t,u-2}\left(\frac{\rho_k\phi_{11k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\phi_{21k}^{(2)}}{\sigma_{e_{Y_k}}}\right)\right\}Y_{t,u-1}$$
$$-\left\{\left(\frac{\rho_k\phi_{12k}^{(1)}}{\sigma_{e_{X_k}}}-\frac{\phi_{22k}^{(1)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-1}+\left(\frac{\rho_k\phi_{12k}^{(2)}}{\sigma_{e_{X_k}}}-\frac{\phi_{22k}^{(2)}}{\sigma_{e_{Y_k}}}\right)Y_{t,u-2}\right\}Y_{t,u-1}\Bigg]=0 \tag{7.10}$$

$$\sum_{t=1}^{N} P\big(\omega_k|Z_t,\Theta(r)\big)$$
$$\Bigg\{(n-2)-\frac{1}{1-\rho_k^2}\sum_{u=3}^{n}\Bigg[\frac{X_{t,u}^2}{\sigma_{e_{X_k}}^2}+\frac{Y_{t,u}^2}{\sigma_{e_{Y_k}}^2}+\left(\frac{\left(\phi_{11k}^{(1)}\right)^2}{\sigma_{e_{X_k}}^2}+\frac{\left(\phi_{21k}^{(1)}\right)^2}{\sigma_{e_{Y_k}}^2}\right)$$
$$X_{t,u-1}^2+\left(\frac{\left(\phi_{11k}^{(2)}\right)^2}{\sigma_{e_{X_k}}^2}+\frac{\left(\phi_{21k}^{(2)}\right)^2}{\sigma_{e_{Y_k}}^2}\right)X_{t,u-2}^2$$
$$+\left(\frac{\left(\phi_{12k}^{(1)}\right)^2}{\sigma_{e_{X_k}}^2}+\frac{\left(\phi_{22k}^{(1)}\right)^2}{\sigma_{e_{Y_k}}^2}\right)Y_{t,u-1}^2+\left(\frac{\left(\phi_{12k}^{(2)}\right)^2}{\sigma_{e_{X_k}}^2}+\frac{\left(\phi_{22k}^{(2)}\right)^2}{\sigma_{e_{Y_k}}^2}\right)Y_{t,u-2}^2$$
$$-2\left(\left(\frac{\phi_{11k}^{(1)}X_{t,u-1}}{\sigma_{e_{X_k}}^2}+\frac{\phi_{12k}^{(1)}Y_{t,u-1}}{\sigma_{e_{Y_k}}^2}+\frac{\phi_{11k}^{(2)}X_{t,u-2}}{\sigma_{e_{X_k}}^2}+\frac{\phi_{12k}^{(2)}Y_{t,u-2}}{\sigma_{e_{Y_k}}^2}\right)X_{t,u}\right.$$
$$\left.+\left(\frac{\phi_{21k}^{(1)}X_{t,u-1}}{\sigma_{e_{Y_k}}^2}+\frac{\phi_{22k}^{(1)}Y_{t,u-1}}{\sigma_{e_{Y_k}}^2}+\frac{\phi_{21k}^{(2)}X_{t,u-2}}{\sigma_{e_{Y_k}}^2}+\frac{\phi_{22k}^{(2)}Y_{t,u-2}}{\sigma_{e_{Y_k}}^2}\right)Y_{t,u}\right)\Bigg]$$

$$+\frac{1+\rho_k^2}{1-\rho_k^2}\,\frac{1}{\sigma_{e_{X_k}}\sigma_{e_{Y_k}}}\sum_{u=3}^{n}\Big[\big(Y_{t,u}-\phi_{21k}^{(1)}X_{t,u-1}+\phi_{21k}^{(2)}X_{t,u-2}$$
$$-\phi_{22k}^{(1)}Y_{t,u-1}-\phi_{22k}^{(2)}Y_{t,u-2}\big)X_{t,u}-\big(\phi_{11k}^{(1)}X_{t,u-1}+\phi_{11k}^{(2)}X_{t,u-2}$$
$$-+\phi_{11k}^{(1)}X_{t,u-1}\ \phi_{12k}^{(1)}Y_{t,u-1}+\phi_{12k}^{(2)}Y_{t,u-2}\big)Y_{t,u}$$
$$+\phi_{11k}^{(1)}\big(\ \phi_{21k}^{(1)}X_{t,u-1}+\phi_{21k}^{(2)}\ X_{t,u-2}+\phi_{22k}^{(1)}Y_{t,u-1}+\phi_{22k}^{(2)}\ Y_{t,u-2}\big)X_{t,u-1}$$

$$+\phi_{11k}^{(2)}\left(\phi_{21k}^{(1)}X_{t,u-1}+\phi_{21k}^{(2)}X_{t,u-2}+\phi_{22k}^{(1)}Y_{t,u-1}+\phi_{22k}^{(2)}Y_{t,u-2}\right)X_{t,u-2}$$

$$+\phi_{12k}^{(1)}\left(\phi_{21k}^{(1)}X_{t,u-1}+\phi_{21k}^{(2)}X_{t,u-2}+\phi_{22k}^{(1)}Y_{t,u-1}+\phi_{22k}^{(2)}Y_{t,u-2}\right)Y_{t,u-1}$$

$$\left.+\phi_{12k}^{(2)}\left(\phi_{21k}^{(1)}X_{t,u-1}+\phi_{21k}^{(2)}X_{t,u-2}+\phi_{22k}^{(1)}Y_{t,u-1}+\phi_{22k}^{(2)}Y_{t,u-2}\right)Y_{t,u-2}\Big]\right\}$$

$$= 0 \qquad (7.11)$$

Using the initialization of parameters discussed in section 5, the initial estimates of parameters $P(\omega_k), \sigma^2_{e_{X_k}}, \sigma^2_{e_{Y_k}}$, $\phi^{(j)}_{11k}, \phi^{(j)}_{12k}, \phi^{(j)}_{21k}, \phi^{(j)}_{22k}$ and ρ_k are obtained for three regions. The computed initial estimates of the model parameters are presented in Table 7.1.

Table 7.1
Initial Estimates of the Model Parameters

Parameter	Cluster 1 (Region I)	Cluster 2 (Region II)	Cluster 3 (Region III)
$P(\omega_k)$	0.3333	0.3333	0.3333
$\sigma^2_{e_X}$	13.5424	10.8860	13.6331
$\sigma^2_{e_Y}$	15.8794	20.9672	9.3483
ρ	0.1726	0.1826	0.4143
$\Phi^{(1)}{}_{11}$	-0.5576	-0.1566	0.2256
$\Phi^{(1)}{}_{12}$	-0.0320	-0.0075	-0.544
$\Phi^{(1)}{}_{21}$	-0.1198	0.0089	0.1110
$\Phi^{(1)}{}_{22}$	-0.2174	0.0329	-0.0054
$\Phi^{(2)}{}_{11}$	0.3700	0.3327	0.0266
$\Phi^{(2)}{}_{12}$	-0.1738	0.2782	-0.5018
$\Phi^{(2)}{}_{21}$	0.3572	0.0522	0.0382
$\Phi^{(2)}{}_{22}$	0.2559	0.075	0.3254

Using the initial estimates of the parameters and EM algorithm, the refined estimates of parameters for each class are obtained and presented in Table 7.2.

Table 7.2
Final Estimates of the Model Parameters

Parameter	Cluster 1 (Region I)	Cluster 1 (Region II)	Cluster 1 (Region III)
$P(\omega_k)$	0.342	0.412	0.246
$\sigma^2_{e_X}$	12.243	9.584	14.6996
$\sigma^2_{e_Y}$	8.9341	22.467	10.040
ρ	-0.156	-0.05	0.024
$\Phi^{(1)}{}_{11}$	-0.3891	0.036	-0.4053
$\Phi^{(1)}{}_{12}$	0.0493	0.016	0.0531
$\Phi^{(1)}{}_{21}$	-0.0294	0.142	-0.0295
$\Phi^{(1)}{}_{22}$	-0.0925	0.148	-0.0925
$\Phi^{(2)}{}_{11}$	-0.2965	-0.0009	-0.321
$\Phi^{(2)}{}_{12}$	0.4321	0.009	0.4284
$\Phi^{(2)}{}_{21}$	-0.1815	-0.140	-0.1815
$\Phi^{(2)}{}_{22}$	0.1395	0.019	0.1395

With these final estimates, the models characterizing the three groups of time series is estimated as

Region I

$$\begin{pmatrix}X_t\\Y_t\end{pmatrix}=\begin{pmatrix}-0.3891 & 0.0493 & -0.2965 & 0.4321\\-0.0294 & -0.0925 & -0.1815 & 0.1395\end{pmatrix}\begin{pmatrix}X_{t-1}\\Y_{t-1}\\X_{t-2}\\Y_{t-2}\end{pmatrix}+\begin{pmatrix}e_{X_t}\\e_{Y_t}\end{pmatrix}$$

Region II

$$\begin{pmatrix}X_t\\Y_t\end{pmatrix}=\begin{pmatrix}0.036 & 0.016 & -0.0009 & 0.009\\0.142 & 0.148 & -0.0140 & 0.019\end{pmatrix}\begin{pmatrix}X_{t-1}\\Y_{t-1}\\X_{t-2}\\Y_{t-2}\end{pmatrix}+\begin{pmatrix}e_{X_t}\\e_{Y_t}\end{pmatrix}$$

Region III

$$\begin{pmatrix}X_t\\Y_t\end{pmatrix}=\begin{pmatrix}-0.4053 & 0.0531 & -0.321 & 0.4284\\-0.0295 & -0.0925 & -0.1815 & 0.1395\end{pmatrix}\begin{pmatrix}X_{t-1}\\Y_{t-1}\\X_{t-2}\\Y_{t-2}\end{pmatrix}+\begin{pmatrix}e_{X_t}\\e_{Y_t}\end{pmatrix}$$

Here, X_t is Temperature in a region at time t and Y_t is Humidity level at time t.

Therefore the model that characterizes the whole data set is a three component mixture of bivariate autoregressive process of order 2 (BAR(2)) with component weights as $P(\omega_1)$= 0.342, $P(\omega_2)$= 0.412 and $P(\omega_3)$ = 0.246 respectively for regions I, II and III.

For evaluating the performance of the proposed algorithm, true positive rate (TPR), false positive rater (FPR), false discovery rate (FDR) and F- measure are used. For the proposed supervised learning algorithm with bivariate autoregressive process of order 2, the performance measures for each class are computed and presented in Table 7.3.

Table 7.3
Performance Measures of the Bivariate AR(2)Classifier

	True Positive Rate (TPR)	False Positive Rate (FPR)	False Discovery Rate(FDR)	F Measure
Cluster 1	0.9667	0.0143	0.0333	0.9667
Cluster 2	0.9800	0.0400	0.0392	0.9703
Cluster 3	0.9500	0.0250	0.0952	0.9268

To compare the efficiency of the developed bivariate AR(2) classifier with the earlier AR(2) classifiers, the true positive rate (TPR), false positive rate (FPR), false discover rate (FDR) and F-Measure are computed and presented in Table 7.4 and Table 7.5.

Table 7.4
Performance Measures of the Mixture of Univariate AR(2) Classifier for Temperature

	True Positive Rate (TPR)	False Positive Rate (FPR)	False Discovery rate (FDR)	F Measure
Cluster 1	0.9333	0.0571	0.1250	0.9032
Cluster 2	0.9400	0.0400	0.0408	0.9495
Cluster 3	0.8947	0.0370	0.1500	0.8718

Table 7.5
Performance Measures of the Mixture of Univariate AR(2) Classifier for Humidity

	True Positive Rate (TPR)	False Positive Rate (FPR)	False Discovery rate (FDR)	F Meaure
Cluster 1	0.9355	0.0435	0.0968	0.9355
Cluster 2	0.9592	0.0196	0.0208	0.9691
Cluster 3	0.9000	0.0250	0.1000	0.9000

From Table 7.3, Table 7.4 and Table 7.5, it is observed that the F value for all categories using the proposed classifier are more compared to that of the classifier with univariate AR(2) models for both the variables. This indicates that the proposed classifier with bivariate AR(2) model is much better in classifying time series data than the other two classifiers.

For evaluating the developed algorithm discussed in section 6, the test data consisting of 100 time series is considered. The developed unsupervised learning algorithm with bivariate AR(2) model identified 29 as region I, 49 as region II and 19 region III. To compare the efficiency of developed algorithm with existing univariate unsupervised learning algorithms with AR(2) model for both the variables temperature and humidity, the same test data have been considered and the misclassification rates are computed. Table 7.6 presents the misclassification rates of BAR(2) classifier, AR(2) classifiers for Temperature and Humidity separately.

Table 7.6
Performance Evaluation of Misclassification Rate

Classifier with	Misclassification Rate
BAR(2)	3%
AR(2) of temperature	8%
AR(2) of humidity	6%

From Table 7.6, it is observed that the misclassification rate for bivariate AR(2) classifier is less compared to misclassification rate of AR(2) classifiers of the individual variables Temperature and Humidity. Therefore the developed unsupervised learning algorithm with bivariate autoregressive process of order 2 outperforms the existing unsupervised learning algorithm with autoregressive processes of order 2. This algorithm is useful in other domains like medical, biological, financial applications etc.

8. Conclusion

In this paper we investigated the problem of clustering Bivariate Time Series data. Traditional static data clustering algorithms are inadequate in clustering time series data. Also the algorithms of univariate data sets assumes that there exists no correlation among the attributes under consideration. But in real world the inherent correlations influence the developed models. Considering this, we developed an algorithm for clustering bivariate time series using autoregessive model of order (p). The initial clusters are found using K-means algorithm and the model parameters are estimated using the EM algorithm. The clustering algorithm is developed by utilizing component maximum likelihood.

References

[1] Bandyopadhyay,S. , R. Baragona, U. Maulik (2010). Fuzzy clustering of univariate and multivariate time series by genetic multiobjective optimization. Computational Optimization methods in Statistics, Econometrics and Finance, WPS(028).

[2] Christoph Pamminger and Sylvia FrAuhwirth-Schnatter (2010). Model-based Clustering of Categorical Time Series. Bayesian Analysis, 5(2), 345-368.

[3] Douglas C. Montigomry, Lunwood A.Johnson and John S.Gardiner (1990). Forecasting and Time Series Analysis, Mc.Graw Hill.

[4] Eamonn Keogh, Jessica Lin, Ada Waichee Fu, and Helga Van Herle (2006). Finding Unusual Medical Time-Series Subsequences: Algorithms and Applications. IEEE Transactions on Information Technology in Biomedicine, 10(3), 429-439.

[5] Elizabeth E. Holmes, Eric J. Ward, Kellie Wills (2012). MARSS: Multivariate Autoregressive State-space Models for Analyzing Time-series Data. The R Journal, 4(1), 11-19.

[6] Emrah Bulut, Okan Duru, and Shigeru Yoshida (2012). A Fuzzy Time Series Forecasting Model for Multi-Variate Forecasting Analysis with Fuzzy C-Means Clustering. World Academy of Science, Engineering and Technology International Journal of Computer, Electrical, Automation, Control and Information Engineering, 6(3), 390-396.

[7] Jose A. Vilar and Juan M. Vilar (2013). Time series clustering based on nonparametric multidimensional forecast densities. Electronic Journal of Statistics, 7, 1019–1046.

[8] Kavitha,V., M.Punithavalli (2010). Clustering Time Series Data Stream – A Literature Survey. International Journal of Computer Science and Information Security, 8(1), 281-294.

[9] Luciana Alvim S. Romani, Ana Maria H. de Avila, Daniel Y. T. Chino, Jurandir Zullo, Jr., Richard Chbeir, Caetano Traina, Jr., and Agma J. M. Traina (2014). A New Time Series Mining Approach Applied to Multitemporal Remote Sensing Imagery. IEEE Transactions on Geoscience and Remote Sensing, 51(1), 140-150.

[10] Luis E. Nieto-Barajas and Alberto Contreras-Cristan (2014). A Bayesian Nonparametric Approach for Time Series Clustering. Bayesian Analysis, 9(1), 147 -170.