Yang Zong-chang*

J. Electrical Systems 9-2 (2013): 223-242

Regular paper

Journal of Electrical Systems

Eigenload Model Based Yearly Electric Load Demand Evaluation and Forecasting for State Grid Corporation of China

This study addresses the yearly load movement evaluation and forecasting based on the Principal Component Analysis (PCA). A called Eigenload model for describing the annual load movement by employing PCA is introduced. Principal orthogonal eigenvectors of covariance matrix of the load data called "Eigenloads" are used to build subspace for the load movement evaluation. Each load movement is projected onto subspace of the eigenspace and described by a linear combination of the "Eigenloads". Incorporated with the polynomial curve fitting algorithm to estimate its subsequent representation weights with respect to the "Eigenloads" generated by previous load movements, the Eigenload model is extended to forecasting subsequent movement of the load demand. The proposed method is applied to experiments of annual load demand evaluation for State Grid Corporation of China (SG) and its five branches in 2004-2006, and forecasting for their annual load demands in 2007. Experimental results agreeing well with their actual load demands show workability of the proposed model. Result analysis indicates that the Eigenload model is outperforming the classical autoregressive (AR) model on the forecast tasks.

Keywords: Principal Component Analysis; Load Movement; Eigenload Model; Evaluation and Forecasting.

1. Introduction

The load demand forecasting has a significant effect on both economy of operations and control of power systems [1-3]. It is shown that [2, 4] the reduction in the average forecast error can save cost. We may regard the load movement as a random non-stationary process comprising a wide variety of individual components. A number of factors may influence the load movement, which may be categorized as [1-3, 5-7]: time, season and climatic factors, economic factors and random effects.

A wide variety of methodologies and techniques [8-16] have been used for the load forecasting with varied degrees of success, which can be grouped as two main categories [3]: one is methods and models belonging to a more classical approach by applying concepts from time-series and regression analysis such as the 3 classical models called the autoregressive (AR), moving average (MA) and ARMA, the other is methods following the emerging technology of artificial and computational intelligence. The authors of [2] in 2002,ever prospected that new hopes in this direction of load forecast research would be provided by the new methods based on fuzzy logic, expert systems, genetic algorithms and artificial neural networks (ANN). Until now, a great deal of attention has been given to ANN and it has been reported fairly good performances for its nonlinear learning capability. Besides its computational burden, two major risks called under-fitting and over-fitting regarding the possibilities of less or excessive training data approximation still exist in employing ANN models, which increase the errors of out-of-sample forecasting. The authors of [3] in 2009 remarked that to select a proper model is one of the first decisions to be made. However, it is a problem &situation dependent consideration. The authors of [3] then indicated that no general recommendations could be recommended. The load forecast may be characterized by so various factors [6, 7] including not only the power system itself but also social-economic and climatic factors. Existing one method or model can meet all

* Corresponding author: Dr. Yang Zong chang, School of Information and Electronical Engineering, Hunan University of Science and Technology, P.R. China, email : yzc233@163.com

requirements seems to be inadequate. Hybrid methods by combining two or more of these techniques seems to be one direction, and methods following the iterative reweighted least-squares and adaptive load forecasting are also considerably less emphasized [2]. Incorporated with the polynomial curve fitting algorithm, in this study, a PCA-based called Eigenload model is presented for the yearly electric load movement evaluation and forecast.

The PCA (principal component analysis) technology [17] thinks of an object as being described by a weighted combination of some principal ("basis") components, while try to keep information losses as little as possible in the mean-square sense. In the PCA, data structure is decomposed into orthogonal components that the most effective structure may be discovered in low dimensions. A regression method with PCA of the daily demand profiles is intruded in [18-19], where PCA is employed to capture the intra-day variation in the electric demand [18-19]. In [25], we introduce an Eigen-temperature model for describing the annual air temperature movement by employing PCA. It is shown in [25] that the annual temperature movement can be well evaluated (reconstructed) by a linear combination of its few principal orthogonal eigenvectors called "Eigen-temperatures". Since the air temperature movement is the most important weather influence on the load movement, in this study, we try to employ the PCA technique in creating eigenvectors and using them for the yearly electric load movement evaluation and forecast incorporated with the polynomial curve fitting algorithm, and apply it to State Grid Corporation of China (SG) and its five branches.

The paper is organized as follows. In Section 2, it describes the principal component analysis (PCA). In Section 3, an Eigenload model for the load movement evaluation and forecasting is presented. Principal orthogonal eigenvectors of covariance matrix of the load data are selected to generate the subspace for evaluation and forecasting. The principal orthogonal eigenvectors are called "Eigenloads", because they are eigenvectors in the eigenspace and each load movement is described in this basis. In Section4, the proposed method is applied to experiments including load demand evaluation for SG and its five branches in 2004-2006, and forecasting of their annual load demands in 2007. Section 5 depicts result analysis and discussion. Section 6 lists concluding remarks and future study.

2. Principal Component Analysis (PCA)

The KLT (Karhunen-Loève transform) or PCA [17] is an important orthogonal linear transformation, in which the data is transformed to a new coordinate system. It is one of the most popular methods for feature generation and dimensionality reduction in data analysis and pattern classification [17, 20-23]. By means of data compression basics, the PCA technology [17] can reduce data-dimension. The reduction in dimensions can remove the less useful information and the most effective structure of data structure may be discovered in low dimensions [17].

Given a set of l (l>1) measurement vectors: $\boldsymbol{X}_{N\times l} = (\boldsymbol{x}_1, \boldsymbol{x}_2, .., \boldsymbol{x}_l)$, its PCA is presented as follows [25]:

$$\boldsymbol{X} = \boldsymbol{X}_{N\times l} = (\boldsymbol{x}_1, \boldsymbol{x}_2, .., \boldsymbol{x}_l) \text{ And } \boldsymbol{\Psi} = \frac{1}{l}\sum_{i=1}^{l} \boldsymbol{x}_i \quad \textbf{(1)}$$

Where $\boldsymbol{\Psi}_{N\times 1} = (\psi_1, \psi_2, ..., \psi_N)^T$ is the average vector of X.

Compute Y by the linear transformation using an orthogonal basis: $\boldsymbol{A}$ $(\boldsymbol{A}^{-1} = \boldsymbol{A}^T)$,

$$\boldsymbol{Y}_{N\times l} = \boldsymbol{A}^T(\boldsymbol{X} - \boldsymbol{\Psi}) \quad (\boldsymbol{Y}_{N\times l} = (\boldsymbol{y}_1, \boldsymbol{y}_2, .., \boldsymbol{y}_l)) \quad \textbf{(2)}$$

Where $\boldsymbol{y}_i$, $\boldsymbol{y}_j$ are columns (column vectors) of $\boldsymbol{Y}$.

Define correlation matrix $\boldsymbol{R}_Y$ and covariance matrix $\boldsymbol{C}_X$,

$$\boldsymbol{R}_Y = E\left[\boldsymbol{YY}^T\right] = E\left[\boldsymbol{A}^T(\boldsymbol{X} - \boldsymbol{\Psi})(\boldsymbol{X} - \boldsymbol{\Psi})^T \boldsymbol{A}\right] = \boldsymbol{A}^T \boldsymbol{C}_X \boldsymbol{A} \tag{3}$$

If $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_N)$ is selected so that its columns (vectors) $\boldsymbol{a}_i$ are the orthogonal eigenvectors of $\boldsymbol{C}_X$, we have,

$$\boldsymbol{R}_y = E\left[\boldsymbol{YY}^T\right] = \boldsymbol{A}^T \boldsymbol{C}_X \boldsymbol{A} = \boldsymbol{\Lambda} \tag{4}$$

Where $\boldsymbol{\Lambda}$ is diagonal with elements the respective eigenvalues λ_i of $\boldsymbol{C}_X$.

Due to the orthogonality of $\boldsymbol{A}$, we get,

$$\boldsymbol{Y} = \boldsymbol{A}^T(\boldsymbol{X} - \boldsymbol{\Psi}) \rightarrow \boldsymbol{X} = (\boldsymbol{A}^T)^{-1}\boldsymbol{Y} + \boldsymbol{\Psi} = \boldsymbol{AY} + \boldsymbol{\Psi} \tag{5}$$

That is,

$$\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, .., \boldsymbol{x}_l) = \boldsymbol{AY} + \boldsymbol{\Psi} \quad \rightarrow \boldsymbol{x}_j = \sum_{i=1}^{N} \boldsymbol{a}_i y_{ij} + \boldsymbol{\Psi}(j = 1, ..., l) \tag{6}$$

Where y_{ij} are matrix elements of $\boldsymbol{Y}$ in *No.i* row and *No.j* column

When choose its principal m eigenvectors $\boldsymbol{a}_i(i = 1, 2, .., m)$ of $\boldsymbol{C}_X$ to approximate or reconstruct $\boldsymbol{X}$, defined as $\hat{\boldsymbol{X}}$,

$$\hat{\boldsymbol{X}} = (\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, ..., \hat{\boldsymbol{x}}_l) = (\sum_{i=1}^{m} \boldsymbol{a}_i y_{i1}, \sum_{i=1}^{m} \boldsymbol{a}_i y_{i2}, ..., \sum_{i=1}^{m} \boldsymbol{a}_i y_{il}) + \boldsymbol{\Psi} \tag{7}$$

The K-L transform minimizes the square error by defining MSE (the mean square error) for the matrix $\boldsymbol{X}$ and its approximation $\hat{\boldsymbol{X}}$ as follows [17]:

$$E\left[\left\|\boldsymbol{X} - \hat{\boldsymbol{X}}\right\|^2\right] = \sum_{i=m+1}^{N} \boldsymbol{a}_i^T \boldsymbol{C}_X \boldsymbol{a}_i = \sum_{i=m+1}^{N} \lambda_i \tag{8}$$

That is, when m principal eigenvectors (that correspond to the m largest eigenvalues of the covariance matrix $\boldsymbol{C}_X$) are chosen, the error (*MSE*) in *Eq.*(8) is minimized, which is the sum of the N-m smallest eigenvalues 8.

3. Eigenload Model for Annual Load Movement Analysis

In this section, a called "Eigenload" model based on PCA is proposed for the annual load movement evaluation and forecasting.

Eigenload Evaluation Model for the Load Movement

Assume the load movement of N period (normally, set N=12 for the yearly load movement that it has 12 months in each year), given a set of l vectors of the annual load moments: $\boldsymbol{X} = \boldsymbol{X}_{N\times l} = (\boldsymbol{x}_1, \boldsymbol{x}_2, .., \boldsymbol{x}_l)$, where each vector (time series) $\boldsymbol{x}_k$ ($1 \leq k \leq l$) of N load values denotes one yearly load movement.

From *Eq*. (2) that $\boldsymbol{Y}_{N\times l} = \boldsymbol{A}^T(\boldsymbol{X} - \boldsymbol{\Psi}) = \begin{pmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \\ ... \\ \boldsymbol{a}_N^T \end{pmatrix} (\boldsymbol{X} - \boldsymbol{\Psi}) = \begin{pmatrix} \boldsymbol{a}_1^T(\boldsymbol{X} - \boldsymbol{\Psi}) \\ \boldsymbol{a}_2^T(\boldsymbol{X} - \boldsymbol{\Psi}) \\ ... \\ \boldsymbol{a}_N^T(\boldsymbol{X} - \boldsymbol{\Psi}) \end{pmatrix}$, we have

$y_{ik} = \boldsymbol{a}_i^T(\boldsymbol{x}_k - \boldsymbol{\Psi}) = (\boldsymbol{x}_k - \boldsymbol{\Psi})^T \boldsymbol{a}_i$, where y_{ik} is scalar quantity.

From *Eq*.(7) that $\hat{\boldsymbol{X}} = (\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, ..., \hat{\boldsymbol{x}}_l) = (\sum_{i=1}^{m} \boldsymbol{a}_i y_{i1},\ \sum_{i=1}^{m} \boldsymbol{a}_i y_{i2},\ ...,\ \sum_{i=1}^{m} \boldsymbol{a}_i y_{il}) + \boldsymbol{\Psi}$, we have

$$\hat{\boldsymbol{x}}_k = \sum_{i=1}^{m} \boldsymbol{a}_i y_{ik} + \boldsymbol{\Psi} \qquad \textbf{(12)}$$

It indicates that $\boldsymbol{x}_k$ can be evaluated (approximated or reconstructed) by a linear combination of the *m* principal orthogonal eigenvectors $(\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m)$. The basis vectors $(\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m)$ are now referred as the "**Eigenloads**", since they are eigenvectors and each one load movement $\boldsymbol{x}_k$ is described by a linear combination of the basis vectors (*Eq*.(12)), and represented in this basis by a weight vector of $\boldsymbol{\Omega}_k$,

$$\boldsymbol{\Omega}_k = \begin{pmatrix} y_{1k} \\ y_{2k} \\ \cdots \\ y_{mk} \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_1^T(\boldsymbol{x}_k - \boldsymbol{\Psi}) \\ \boldsymbol{a}_2^T(\boldsymbol{x}_k - \boldsymbol{\Psi}) \\ \cdots \\ \boldsymbol{a}_m^T(\boldsymbol{x}_k - \boldsymbol{\Psi}) \end{pmatrix} \qquad \textbf{(13)}$$

That is, project each load movement $\boldsymbol{x}_k$ onto the eigenspace, we can evaluate it by a linear combination (*Eq*.(12)) of the *m* principal orthogonal components $(\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m)$ called "**Eigenloads**" in the eigenspace. **Fig.1** illustrates the "Eigenloads" (*Eq*.(12)) used for evaluation of annual load demand (electricity demands, 10^8 kWh) of State Grid Corporation of China (SG) in 2004-2006.

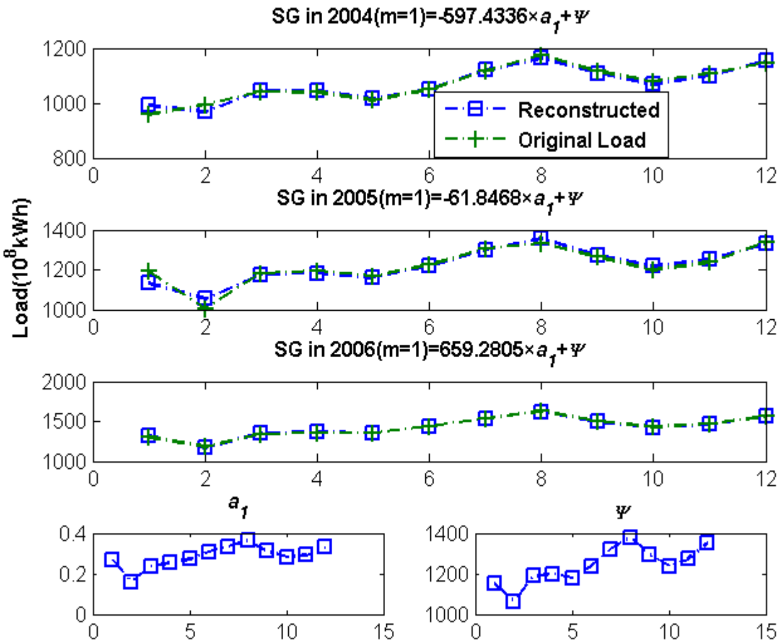

Fig.1:The Eigenload evaluation for annual load demands (electricity demands, 10^8 kWh) of SG in 2004-2006 (***m*=1**).

Eigenload Forecast Model for the Load Movement

Given a set of *l* vectors of the annual load moments: $\boldsymbol{X} = \boldsymbol{X}_{N \times l} = (\boldsymbol{x}_1, \boldsymbol{x}_2, .., \boldsymbol{x}_l)$, where each vector $\boldsymbol{x}_k$ ($1 \le k \le l$) of *N* load values is a time-series, and denotes one yearly load movement, we now focus on how to forecast its next annual load movement $\hat{\boldsymbol{x}}_{l+1}$ of *N* monthly (average) load values at the next N coming months. The forecast is of N-step-length (normally, N=12).

In ***Eq*.(12)**, it indicates that each load movement $\boldsymbol{x_k}$ of N values is described as a linear combination of the m principal orthogonal eigenvectors $(\boldsymbol{a_1}, \boldsymbol{a_2}, ..., \boldsymbol{a_m})$ called "**Eigenloads**" in the eigenspace, and represented in this basis by a weight vector of $\boldsymbol{\Omega}_k = (y_{1k}, y_{2k} ... y_{mk})^T$ (1≤k≤l) (***Eq*.(13)**).
For the given l load moments: $\boldsymbol{X} = \boldsymbol{X}_{N \times l} = (\boldsymbol{x}_1, \boldsymbol{x}_2, .., \boldsymbol{x}_l)$, there are l corresponding representation vectors $(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2 ... \boldsymbol{\Omega}_l)$, denote them as,

$$\boldsymbol{\Omega} = (\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2 ... \boldsymbol{\Omega}_l) = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1l} \\ y_{21} & y_{22} & \cdots & y_{2l} \\ \cdots & \cdots & \cdots & \cdots \\ y_{m1} & y_{m2} & \cdots & y_{ml} \end{pmatrix} = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \cdots \\ \boldsymbol{y}_m \end{pmatrix} \quad \textbf{(14)}$$

Where $y_{ik} = \boldsymbol{a}_i^T (\boldsymbol{x_k} - \boldsymbol{\Psi})$

In the matrix of **Ω**= (**Ω**$_1$, **Ω**$_2$…**Ω**$_l$), each row vector $\boldsymbol{y}_i$= (y_{i1}, y_{i2}… y_{il}) denotes respective presentation weights of the l load movements with respect to one "Eigenload" $\boldsymbol{a}_i$ (1≤ i ≤m).
From ***Eq*.(12)** that $\hat{\boldsymbol{x}}_{\boldsymbol{k}} = \sum_{i=1}^{m} \boldsymbol{a}_i y_{ik} + \boldsymbol{\Psi}$ (1≤k≤l), when use the average vector $\boldsymbol{\Psi}$ of $\boldsymbol{X}$ and the m "Eigenload" ($\boldsymbol{a}_1$,$\boldsymbol{a}_2$…$\boldsymbol{a}_m$) to forecast its subsequent annual load movement $\hat{\boldsymbol{x}}_{\boldsymbol{l+1}}$ with N monthly (average) load values in the next N coming months, obviously, we should get its representation weights **Ω**$_{l+1}$ with respect to the m "Eigenload" ($\boldsymbol{a}_1$, $\boldsymbol{a}_2$…$\boldsymbol{a}_m$). However, the subsequent annual load movement $\hat{\boldsymbol{x}}_{\boldsymbol{l+1}}$ is unavailable, we can NOT directly get its representation weight vector $\boldsymbol{\Omega}_{l+1} = (y_{1l+1}, y_{2l+1} ... y_{ml+1})^T$ by the formulation of $y_{il+1} = \boldsymbol{a}_i^T (\boldsymbol{x}_{\boldsymbol{l+1}} - \boldsymbol{\Psi})$ (*Eq*.(13), *Eq*.(14)).
We now propose one solution by using the polynomial curve fitting method [25] to estimate the representation weights **Ω**$_{l+1}$ for its subsequent annual load movement $\hat{\boldsymbol{x}}_{\boldsymbol{l+1}}$ with respect to the m "Eigenload" ($\boldsymbol{a}_1$, $\boldsymbol{a}_2$…$\boldsymbol{a}_m$), which are generated by its previous load movements. That is, for each row vector $\boldsymbol{y}_i$= (y_{i1}, y_{i2}… y_{il}) in the matrix of $\boldsymbol{\Omega}$ (*Eq*.(14))**,** we find a polynomial of degree n with coefficients $\boldsymbol{p}_i$= (p_{i0}, p_{i1}…p_{in}) that fits the data $\boldsymbol{k}$= (1, 2…l) to $\boldsymbol{y}_i$= (y_{i1}, y_{i2}… y_{il}) in a least squares sense,

$$\hat{y}_{ik} = \hat{y}_i(\boldsymbol{p}_i, k) = p_{in}k^n + p_{in-1}k^{n-1} + ... + p_{i1}k + p_{i0} \quad (1 \le k \le l,\ 1 \le i \le m) \quad \textbf{(15)}$$

Where, $\hat{y}_{ik}$ is value of the polynomial of degree n approximated at k for y_{ik}.
With the found coefficients $\boldsymbol{p}_i$= (p_{i0}, p_{i1}…p_{in}) of a polynomial of degree n for fitting each row vector $\boldsymbol{y}_i$= (y_{i1}, y_{i2}… y_{il}) in the matrix of **Ω**, from ***Eq*.(15)**, we can estimate the representation weight vector $\boldsymbol{\Omega}_{l+1}$ for its subsequent annual load movement $\hat{\boldsymbol{x}}_{\boldsymbol{l+1}}$ with respect to the m "Eigenload" ($\boldsymbol{a}_1$, $\boldsymbol{a}_2$…$\boldsymbol{a}_m$) by,

$$\hat{\boldsymbol{\Omega}}_{l+1} = (\hat{y}_{1l+1}, \hat{y}_{2l+1} ... \hat{y}_{ml+1})^T \quad \textbf{(16)}$$

Where, $\hat{y}_{il+1} = \hat{y}_i(\boldsymbol{p}_i, l+1) = p_{in}(l+1)^n + p_{in-1}(l+1)^{n-1} + ... + p_{i1}(l+1) + p_{i0}$ (1≤i≤m)

From ***Eq*.(12)**, with the estimated weight vector $\hat{\boldsymbol{\Omega}}_{l+1}$ (***Eq*.(16))** for its subsequent annual load movement $\hat{\boldsymbol{x}}_{\boldsymbol{l+1}}$ with respect to the m "Eigenload" ($\boldsymbol{a}_1$,$\boldsymbol{a}_2$…$\boldsymbol{a}_m$), we can forecast its subsequent load movement $\hat{\boldsymbol{x}}_{\boldsymbol{l+1}}$ with N(N=12)monthly load values at the next N coming months by,

$$\hat{\boldsymbol{x}}_{\boldsymbol{l+1}} = \sum_{i=1}^{m} \boldsymbol{a}_i \hat{y}_{il+1} + \boldsymbol{\Psi} \quad \textbf{(17)}$$

4. Experimental Result

Experimental data include 4 annual load demands (electricity sales, 10^8 kWh) of State Grid Corporation of China (SG) [24] and five branches - North Branch, East-North Branch, East Branch, Central Branch, and Western-North Branch from 2004 to 2007. It should be noted that the "load demand" is usually in terms of MW(10^3kW), and the "energy (electricity) demand" is often in terms of MWh (10^3kW.hour). Numerically, in one hour, the (average) energy (electricity) demand equals the load demand. For the convenient expression, in the following, we use the records of electricity demands (10^8 kWh) of the State Grid Corporation of China (SG) as the experimental data for the load demand evaluation and forecasting.

4.1. Yearly Load Demand Evaluation

The previous 3 annual electric load demands (electricity demands, 10^8 kWh) of State Grid Corporation of China (SG) and its five branches from 2004 to 2006 are used for yearly load demand evaluation. **Table 1** lists annual load demands of State Grid Corporation of China (SG) form 2004 to 2006.

Table 1: Annual Electric Load Demands (Electricity demands, 10^8 kWh) of SG from 2004 to 2006.

Month	2004	2005	2006
1	957.51	1191.11	1301.64
2	994.36	1006.19	1188.16
3	1042.25	1182.84	1343.37
4	1038.42	1194.36	1358.72
5	1011.06	1167.37	1353.88
6	1048.72	1223.72	1437.06
7	1116.60	1306.78	1536.26
8	1174.77	1333.89	1626.67
9	1114.89	1262.52	1503.04
10	1080.25	1197.01	1432.05
11	1108.19	1237.58	1472.78
12	1146.78	1340.88	1564.19

For each of SG and its five branches, each one yearly load demand $\boldsymbol{x}_k$ (k=1, 2, 3) of N (N=12) monthly load values is a time series and denoted as one vector. $\boldsymbol{X} = \boldsymbol{X}_{N\times l} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$ (l=3) of three annual load demands are used to build the called "Eigenload model" (***Eq*.(12)**) for the load demand evaluation. Note that where, k=1, 2, 3 correspond to the years from 2004 to 2006.

Figs.2-4 display evaluation results with setting of m=1 for the 3 annual load demands of SG and its five braches from 2004 to 2006, respectively. **Figs.5-7** display evaluation results with setting of m=2 for the 3 annual load demands of SG and its five braches from 2004 to 2006, respectively.

In **Figs.2-7,** it is shown that for each of SG and its five branches, each one yearly load demand(electricity demands) $\boldsymbol{x}_k$ (k=1, 2, 3, correspond to the years from 2004 to 2006), is well evaluated (described) by a linear combination of its few m (m=1, 2) principal orthogonal eigenvectors $(\boldsymbol{a_1}, \boldsymbol{a_2}, ..., \boldsymbol{a_m})$ called "Eigenloads" and the average vector $\boldsymbol{\Psi}$ of $\boldsymbol{X}$ (***Eq.*(12)**).

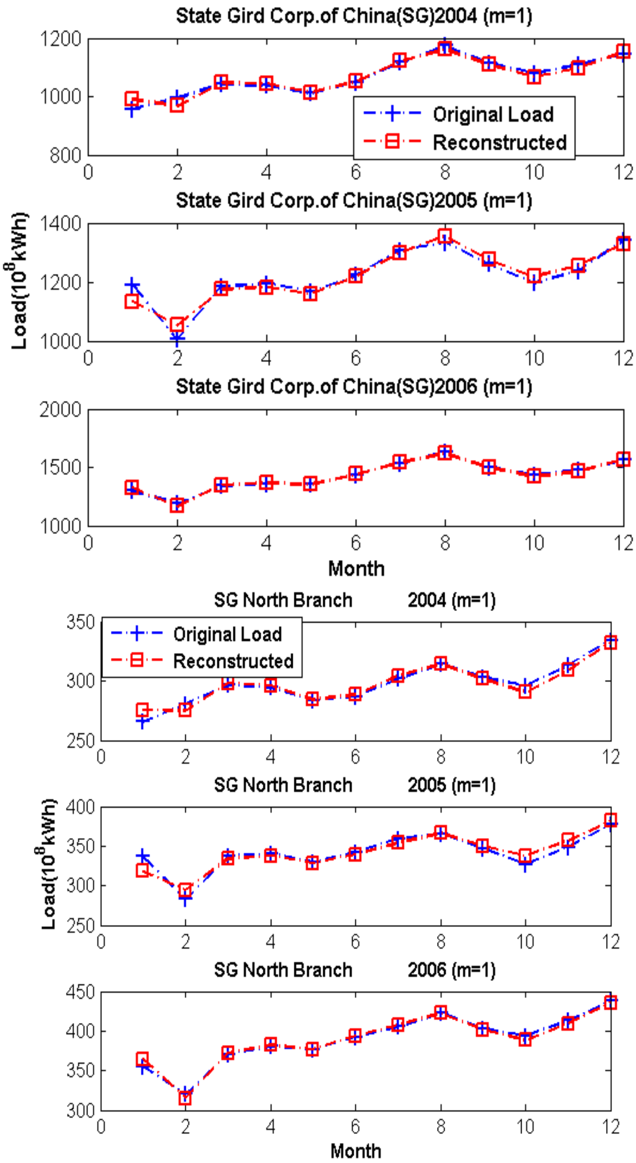

Fig.2: Plot of evaluation results for 3 annual electric load demands (electricity demands, 10^8 kWh) of State Grid Corporation of China (SG) and SG North Brach form 2004 to 2006 (**m=1**), respecitvely.

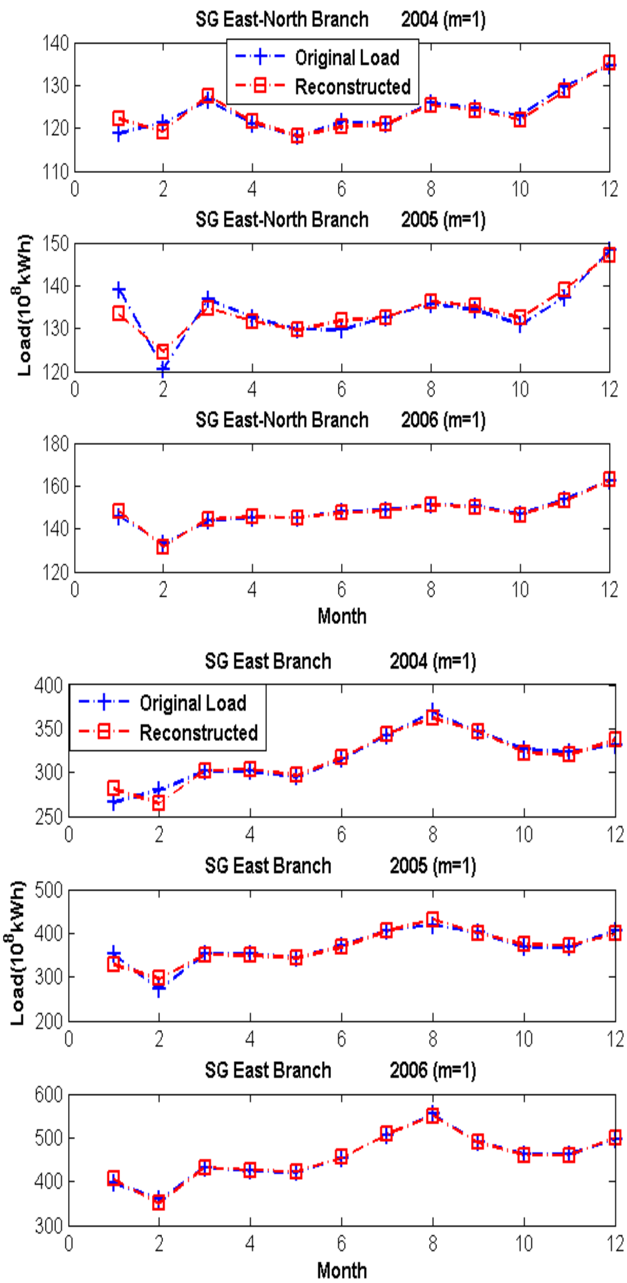

Fig.3: Plot of evaluation results for 3 annual electric load demands (electricity demands, 10^8 kWh) of SG East-North Brach and SG East Brach form 2004 to 2006 (**m=1**), respecitvely.

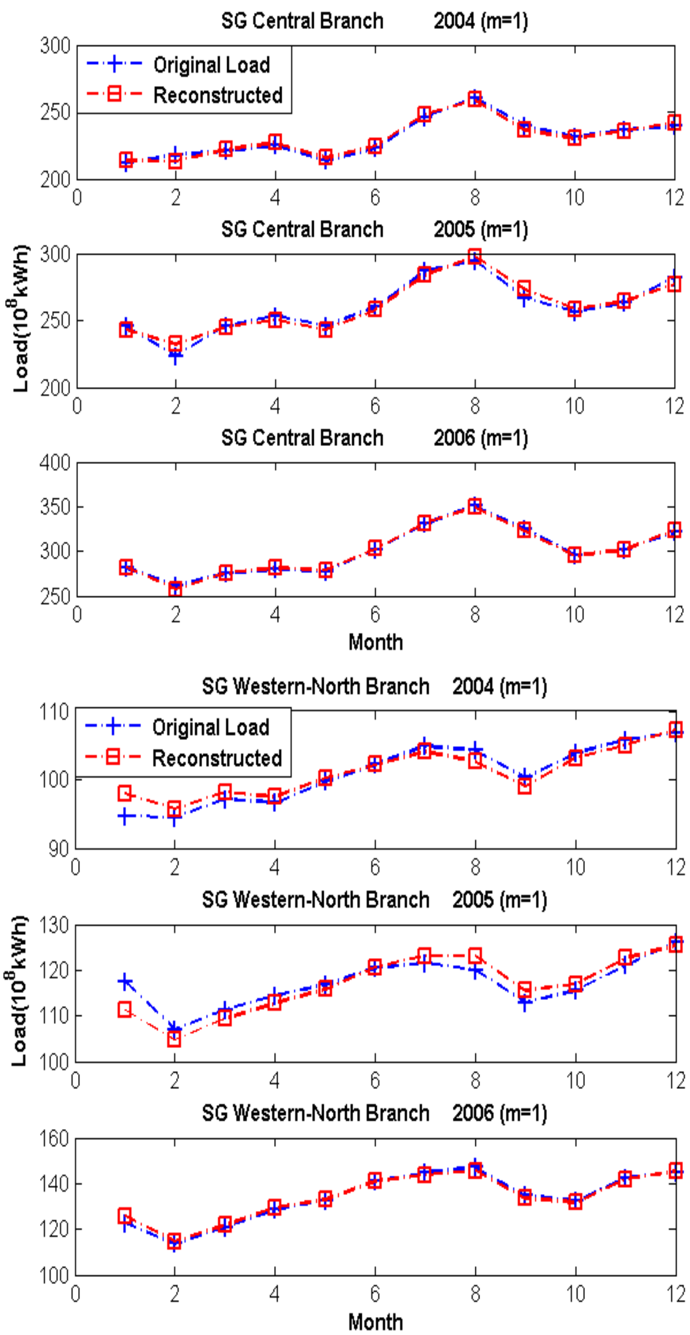

Fig.4: Plot of evaluation results for 3 annual electric load demands(electricity demands, 10^8 kWh) of SG Central Brach and SG Western-North Brach form 2004 to 2006 (**m=1**), respecitvely.

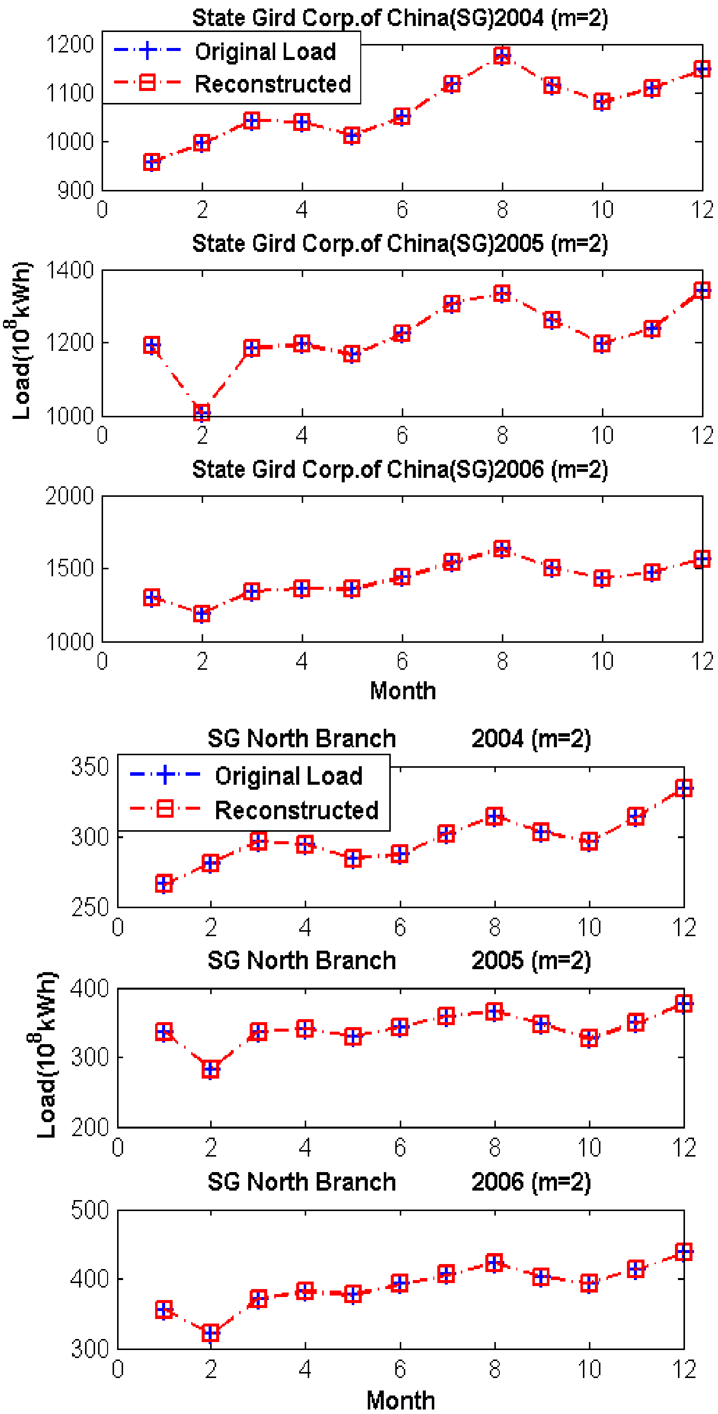

Fig.5: Plot of evaluation results for 3 annual electric load demands (electricity demands, 10^8 kWh) of State Grid Corporation of China (SG) and SG North Brach form 2004 to 2006 (**m=2**) , respecitvely.

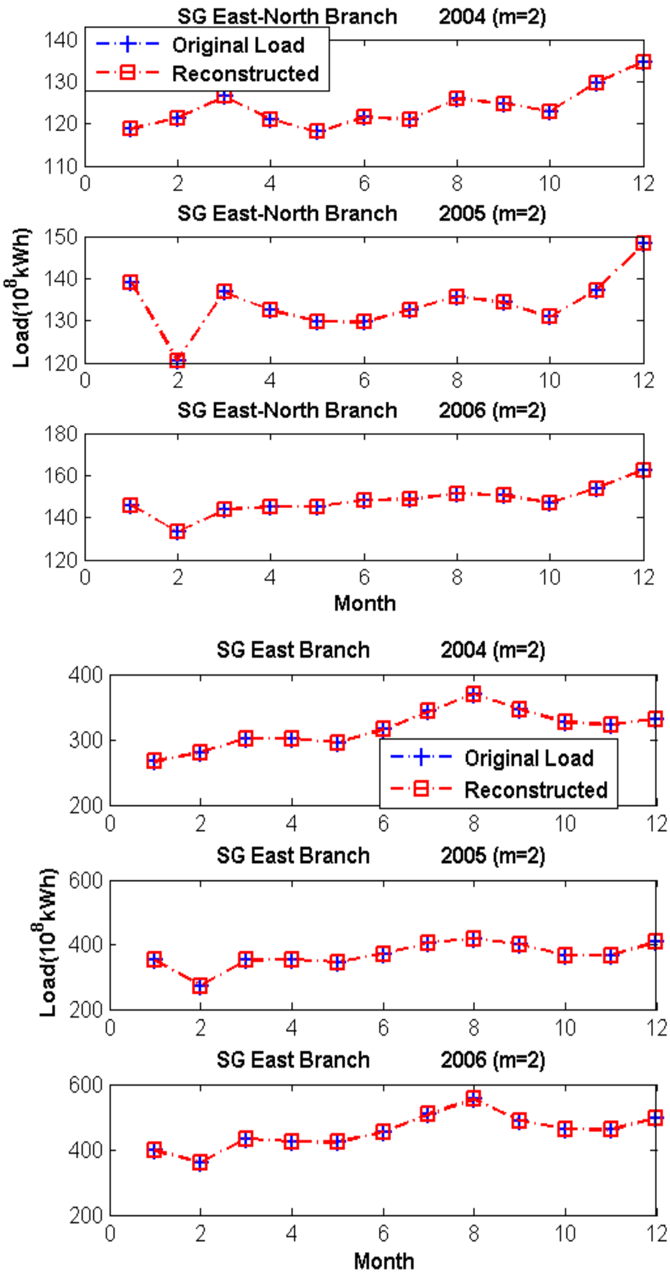

Fig.6: Plot of evaluation results for 3 annual electric load demands (electricity demands, 10^8 kWh) of SG East-North Brach and SG East Brach form 2004 to 2006 (**m=2**) , respecitvely.

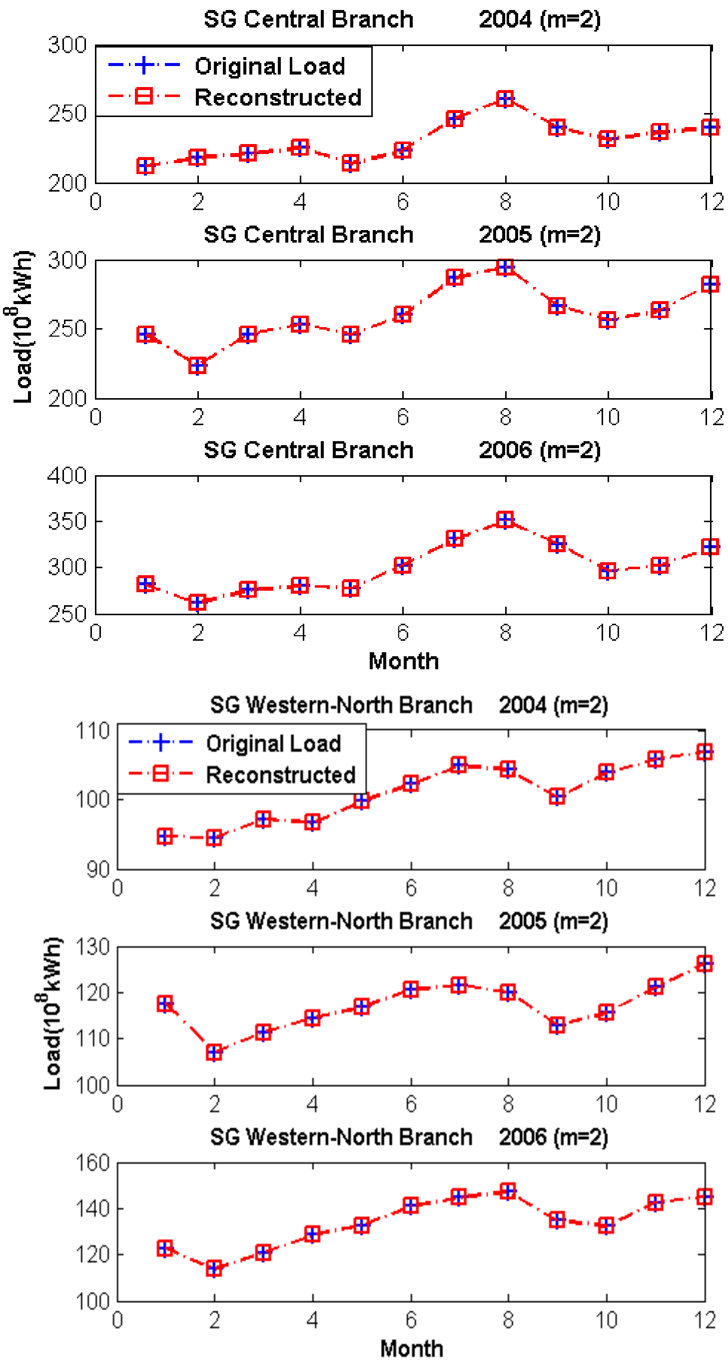

Fig.7: Plot of evaluation results for 3 annual electric load demands (electricity demands, 10^8 kWh) of SG Central Brach and SG Western-North Brach form 2004 to 2006 (**m=2**), respecitvely.

4.2. **Yearly Load Demand Forecast**

In the load demand(electricity demands) evaluation , it is shown that for each of State Grid Corporation of China(SG) and its five branches, each one yearly load demand is well described by a linear combination of its few m (m=1, 2) "Eigenloads"and the average vector.

For each of SG and its five branches, based on the "Eigenloads" generated by its previous 3 annual load demands $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$, we now employ the proposed "Eigenload" forecast model (***Eq.*(17)**) to forecast its next annual load demand $\hat{\boldsymbol{x}}_4$ with N

(N=12) monthly load values in 2007. Each forecast is of N-step-length (N=12). Where, k=1, 2, 3 correspond to the years from 2004 to 2006, and k=4 denotes the year 2007.

Under different settings with (m=1, 2; n=1)(***Eqs*.(14)-(17)**), forecast results of annual load demands for each of SG and its five branches in 2007 are displayed in **Figs.8-9** and **Figs.10-11**, respectively. **Table2** lists the forecast results of annual load demand for SG in 2007 with the different settings.

Table 2: Forecast results of annual electric load (electricity demands, 10^8 kWh) demand of State Grid Corporation of China (SG) in 2007.

Chinese State Grid Corporation(SG), 2007			
Month	**Actual Load**	**Forecast (m=1,n=1)**	**Forecast (m=2,n=1)**
1	1629.25	1485.72	1494.22
2	1303.43	1263.62	1256.70
3	1483.35	1489.41	1490.61
4	1583.90	1515.57	1517.47
5	1594.30	1519.26	1520.26
6	1668.29	1623.91	1624.84
7	1771.49	1738.43	1739.54
8	1866.88	1833.61	1830.34
9	1722.42	1683.37	1681.63
10	1647.14	1591.48	1588.24
11	1694.33	1639.98	1637.44
12	1777.11	1766.44	1768.03

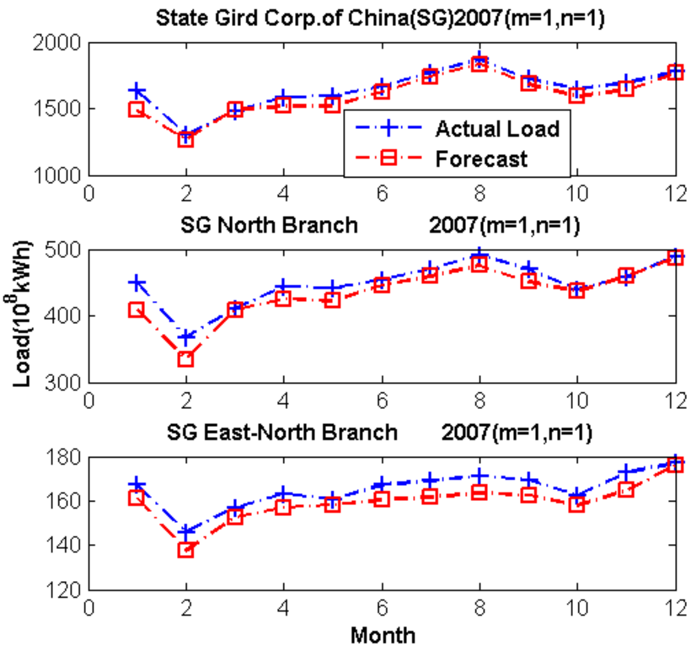

Fig.8. Forecast results for the electric load demands(electricity demands, 10^8 kWh) of SG and North Branch, East-North Branch in 2007 by the Eigenload model based on previous years of 2004-2006 with **(m=1,n=1)**.

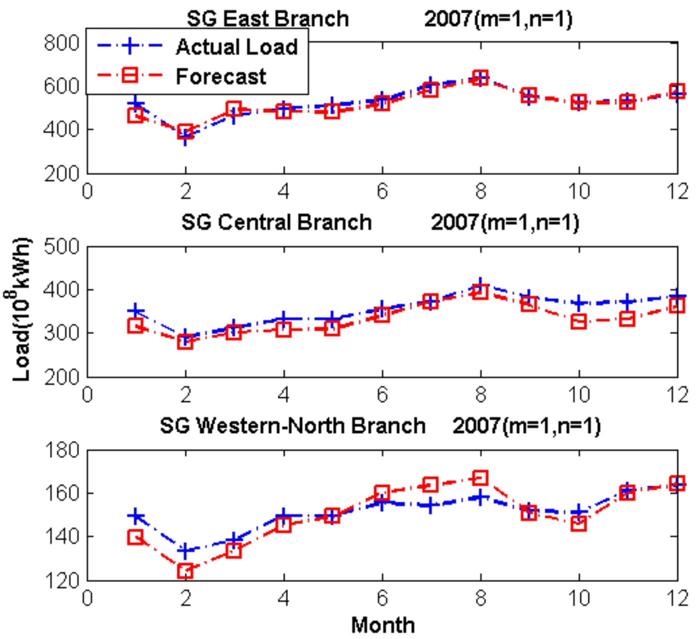

Fig.9. Forecast results for the electric load demands(electricity demands, 10^8 kWh) of SG East Branch, Central Branch and Western-North Branch in 2007 by the Eigenload model based on previous years of 2004-2006 with **(m=1,n=1)**

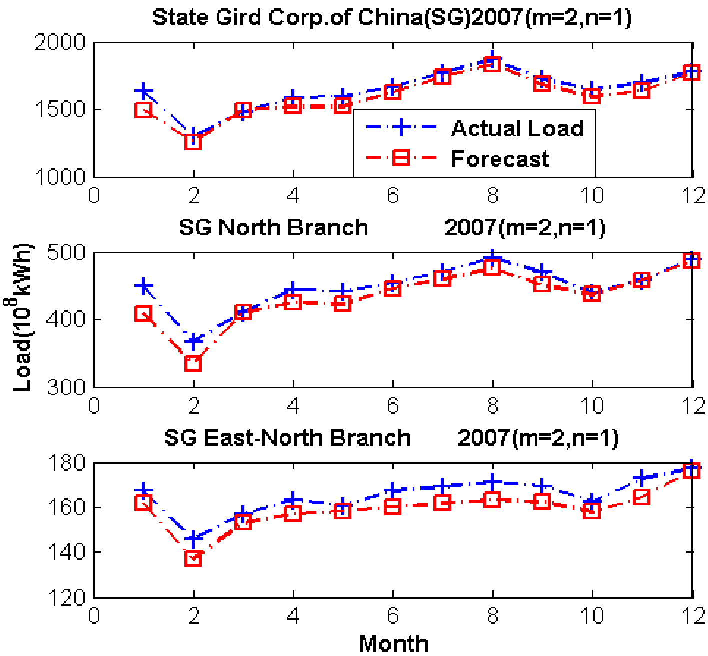

Fig.10. Forecast results for the electric load demands(electricity demands, 10^8 kWh) of SG and North Branch, East-North Branch in 2007 by the Eigenload model based on previous years of 2004-2006 with **(m=2,n=1)**.

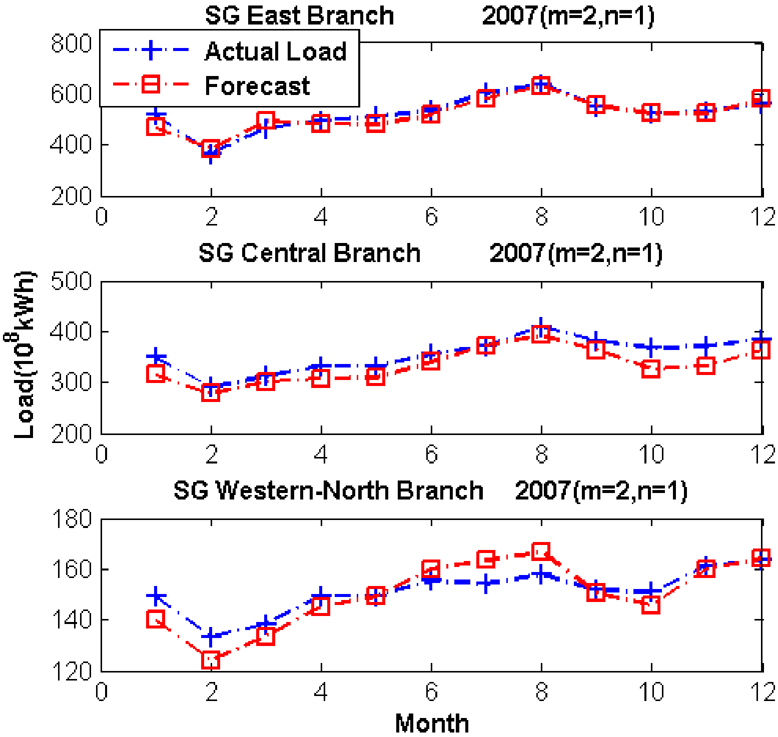

Fig.11 Forecast results for the electric load demands(electricity demands, 10^8 kWh) of SG East Branch, Central Branch and Western-North Branch in 2007 by the Eigenload model based on previous years of 2004-2006 with **(m=2,n=1)**.

5. Result Analysis and Discussion

The proposed "Eigenload" model (***Eq*.(12)**) has been applied to annual load demand evaluations for State Grid Corporation of China (SG) and five branches -North Branch, East-North Branch, East Branch, Central Branch, and Western-North Branch from 2004 to 2006. For each of SG and its five branches, each one yearly load demand $\boldsymbol{x}_k$ (k=1, 2, 3, correspond to the years from 2004 to 2006) with N (N=12) monthly load values, is a time series and denoted as one vector. $\boldsymbol{X}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)$ of three annual load demands are used to build the "Eigenload model" (***Eq*.(12)**) for the load demand evaluation. It is demonstrated in **Figs.2-11** that for each of SG and its five branches, each one yearly load demand $\boldsymbol{x}_k$ (k=1, 2, 3), is well evaluated (described) by a linear combination of its few m (m=1, 2) principal orthogonal eigenvectors $(\boldsymbol{a}_1,\boldsymbol{a}_2,...,\boldsymbol{a}_m)$ called "Eigenloads" and the average vector $\boldsymbol{\Psi}$ of $\boldsymbol{X}$.

For each of SG and its five branches, based on the "Eignloads" generated by its previous 3 annual load demands $\boldsymbol{X}$=($\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$), the "Eigenload" forecast model (***Eq*.(17)**) is employed to forecast its next annual load demand $\hat{\boldsymbol{x}}_4$ with N (N=12) monthly load values in 2007. Each forecast is of N-step-length (N=12). Where, k=4 denotes the year 2007.

For each of State Grid Corporation of China (SG) and its five Branches in 2007, **Table 3** lists correlation coefficients and errors in terms of *RMSE* (root mean square error, ***known as standard error σ)*** between actual load demand($\boldsymbol{x}_4$) and its forecast result ($\hat{\boldsymbol{x}}_4$), and ***ratio (percent) of*** *RMSE* to mean of actual load demand.

Table3: Correlation coefficients and RMSE between actual load demands and forecast results for SG and branches in 2007 by employing the Eigenload model

Forecast Result in 2007	*Correlation Coefficients*	*RMSE (σ)*	*Percent (RMSE/Mean)*
State Grid Corp., China (m=1,n=2)	97.39%	*56.265*	3.420%
North Branch (m=1,n=2)	95.55%	*15.147*	3.375%
East-North Branch (m=1,n=2)	96.79%	*2.677*	1.620%
East Branch (m=1,n=1)	94.20%	*22.919*	4.366%
Central Branch (m=1,n=2)	93.80%	*13.305*	3.749%
Western-North Branch (m=2,n=1)	91.86%	*5.6729*	3.520%

The proposed Eigenload model(*Eq*.(12), *Eq*.(17)) for the load demand evaluation and forecasting, only involves the previous load data itself, however, many presented methods for the load forecast usually use some of other parameters[1,2,3], such as time, temperature, light intensity, wind speed, humidity, season, day type(workday, weekend) and hour of the day, etc. So, we provide a comparison of the proposed Eigenload model with the classical autoregressive (AR) model [2,3]. In the AR model, the load is assumed to be a linear combination of previous loads. The AR model is used to model the load profile [2], which is given by the authors of [16] as,

$$\hat{L}(k) = -\sum_{i=1}^{n} a_i L(k-i) + e(k) \quad \textbf{(18)}$$

Where, $\hat{L}(k)$ is the predicted load at time *k*, and $e(k)$ is a random load disturbance. The AR model (*Eq*.(18)) is of order *n*. The unknown coefficients a_i are estimated with the least-squares algorithm.

For the same tasks of forecasts of *N*-step-length (*N*=12) that for each of SG and its five branches, based on its previous 3 annual load demands to forecast its annual load demand in 2007, **Table 4** list errors in terms of *RMSE* (***the standard error σ***) between the actual load demand and forecast result and ***percent (ratio) of*** *RMSE* to mean of actual load demand by employing the two models.

Table4: Errors in terms of *RMSE* between actual load demands and forecast results for State Grid Corporation of China and its five branches in 2007.

Forecast Result in 2007	*The Proposed Eigenload Model*		*The AR Model (optimal order n) (range: >=2)*	
	σ	σ/mean	σ	σ/mean
State Grid Corp., China	56.265	3.420%	86.673	5.268%
North Branch	15.147	3.375%	19.798	4.411%
East-North Branch	2.677	1.620%	3.398	2.056%
East Branch	22.919	4.366%	80.089	15.258%
Central Branch	13.305	3.749%	12.241	3.449%
Western-North Branch	5.173	3.420%	5.237	3.462%

In **Table4**, it is shown that in terms of *RMSE* between actual load demand and forecast result and ***percent (ratio) of*** *RMSE* to mean of actual load demand, the proposed Eigenload model performs better than the AR model in the forecast tasks for State Grid Corporation of China and its five branches in 2007.

Table 5 displays correlation coefficients between actual load demands and forecast results by the employing the two models. In terms of correlation coefficients (in **Table5**) between the actual load demands and the forecast results for State Grid Corporation of China and five branches in 2007, ANOVA (Analysis of variance) statistical test for performance between the proposed Eigenload model and the AR model is listed in **Table 6.**

Table 5: Correlation coefficients between actual load demands and forecast results for State Grid Corporation of China (SG) and five Branches in 2007.

Forecast Result in 2007	***The Proposed Eigenload Model***	***The AR Model (optimal order n) (range: >=2)***
State Grid Corp., China	**97.39%**	**88.93%**
North Branch	**95.55%**	**84.86%**
East-North Branch	**96.79%**	**90.69%**
East Branch	**94.20%**	**83.52%**
Central Branch	**93.80%**	**93.04%**
Western-North Branch	**91.86%**	**88.58%**

Table 6: ANOVA statistical test for performance of the load movement demand forecasts between the proposed Eigenload model and the AR model

Source of Variation	***Sum of Squares***	***Degree of freedom.***	***Mean Squares***	***F***	***P***
Between	0.013313	1	0.013313	15.761	0.0026435
Within	0.0084468	10	0.00084468		
Total	0.021759	11			

In **Table6**, the very small *p*-value (significant level) of 0.0026435<<0.01 indicates that the proposed Eigenload model is outperforming than the AR model in the forecast tasks for State Grid Corporation of China and its 5 branches in 2007.

6. Conclusion and Future Study

PCA technology views of an object as being a weighted combination of some "basis" elements with information loss as little as possible in the mean-square sense. It exhibits vital importance in a variety of applications, and the eigenface is a typical one. The PCA approach may discovery the most effective structure in low dimensions by decomposing the data structure into orthogonal components. In this study, we employ the PCA in creating eigenvectors and using them for annual electric load movement evaluation and forecasting.

(1) Introduce a called Eigenload model for describing the yearly load movement based on the PCA. That is, given a set of *l* vectors of annual load moments where each vector denotes one yearly load movement, project each load movement onto the eigenspace and describe it by a linear combination of the *m* principal orthogonal eigenvectors in the eigenspace. The *m* basis components are called "Eigenloads" for they are eigenvectors and each load movement is described in this basis. Based on the called "Eigenloads", an Eigenload model for annual load movement evaluation (*Eq*.(12)) is presented.

(2) Propose a forecast method (*Eq*.(17)) employing the Eigenload model incorporated with the polynomial curve fitting. The forecast method employs the polynomial curve fitting algorithm to estimate its subsequent representation weights with respect to the "Eigenloads" generated by previous load demands. With the estimated subsequent representation weights with respect to the "Eigenloads", the subsequent movement of the annual load can be then predicted.

(3) Show workability and effectiveness of the proposed. It is shown that for each of State Grid Corporation of China (SG) and its five branches, each one yearly load demand can be well evaluated (*Eq*.(12)) by a linear combination of its few m principal orthogonal eigenvectors called "Eigenloads". For each of SG and its five branches, based on the "Eigenloads" generated by its previous 3 annual load demands, we employ the proposed Eigenload model ((*Eq*.(17))) to forecast its future annual load demand. It is manifested that the proposed forecast method yields satisfying results, which agree well with their actual load demand movements. Analysis of the results indicates that the proposed Eigenload model is outperforming than the classical autoregressive (AR) model on the forecast tasks.

(4) Only the load data itself (its previous data) is used in the proposed method, and other parameters are not directly involved, thus being concise with well effectiveness may be the major advantage of the proposed PCA-based model.

However, many presented methods for the load forecast usually use some of other parameters[1-3,18,19], such as time, temperature, light intensity, wind speed, humidity, season, day type(workday, weekend) and hour of the day, etc. To further extend and improve the proposed Eigenload model for the load analysis is still included in our future work:

(1) In current state of the proposed model, it could not treat well with the special situations, such as extreme weather conditions, which may be considerably emphasized in our future study.

(2) Towards hybrid methods with other analyze techniques, may be also a consideration.

ACKNOWLEDGMENTS

The project was supported by Scientific Research Fund of Hunan Provincial Science and Technology Department (2013GK3090) and Scientific Research Fund of Hunan Provincial Education Department (09C399) and Research fund of Hunan University of Science and Technology (E50811).The author is grateful to State Grid Corporation of China.

References

[1] Eugene A. Feinberg, Dora Genethliou. Load Forecasting. Applied Mathematics for Restructured Electric Power Systems: Optimization, Control, and Computational Intelligence, Spinger, pp. 269-285, 2005.

[2] H. K. Alfares and M. Nazeeruddin. Electric load forecasting: literature survey and classification of methods. *International Journal of Systems Science*, vol. 33, pp.23–34, 2002.

[3] Heiko Hahn, Silja Meyer-Nieberg, Stefan Picklaetc. Electric load forecasting methods: Tools for decision making, *European Journal of Operational Research,* vol.199, No.3, pp.902-907, 2009.

[4] HOBBS, B, F., JITPRAPAIKULSARN, S., KONDA, S., CHANKONG, V., LOPARO, K. A., and MARATUKULAM, D. J. Analysis of unit commitment of improved load forecasts. *IEEE Transactions on Power Systems*, vol.14, no.4, pp.1342-1348, 1999.

[5] Nagi, J. ,Yap, K.S. ,Nagi, F. ,Tiong, S.K. ,Ahmed, S.K. A computational intelligence scheme for the prediction of the daily peak load. *Applied Soft Computing Journal*, vol.11, no.8, pp. 4773-4788, 2011.

[6] Ádamo L. Santana, etc. PREDICT- Decision support system for load forecasting and inference: A new undertaking for Brazilian power suppliers. *International Journal of Electrical Power & Energy Systems*, Vol.38, No.1, pp. 97–104, 2012.

[7] Kanchan, A. Singh, K.B. Load modeling, estimation and forecasting .Universities Power Engineering Conference (UPEC), pp.1-4, Aug. 31 -Sept. 3, 2010.

[8] Tharam S. Dillon, M. A. Laughton. Expert Systems Applications in Power Systems. Prentice Hall, 1990.

[9] Song Yi, Sun Ya-ming. Wavelet-based short-term load forecasting. Proceedings of the Chinese Society of Universities for Electric Power System and Automation, Wuhan, China, 2002, pp.184-189.

[10] Juan M. Vilar , Ricardo Cao and Germán Aneiros. Forecasting next-day electricity demand and price using nonparametric functional methods. *International Journal of Electrical Power & Energy Systems*, vol.39, no.1, pp.48-55, 2012.

[11] D.C. Sansom, T. Downs, T.K. Saha. Evaluation of support vector machine based forecasting tool in electricity price forecasting for Australian national electricity market participants. *Australian Journal of Electrical & Electronics Engineering,* Vol. 22, No. 3, pp.227-233, 2003.

[12] Tharam S. Dillon and Dagmar Niebur. Artificial Neural Networks with Applications to Power Systems. London CRL Pub., 1996.

[13] Kusum Verma, K.R. Niazi. Supervised learning approach to online contingency screening and ranking in power systems. *International Journal of Electrical Power & Energy Systems*, Vol.38, No.1, pp. 97–104, 2012.

[14] De Jonghe, C., Hobbs, B. F., Belmans, R. Optimal Generation Mix with Short-Term Demand Response and Wind Penetration. *IEEE Transactions on Power Systems*, vol.27, no.2, pp. 830 – 839, 2012.

[15] J.C. Sousaa, L.P. Nevesa, and H.M. Jorgeb. Assessing the relevance of load profiling information in electrical load forecasting based on neural network models. *International Journal of Electrical Power & Energy Systems*. Available online 13 March 2012.

[16] LIU, K., SUBBARAYAN, S., SHOULTS, R. R., MANRY, M. T., KWAN, C., LEWIS, F. L., and NACCARINO, J. Comparison of very short-term load forecasting. *IEEE Transactions on Power Systems*, vol.11, pp.877-882, 1996.

[17] S. Theodoridis and K. Koutroumbas. Pattern Recognition (Fourth Edition). Elsevier, China Machine Press, 2009. ISBN 978-7-111-26896-3.

[18] J. W. Taylor, L.M. de Menezes, P. E. McSharry.A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*, vol.22, pp.1-16, 2006.

[19] J. W. Taylor, P. E. McSharry.Short-Term Load Forecasting Methods: An Evaluation Based on European Data. *IEEE Transaction on Power Systems*, vol.22, pp.2213-2219, 2008.

[20] M. Turk and A. Pentland, "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.

[21] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms," *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 22, No. 10: 1090-1104,2000.

[22] T.Sudhanshu, L. Singh and H. Arora. "Face Recognition Machine Vision System Using Eigenfaces." *International Journal of Recent Trends in Engineering*, Vol.2, No.2, pp.1-3, 2009.

[23] T. Orczyk, P. Porwik, "An attempt to improve eigenface algorithm efficiency for colour images," *Journal of Medical Informatics &Technologies*, Vol.16, pp. 201-207, 2010.

[24] State Grid Corporation of China. Electricity Sale in January of 2005. *http://www.sp.com.cn/dlsc/dltj/gjdwgs/200501yb/200805/t20080515_104510.htm*,Retrieved 2012.

[25] Yang Zong-chang. Eigen-temperature Model for the Annual Air Temperature Movement Evaluation and Forecast. *International Journal of Modeling, Simulation, and Scientific Computing*. DOI: 10.1142/S1793962313500086, in Press, 2013.