

پیش‌بینی بار کوتاه‌مدت با انتخاب ورودی به روش

LLE و موتور پیش‌بینی ترکیبی RBF-Fuzzy

حامد کبریائی^{۱*}، حوراء کمالی نژاد^۲ و بابک نجار اعرابی^۳

^۱دانشکده مهندسی برق و کامپیوتر، پردیس دانشکده‌های فنی، دانشگاه تهران

^۲دفتر بازار برق، معاونت بهره برداری، شرکت برق منطقه‌ای تهران

چکیده

در این مقاله یک روش یادگیری خم تحت عنوان Locally Linear Embedding (LLE) برای استخراج ورودی‌های دارای اطلاعات بیشتر و کاهش ابعاد فضای ورودی جهت پیش‌بینی بار کوتاه‌مدت، پیشنهاد شده است. روش LLE رابطه غیرخطی بین ویژگی‌ها را با تصویرکردن یک خم خطی محلی در فضای ویژگی، می‌یابد. برای بررسی تأثیر روش پیشنهادی در خطای پیش‌بینی بار، یک سامانه پیش‌بینی ترکیبی، از شبکه‌ای با یک تابع پایه رادبال (RBF) و سامانه‌ای فازی، پیشنهاد شده است. شبکه RBF هسته موتور پیش‌بینی و ورودی آن تاریخچه بار است. سامانه استنتاج فازی جهت دخالت دادن اثر دما بر بار، با شبکه RBF ترکیب شده است. شبیه‌سازی با داده‌های واقعی بار منطقه مازندران، کارایی موتور پیش‌بینی پیشنهادی را در مقایسه با روش‌های شبکه عصبی مصنوعی، سری زمانی و شبکه نورو- فازی نشان می‌دهد؛ علاوه بر این، روش انتخاب ورودی (LLE) با روش تحلیل مؤلفه اصلی (PCA) و انتخاب تجربی ورودی‌ها، مقایسه شده است. نتایج شبیه‌سازی با تجزیه و تحلیل آماری معنادار، نشان می‌دهد که روش پیشنهادی، نسبت به سایر روش‌های انتخاب ورودی و موتورهای پیش‌بینی، دارای ابعاد ورودی کوچک‌تر و خطای پیش‌بینی کمتر است.

واژگان کلیدی: یادگیری خم به روش (LLE)، انتخاب ورودی، شبکه RBF، سامانه فازی، پیش‌بینی بار الکتریکی کوتاه‌مدت.

Short term load forecast by using Locally Linear Embedding manifold learning and a hybrid RBF-Fuzzy network

Hamed Kebriaei^{1*}, Howra Kamalinejad² & Babak Nadjar Araabi³

^{1,3}School of Electrical and Computer Engineering, College of Engineering, University of Tehran

²Electricity Market Office, Deputy of Operations, Tehran Regional Electricity Company

Abstract

The aim of the short term load forecasting is to forecast the electric power load for unit commitment, evaluating the reliability of the system, economic dispatch, and so on. Short term load forecasting obviously plays an important role in traditional non-cooperative power systems. Moreover, in a restructured power system a generator company (GENCO) should predict the system demand and its corresponding price for efficient decision making.

The task of a forecasting engine is to find the relation of the inputs and outputs of the system and also predicts the outputs for a given inputs. Therefore, the accuracy of forecasting is highly affected by the inputs of the forecasting engine. This effect can be studied from two points of view; First, extracting the more informative inputs and second, reducing the dimension of input space, both make it possible to learn the forecasting network via more simple models with more generalization. As a result, a reduced informative

* Corresponding author

*نویسنده عهده‌دار مکاتبات

input space leads to lower prediction error. In many previous load forecasting methods, the inputs have been selected empirically. In this manner, the more correlative factors with the load in the forecasting day have been chosen as the inputs. They are generally a combination of load history and weather conditions. Several researches are focused on mathematical approaches of the input selection which are mainly based on principal component analysis (PCA) method as well as some intelligent algorithms.

In this paper, a manifold learning method namely Locally Linear Embedding (LLE) is proposed, aiming to extract more informative inputs and to reduce the dimension of input space for short term load forecasting. Among all methods based on manifold learning, it can be seen that LLE performs very well in extracting the electric load curve features. The aim of this paper is to analyze the features of the load curve for estimating this curve in future. The extensive computational experiments show that the extracted features by LLE results in less prediction error than two other methods. Furthermore, LLE acts faster and makes input dimension lower than the two other methods. In the following section we will discuss the LLE method. The LLE method finds the nonlinear relationships among features by mapping a locally linear manifold in the feature space. Extracting the more informative inputs by extracting the combinational features by finding the nonlinear dependences of the features, results in reducing the dimension of input space. The resulted inputs from feature extraction and dimension reduction are utilized for load forecasting.

To examine the effect of the proposed feature extraction method on load prediction error, a hybrid prediction system is proposed which is a combination of a radial basis function (RBF) network and a fuzzy system. The RBF network is the core of the prediction engine and works with historical load data as its inputs. The fuzzy inference system is combined with the RBF network to incorporate the impact of temperature on load. The case studies are carried out on the real data of electric power load of Mazandaran area in Iran. The efficiency of the proposed forecasting engine is compared with three benchmarks, the artificial neural network, time series and neuro-fuzzy methods. Furthermore, the proposed input selection method (LLE) is compared with principal component analysis (PCA) and empirical selection of inputs. Simulation results with statistical significance analysis show that the LLE method with the proposed forecasting engine is superior to other input selection methods and forecasting engines in sense of lower input dimension and lower prediction error.

Keywords: Manifold learning, input selection, RBF network, fuzzy system, short term load forecasting.

۱- مقدمه

مسأله انتخاب ورودی در بیش تر روش های پیش بینی بار، تجربی است. به این ترتیب که متغیرهایی را که به طور تجربی همبستگی بیشتری با بار روز پیش بینی دارند، به عنوان ورودی انتخاب می شوند که به طور معمول به صورت ترکیبی از تاریخچه بار و شرایط آب و هوایی است. چندین پژوهش در زمینه انتخاب ورودی با رویکردهای ریاضی انجام شده که بر اساس روش تحلیل مؤلفه های اساسی^۱ PCA و همچنین برخی از الگوریتم های هوشمند است.

در [1,2] روش PCA برای انتخاب ورودی یک شبکه عصبی برای پیش بینی بار کوتاه مدت به کار رفته است. در [3] PCA به عنوان انتخاب گر ورودی برای یک LS-SVM^۲ به کار رفته و نشان داده شده است که استفاده از PCA دقت پیش بینی و به نوعی توان تعمیم پذیری سامانه را نسبت به حالتی که انتخاب ورودی صورت نمی گیرد، افزایش داده است. در [4] PCA برای انتخاب ورودی در روش های ARIMA و AR تناوبی جهت تأثیر دادن تناوب فصول در سری های زمانی،

هدف اصلی از پیش بینی بار کوتاه مدت، برآوردی از بار است که برای زمان بندی تولید پایه، ارزیابی قابلیت اطمینان عملکرد سامانه، پخش بار اقتصادی و آگاه سازی به موقع توزیع گر بار، مورد نیاز است. بدیهی است که پیش بینی بار کوتاه مدت، نقش مهمی در سامانه های قدرت غیر رقابتی سنتی ایفا می کند. در یک سامانه قدرت تجدید ساختار یافته نیز، یک تولید کننده، باید تقاضای سامانه و بهای متناظرش را به منظور تصمیم گیری بهینه در بازار، پیش بینی نماید.

وظیفه یک موتور پیش بینی، پیدا کردن رابطه ورودی ها و خروجی های سامانه و پیش بینی خروجی ها برای یک ورودی داده شده است. بنابراین، دقت پیش بینی متأثر از ورودی موتور پیش بینی است. این اثر را می توان از دو نقطه نظر مطالعه کرد. نخست استخراج ورودی های دارای اطلاعات بیشتر و دوم، کاهش ابعاد فضای ورودی. هر دوی این موارد، این امکان را می دهد که یادگیری در شبکه پیش بینی، از طریق مدل های ساده تر با تعمیم پذیری بیشتر، صورت گیرد. در نتیجه، یک فضای ورودی کاهش یافته منجر به خطای پیش بینی پایین تر می شود.

¹ Principal Component Analysis

² Least Square-Support Vector Machine

به‌کار رفته است. در [5] از الگوریتم KPCA¹ برای انتخاب ورودی برای پیش‌بینی بار استفاده شده است. روش KPCA تعمیمی برای روش PCA است، که در آن ابتدا یک تبدیل غیر خطی بر داده‌ها اعمال می‌شود و سپس PCA در فضای جدید کاهش بعد انجام می‌دهد. برای محاسبه ضرب داخلی در ابعاد بالا، از کرنل‌ها استفاده و به همین علت به این روش KPCA گفته می‌شود؛ در نهایت ویژگی‌های انتخاب‌شده توسط KPCA به‌عنوان ورودی‌های یک SVM برای پیش‌بینی بار استفاده شده است. در [6] نیز یک شبکه عصبی سه‌لایه برای پیش‌بینی پیک بار استفاده شده است که در آن یازده پارامتر آب و هوایی، تاریخچه بار به همراه کد روزهای هفته به‌عنوان ورودی‌های شبکه در نظر گرفته شده‌اند. در اینجا نیز از PCA برای کاهش ابعاد ورودی استفاده شده است. در [7] روش آنتروپی اطلاعات² برای انتخاب ورودی موتور پیش‌بینی بار به کار رفته است. به این ترتیب که اگر مجموعه کل ویژگی‌ها را با C و بار پیش‌بینی را D بنامیم، ابتدا مقادیر اهمیت تأثیر پارامتر $a_i \in C$ را در بار با تابع $I(a_i, D)$ محاسبه می‌شود که خود تابعی از آنتروپی است؛ سپس با قراردادن یک حد آستانه، a_i های را که $I(a_i, D)$ آنها از حد آستانه بیشتر باشد به‌عنوان ورودی موتور پیش‌بینی انتخاب می‌شود. علاوه‌براین در [8] از روش تحلیل مؤلفه مستقل (ICA³) جهت کاهش ابعاد ورودی و استخراج ویژگی استفاده شده و همچنین این روش با دو روش PCA و KPCA مقایسه شده است.

در [22] از تجزیه سری زمانی و شبکه عصبی برای پیش‌بینی بار استفاده شده است. برای شبیه‌سازی، از داده‌های شمال آمریکا استفاده و نتایج مدل طراحی شده، با سه مدل بر پایه شبکه عصبی مقایسه شده است.

در [23] از ترکیب شبکه عصبی بهبودیافته با کلونی زنبور عسل برای یافتن بهترین وزن‌ها و بایاس‌ها برای کمینه‌کردن خطای پیش‌بینی بار استفاده شده است. نتایج شبیه‌سازی بر روی داده‌های یک هفته از بار منطقه نیوانگلند نشان می‌دهد که روش مورد نظر از روش شبکه عصبی و عصبی - فازی، نتایج بهتری دارد.

در [24] برای پیش‌بینی کوتاه‌مدت بار در سامانه قدرت استان اصفهان، از شبکه‌های عصبی پیشخور جهت پیش‌بینی، و از روش‌های تجربه خیره، مقایسه آماری و یک الگوریتم بر مبنای تحلیل مؤلفه اصلی برای آموزش داده‌های پالایش شده استفاده و نشان داده شده است که پالایش داده‌های اولیه،

خطای پیش‌بینی بار را کم می‌کند. ایده اصلی این مقاله، حذف داده‌های غیر معمول، قبل از شروع کار پیش‌بینی است. در [25] از یک روش پیشنهادی شامل چهار بخش برای پیش‌بینی هم‌زمان قیمت و بار استفاده شده است. نتایج شبیه‌سازی برای بار منطقه نیوانگلند و نیوسوت‌ویلز استرالیا ارائه شده که دقت پیش‌بینی قابل قبول بوده است. هدف اصلی، در این مقاله در نظر گرفتن تأثیر مدیریت بار و نتایج بار روش‌های دیگر مقایسه نشده است.

در [26] از روش منطقه‌بندی و تأثیر روزهای مشابه و دما، برای پیش‌بینی بار کل شبکه برق ایران استفاده شده است. روش پیش‌بینی، انتخاب روز مشابه بوده و میزان خطا به‌صورت میانگین سالانه در کل کشور محاسبه شده است. در [27] از شبکه عصبی RBF برای اختفای خطا در فریم‌های ویدئو استفاده شده که نشان‌دهنده کاهش حجم محاسبات در این شبکه عصبی است.

در [28] از ترکیب شبکه‌های عصبی و الگوریتم ژنتیک برای پیش‌بینی نیروی گازگرفتن از روی سیگنال الکترومایوگرام استفاده شده که نشان‌دهنده کاربرد شبکه‌های عصبی در حیطه پزشکی است.

آن‌گونه که در بررسی مقالات معتبر داخلی دیده شد، مقالات بسیار کمی برای شبیه‌سازی و آزمایش روش پیش‌بینی بار پیشنهادشده، از داده‌های مربوط به ایران استفاده کرده‌اند. مقالاتی هم که از داده‌های داخل کشور استفاده کرده‌اند، بسیار محدود بوده و بار منطقه‌ای مانند مازندران یا خوزستان را که وابستگی بسیار زیادی به تغییرات وضعیت جوی (دما و رطوبت، یا به‌طور کلی‌تر دمای احساس‌شده در محیط) دارند، مورد استفاده قرار نداده‌اند؛ لذا یکی از جنبه‌های نوآوری در این مقاله، استفاده از روشی است که برای داده‌های داخل کشور، به‌خصوص در منطقه‌ای با وابستگی زیاد بار به وضعیت جوی، پاسخ قابل قبول داشته باشد.

روش‌های انتخاب ویژگی مطرح‌شده در مقالات برای پیش‌بینی بار کوتاه‌مدت، تنها ارتباط خطی بین ویژگی‌ها را می‌یابند. در عین حال، وجود روابط غیر خطی بین ورودی‌ها، اجتناب‌ناپذیر است. بنابراین برای پیدا کردن این نوع روابط غیر خطی، یک تبدیل غیر خطی مورد نیاز است. در این مقاله، یک روش یادگیری خم⁴ محلی به نام روش جاسازی خطی محلی (LLE)⁵ برای این منظور مورد استفاده قرار می‌گیرد. استفاده از این روش مزایای زیر را دارد:

⁴ Manifold learning

⁵ Locally Linear Embedding

در [22] از تجزیه سری زمانی و شبکه عصبی برای پیش‌بینی بار استفاده شده است. برای شبیه‌سازی، از داده‌های شمال آمریکا استفاده و نتایج مدل طراحی شده، با سه مدل بر پایه شبکه عصبی مقایسه شده است.

در [23] از ترکیب شبکه عصبی بهبودیافته با کلونی زنبور عسل برای یافتن بهترین وزن‌ها و بایاس‌ها برای کمینه‌کردن خطای پیش‌بینی بار استفاده شده است. نتایج شبیه‌سازی بر روی داده‌های یک هفته از بار منطقه نیوانگلند نشان می‌دهد که روش مورد نظر از روش شبکه عصبی و عصبی - فازی، نتایج بهتری دارد.

در [24] برای پیش‌بینی کوتاه‌مدت بار در سامانه قدرت استان اصفهان، از شبکه‌های عصبی پیشخور جهت پیش‌بینی، و از روش‌های تجربه خیره، مقایسه آماری و یک الگوریتم بر مبنای تحلیل مؤلفه اصلی برای آموزش داده‌های پالایش شده استفاده و نشان داده شده است که پالایش داده‌های اولیه،

در [22] از تجزیه سری زمانی و شبکه عصبی برای پیش‌بینی بار استفاده شده است. برای شبیه‌سازی، از داده‌های شمال آمریکا استفاده و نتایج مدل طراحی شده، با سه مدل بر پایه شبکه عصبی مقایسه شده است.

در [23] از ترکیب شبکه عصبی بهبودیافته با کلونی زنبور عسل برای یافتن بهترین وزن‌ها و بایاس‌ها برای کمینه‌کردن خطای پیش‌بینی بار استفاده شده است. نتایج شبیه‌سازی بر روی داده‌های یک هفته از بار منطقه نیوانگلند نشان می‌دهد که روش مورد نظر از روش شبکه عصبی و عصبی - فازی، نتایج بهتری دارد.

در [24] برای پیش‌بینی کوتاه‌مدت بار در سامانه قدرت استان اصفهان، از شبکه‌های عصبی پیشخور جهت پیش‌بینی، و از روش‌های تجربه خیره، مقایسه آماری و یک الگوریتم بر مبنای تحلیل مؤلفه اصلی برای آموزش داده‌های پالایش شده استفاده و نشان داده شده است که پالایش داده‌های اولیه،

¹ Kernel Principal Component Analysis

² Information Entropy

³ independent component analysis

۱- استخراج ورودی‌های دارای اطلاعات بیشتر، به کمک استخراج ویژگی‌های ترکیبی؛
 ۲- کاهش ابعاد فضای ورودی، پیدا کردن وابستگی‌های غیر خطی ویژگی‌ها.

ورودی‌های حاصل از استخراج ویژگی‌های ورودی و کاهش ابعاد ورودی، برای پیش‌بینی بار به کار می‌روند. سامانه پیش‌بینی پیشنهادی، شامل یک شبکه RBF و یک سامانه تصحیح فازی برای پیش‌بینی بار روز آینده است. نقش شبکه RBF، پیش‌بینی بار بر اساس داده‌های تاریخیچه بار و نقش سامانه فازی، اصلاح شکل بار بر اساس تغییرات آب و هوایی است. چینش آبخاری این دو سامانه پیش‌بینی و اصلاح، بر اساس مطالعات آماری بر روی الگوی بار در منطقه مازندران است؛ علاوه بر این، مطالعات روان‌شناختی، اجتماعی و اقتصادی منطقه مازندران، تعدادی از قوانین منطقی تغییر بار را بر اساس تغییرات دما، به دست می‌دهد که برای ساختن قواعد سامانه فازی به کار می‌رود. این مقاله از شش بخش تشکیل شده است؛ بخش نخست مقدمه و مرور ادبیات موضوع و انگیزه استفاده از یادگیری خم به جای PCA در بخش دو مورد بحث قرار گرفته است. انتخاب ورودی توسط الگوریتم LLE در بخش سه معرفی شده و سامانه پیش‌بینی، که ترکیبی از یک شبکه RBF و یک سامانه فازی است، در قسمت چهارم معرفی و بررسی شده است. بخش پنجم نتایج شبیه‌سازی را برای مطالعه موردی، نشان می‌دهد و در نهایت بخش ششم نتیجه‌گیری این مقاله است.

۲- پیش‌نیاز مسأله

محبوب‌ترین الگوریتم برای کاهش ابعاد ورودی، روش تجزیه و تحلیل مؤلفه‌های اصلی (PCA) است. با داشتن یک مجموعه داده، PCA جهت‌هایی (بردارهایی) را پیدا می‌کند که در آنها داده‌ها دارای بیشترین واریانس بوده و این جهت‌ها نیز دارای اهمیت نسبی باشند. PCA زمانی مفید و کاربردی است که داده‌ها، بر روی یا در نزدیکی یک زیرمجموعه خطی از مجموعه داده‌ها، قرار گیرند. با در دست داشتن این نوع داده‌ها، PCA یک پایه برای زیرمجموعه خطی پیدا می‌کند و اجازه می‌دهد که ویژگی‌های غیر مرتبط را نادیده بگیریم [9]. روش PCA شامل مراحل زیر است [3]:

مرحله ۱: داده‌های ابتدایی را هنجارسازی می‌کنیم. برای از بین بردن اثر چند مقیاس بودن داده‌ها و تفاوت مقیاس در

مقادیر عددی، داده‌های اولیه باید هنجارسازی شود. مرحله ۲: ماتریس همبستگی را ایجاد کرده و مقادیر ویژه و بردارهای ویژه آن را محاسبه می‌کنیم.

$$R = \frac{1}{N-1} X^* T X^* \quad (1)$$

که در آن N تعداد نمونه‌ها و X^* ماتریس داده‌های هنجارسازی شده است. مقادیر ویژه R عبارتند از $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ و بردارهای ویژه متناظر با آنها عبارتند از u_1, u_2, \dots, u_m .
 مرحله ۳: تعداد مؤلفه‌های اصلی را تعیین می‌کنیم. توزیع واریانس تجمع‌شده را می‌توان به صورت زیر محاسبه کرد:

$$\eta(p) = \sum_{i=1}^p \lambda_i / \sum_{i=1}^m \lambda_i \quad (2)$$

تعداد مؤلفه‌های اصلی به توزیع واریانس تجمع‌شده بستگی دارد. به طور معمول وقتی $\eta(p)$ بزرگ‌تر از 0.85 است، P مؤلفه اصلی متناظر می‌تواند شامل بیشترین اطلاعات از متغیرهای اولیه باشد.

مرحله ۴: ماتریس ویژگی $U_{m \times p}$ را که شامل نخستین P بردار ویژه از ماتریس همبستگی است، تشکیل می‌دهیم. این ماتریس می‌تواند برای تبدیل مجموعه داده‌های اولیه که دارای ابعاد بالاتر است، به مجموعه داده‌های جدید که دارای بعد پایین‌تر است، به کار رود.

$$Z_{N \times p} = X_{N \times m}^* U_{m \times p} \quad (3)$$

که در آن $U_{m \times p} = [u_1, u_2, \dots, u_p]$

با وجود محبوبیت PCA، این روش محدودیت‌هایی دارد. مهم‌ترین نقض PCA این است که ضرورت دارد، داده‌ها بر روی یک زیر فضای خطی قرار گیرند [9]. الگوریتم یادگیری خم یک روش جدید و غیر پارامتری برای کاهش بعد غیرخطی است. الگوریتم‌های یادگیری خم می‌تواند به عنوان یک نسخه غیر خطی PCA در نظر گرفته شود. در این الگوریتم‌ها، یک مجموعه داده با ابعاد بالا که در کاربردهای واقعی مورد استفاده قرار می‌گیرد، می‌تواند به عنوان مجموعه‌ای از داده‌ها که نزدیک به یک خم با ابعاد پایین‌تر قرار دارند، مدل‌سازی شود. فرض کنید مجموعه داده $x_1, x_2, \dots, x_N \in R^D$ داده شده است. اگر داده‌ها بر روی یک زیرفضای R^d با ابعاد پایین‌تر قرار داشته باشند، PCA ابزار مناسبی است. در عوض، فرض می‌کنیم که داده‌ها بر روی یک خم با ابعاد d قرار می‌گیرند که این خم در درون R^D است و $d < D$.

می‌رود. انتخاب ورودی به روش LLE در زیر شرح داده خواهد شد:

فرض می‌کنیم مجموعه داده مربوط به N روز در فضای R^D با بعد بالا داده شده است: $x_1, x_2, \dots, x_N \in R^D$. ما به دنبال بردارهای ویژگی ذاتی جاسازی شده در فضای با بعد پایین، یعنی $y_1, y_2, \dots, y_N \in R^d$ هستیم. LLE روشی غیر پارامتری است که به شرح زیر عمل می‌کند [11,12]:

(۱) نزدیک‌ترین k همسایگی‌ها برای هر نقطه $x_i, 1 \leq i \leq N$ را بر اساس فاصله اقلیدسی پیدا می‌کنیم. مجموعه اندیس‌های نزدیک‌ترین k همسایگی‌های x_i را با N_i نشان می‌دهیم. در زمینه انتخاب k کارهای کمی انجام شده است که از بهترین آنها می‌توان به [20] اشاره کرد؛ با این حال این روش‌ها تنها برای LTSA مورد بحث قرار گرفته‌اند. در این مقاله ما k را به صورت تجربی انتخاب می‌کنیم؛ به نحوی که بیشترین بازدهی را در مجموعه داده‌های آموزش به دست آوریم.

(۲) ترکیبی محدب و بهینه محلی از نزدیک‌ترین k همسایگی x_i را برای بیان کردن داده x_i می‌یابیم، که معادل با کمینه کردن تابع هزینه زیر با ضرایب وزنی w_i به عنوان متغیرهای تصمیم‌گیری هستند:

$$E(w) = \sum_{i=1}^N \left\| x_i - \sum_{j \in N_i} w_{ij} x_j \right\|^2 \quad (5)$$

که در آن $\| \cdot \|$ نرم l_2 است و $\sum_{j \in N_i} w_{ij} x_j = x_i$

ضریب وزن w_{ij} نشان‌دهنده سهم داده i -ام در نمایش داده i -ام است. وزن‌های بهینه را می‌توان از حل مسأله کمینه مربعات مقید به دست آورد، که در نهایت به حل یک مسأله سامانه معادلات خطی، منتهی می‌شود.

(۳) بردارهای ویژگی با بعد کمتر $y_i, 1 \leq i \leq N$ را پیدا می‌کنیم که دارای نمایش بهینه محدب محلی با وزن‌های w_{ij} است که از مرحله قبل به دست می‌آیند. y_i ها از کمینه سازی تابع هدف زیر به دست می‌آیند:

$$\Phi(y) = \sum_{i=1}^N \left\| y_i - \sum_{j \in N_i} w_{ij} y_j \right\|^2 \quad (6)$$

که می‌تواند به فرم درجه دو زیر نوشته شود:

$$\Phi(y) = \sum_{ij} M_{ij}(y_i, y_j) \quad (7)$$

فرض می‌کنیم که داده‌ها از یک خم نمونه برداری

شده‌اند:

$$x_i = f(y_i) \quad i = 1, 2, \dots, N \quad (4)$$

که $y_i \in R^d$. عدد صحیح d بُعد ذاتی نامیده می‌شود. روش مبتنی بر خم، روشی را برای یافتن بردارهای ویژگی با بُعد کمتر $y_i \in R^d$ ، جاسازی شده درون نقاط داده x_i با بعد بالاتر، پیشنهاد می‌دهد.

چندین روش یادگیری غیر پارامتری برای یادگیری خم غیرخطی، وجود دارد، مانند تغییر مقیاس چندبعدی (MDS) [10,11] جاسازی خطی محلی (LLE) [12,13]، Isomap [14]، تبدیل‌های ویژه لاپلاسی [15]، تبدیل‌های ویژه هسیان [16]، همبستگی محلی فضای تانژانت (LTSA) [17] و تبدیل‌های انتشار [18]. مروری بر روش‌های یادشده، در [19] بیان شده است. در میان تمام روش‌های مبتنی بر یادگیری خم، دیده می‌شود که LLE در استخراج ویژگی‌های منحنی بار الکتریکی بسیار خوب عمل می‌کند. هدف از این مقاله، تحلیل ویژگی‌های منحنی بار برای تخمین این منحنی در آینده است. آزمایش‌های محاسباتی متعدد نشان می‌دهد که ویژگی‌های استخراج شده از طریق LLE موجب خطای پیش‌بینی کمتر نسبت به دو روش دیگر می‌شود. علاوه بر این، LLE سریع‌تر عمل می‌کند و ابعاد ورودی در آن، کمتر از دو روش دیگر است. در بخش زیر روش LLE را مورد بحث قرار خواهیم داد.

۳- انتخاب ورودی به روش LLE

در مسأله پیش‌بینی بار، ورودی‌های مختلفی برای سامانه پیش‌بینی در نظر گرفته می‌شود که معمول‌ترین آنها عبارت‌اند از: بار روزهای گذشته، کد روز هفته، دما، رطوبت و... ورودی اصلی بسیاری از سامانه‌های پیش‌بینی بار، بار در روزهای گذشته (تاریخچه بار) است. در این بخش، به روش انتخاب ورودی LLE، با جزئیات اشاره می‌شود.

از آنجایی که هر روز دارای ۲۴ داده ساعتی بار است، ابعاد ورودی، بزرگ خواهد بود؛ بنابراین داده‌های n روز گذشته بار، برای پیش‌بینی بار روز هدف، به عنوان ورودی LLE در نظر گرفته می‌شود. الگوریتم LLE ابعاد ویژگی ورودی را، به فضای ورودی با اطلاعات بیشتر و بُعد کمتر، کاهش می‌دهد. ویژگی‌های به دست آمده توسط شبکه RBF، برای پیش‌بینی بار استفاده می‌شود. دیگر ورودی برای پیش‌بینی بار، بیشینه دمای روز است که به طور مستقیم توسط سامانه فازی به کار

است. سامانه فازی برای مقابله با تأثیر تغییرات دما بر شکل منحنی بار، با شبکه RBF ترکیب شده است. این روش پلکانی استفاده از دو سامانه پیش‌بینی‌کننده و اصلاح‌کننده، بر اساس تاریخچه شکل بار در منطقه مازندران در نظر گرفته شده است. قوانین پایه فازی نیز بر اساس قوانین به‌دست‌آمده از مطالعات روان‌شناختی، اجتماعی و اقتصادی در مورد اثر تغییرات دما بر بار مصرفی در منطقه مازندران، طراحی شده است. نقش سامانه فازی، شیفت‌دادن بار پیش‌بینی‌شده به بالا یا پایین بر اساس افزایش یا کاهش دما است. طرح‌بندی سامانه پیش‌بینی در شکل (۱) نشان داده شده است. پروفایل بار از هفته گذشته تا روز قبل از روز هدف، به‌عنوان ورودی الگوریتم LLE در نظر گرفته شده است. ویژگی‌های استخراج‌شده توسط LLE، به‌عنوان ورودی سامانه پیش‌بینی استفاده می‌شود. $D = 168$ و N برابر با تعداد روزها است. $\hat{u}(t)$ بار تخمین زده شده ساعت t توسط شبکه RBF و $T(t)$ دمای ساعت t است.

پس از کاهش بُعد به کمک روش LLE، فضای ورودی R^D به فضای R^d که $d < D$ است، تبدیل می‌شود. هدف در اینجا کم‌کردن مربعات خطای پیش‌بینی و به تبع آن قدرمطلق خطا است. سامانه پیش‌بینی در این حالت یک شبکه RBF است که خروجی آن، توسط یک سامانه فازی تصحیح می‌شود. شمای سامانه شناسایی به‌صورت شکل (۱) است. ورودی‌های سامانه تصحیح فازی ($\hat{u}(t)$) شامل بیشینه دمای پیش‌بینی شده روز هدف، بیشینه دمای هفته روز گذشته، پروفایل بار هفت روز گذشته و خروجی‌های سامانه RBF است. خروجی‌های سامانه فازی، خروجی‌های تصحیح‌شده سامانه RBF هستند. گفتنی است که علامت منفی نشان‌دهنده ساعت‌های قبل از روز هدف است.

که در بردارنده ضرب داخلی y_i است. ماتریس M ، ماتریسی مربعی با بعد N است و از رابطه زیر به‌دست می‌آید:

$$M_{ij} = \delta_{ij} - w_{ij} - w_{ji} + \sum_k w_{ik} w_{jk} \quad (8)$$

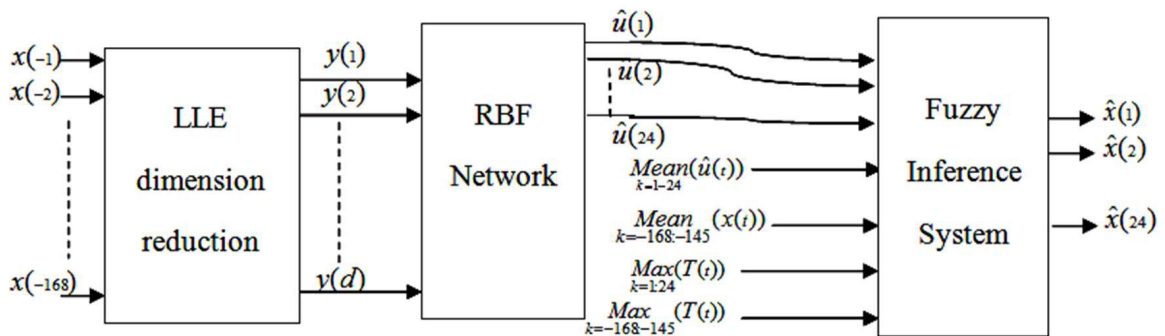
که در δ_{ij} آن برابر ۱ است اگر $i = j$ باشد و در غیر این صورت برابر صفر است. برای آنکه y_i ها را به‌صورت یکتا تعیین کنیم، شروط محدودکننده $\sum_i y_i = 0$ و $\frac{1}{N} \sum_i y_i y_i^T = I$ را قائل می‌شویم تا به‌ترتیب درجه آزادی انتقالی و گردشی را حذف کنیم. بهینه‌سازی درجه دو در رابطه (۵) همراه با قیود محدودکننده با ضریب لاگرانژ، قابل حل هستند؛ درنهایت، مسأله به پیدا کردن d بردار ویژه متناظر با کوچک‌ترین d مقادیر ویژه غیر صفر ماتریس M ، که مختصات d -بعدی y_i ها را بهینه می‌کند، منتهی می‌شود؛ بنابراین مختصات، y_i ها متعامد هستند. برای جزئیات بیشتر این نوع فرمولاسیون مسأله، به [11] مراجعه کنید.

پس از کاهش بعد توسط روش LLE، فضای ورودی از $(x_1, x_2, \dots, x_N) \in R^D$ به $(y_1, y_2, \dots, y_N) \in R^d$ تغییر می‌کند که $d < D$ است.

روش LLE هیچ‌گونه مدل احتمالی را بر داده‌ها تحمیل نمی‌کند. هرچند به‌طور ضمنی محدب‌بودن خم را فرض می‌کند.

۴- سامانه پیش‌بینی ترکیبی

هدف از سامانه پیش‌بینی، به کمینه‌رساندن درصد میانگین خطای مطلق (MAPE) است. موتور پیش‌بینی پیشنهادی، یک شبکه RBF (با ۲۴ خروجی که پیش‌بینی بار ساعتی روز آینده است) می‌باشد، که با یک سامانه فازی ترکیب شده



(شکل-۱): طرح‌بندی سیستم پیش‌بینی پیشنهادی
(Figure-1): The schematic of the proposed prediction system

۴-۱- شبکه RBF

ساعت و اندیس k به به شماره داده (که در بخش ۳ مطرح شد) اشاره دارند):

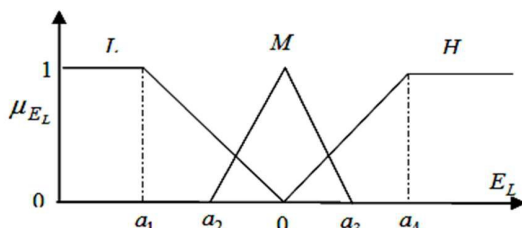
$$\begin{aligned} E_L &= \text{Mean}(\hat{u}(t)) - \text{Mean}(x(t)) \\ &= \frac{1}{k-1:24} - \frac{1}{k-168:-145} \end{aligned} \quad (11)$$

$$E_T = \text{Max}(T(t)) - \text{Max}(T(t))$$

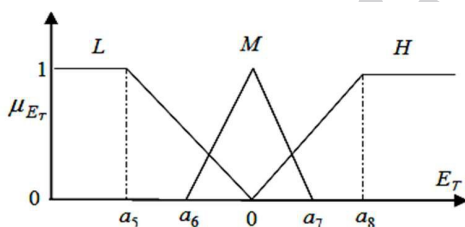
$$= \frac{1}{k=1:24} - \frac{1}{k=-168:-145}$$

به عبارت دیگر تفاضل مقدار میانگین بار پیش‌بینی شده و مقدار میانگین بار هفته قبل به‌عنوان یک ورودی سامانه فازی و تفاضل مقدار بیشینه دما در روز پیش‌بینی و مقدار بیشینه دما هفته قبل به‌عنوان ورودی دیگر سامانه فازی در نظر گرفته شده است.

حال فرض می‌کنیم که E_L و E_T می‌توانند مقادیر برجسب‌های فازی "کم" و "متوسط" و "زیاد" را اختیار کنند. توابع عضویت مربوط به این مجموعه‌های فازی در شکل‌های (۲) و (۳) نمایش داده شده است.

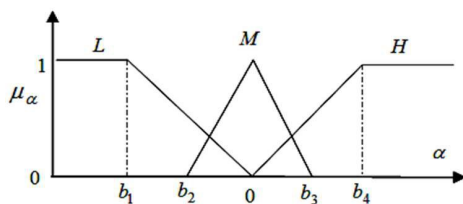


(شکل-۲): توابع عضویت ورودی نخست (E_L)
(Figure-2): The membership functions for EL



(شکل-۳): توابع عضویت ورودی دوم (E_T)
(Figure-3): The membership functions for ET

همچنین خروجی این سیستم فازی (α) نیز با توابع دوزنقه‌ای به صورت زیر بیان می‌شود:



(شکل-۴): توابع عضویت خروجی (α)
(Figure-4): The fuzzy output (α) membership function

برای آموزش شبکه RBF، روش‌های مختلفی وجود دارد. برای آموزش لایه پنهانی شبکه (مرکز و واریانس تابع گوسی هر نورون)، می‌توان از روش‌های متنوع خوشه‌بندی استفاده کرد. در این مقاله از روش خوشه‌بندی معروف K-means برای محاسبه مراکز توابع گوسی استفاده شده است. واریانس هر تابع گوسی را می‌توان به صورت ذهنی محاسبه کرد. به‌عنوان مثال می‌توان از میانگین فاصله هر نورون تا نورون‌های دیگر، استفاده کرد. به‌طور هم‌ارز:

$$\sigma_i = \text{Mean}(\|c_i - c_j\|), j = 1, 2, \dots, m \quad (9)$$

که m تعداد نورون‌ها و c_i مرکز هر خوشه، برداری با d مؤلفه است:

$$c_i = (c_{i1}, c_{i2}, \dots, c_{id}) \quad i = 1, 2, \dots, m \quad (10)$$

بعد از خوشه‌بندی فضای ورودی و محاسبه مراکز و واریانس‌های توابع گوسی تمام نورون‌ها، خروجی هر نورون قابل محاسبه است. پارامترهای لایه خروجی (بایاس‌ها و وزن‌ها) با روش‌های بهینه‌سازی مبتنی بر گرادیان قابل تخمین هستند. در این مقاله برای تخمین وزن‌های لایه خروجی شبکه، از روش کمینه مربعات استفاده شده است.

۴-۲- سامانه فازی

بار الکتریکی در مناطق مرطوبی که در آن‌ها دما نوسان زیادی دارد، به شدت تحت تأثیر دمای احساس شده^۱ است. در این مقاله، اثر دما (به‌عنوان عامل غالب برای تنظیم مقدار بار پیش‌بینی‌شده، در مناطقی که بار به شدت تحت تأثیر شرایط آب و هوایی است)، از طریق استفاده از یک سامانه خیره فازی مورد بررسی قرار می‌گیرد.

تخمین بار تنها بر اساس پیشینه بار، به‌خصوص در فصلی که دما به‌طور محسوسی در مقدار بار مصرفی تأثیرگذار است، موجب خطای زیادی در پیش‌بینی بار است. به این منظور پارامتر دما را برای تصحیح این پیش‌بینی به کار می‌بریم. دلیل اینکه این پارامتر را در ورودی شبکه RBF قرار ندادیم، دانش ما از نحوه تأثیرگذاری دما بر روی بار است. به همین منظور، تصحیح بار به کمک یک سامانه خیره فازی صورت می‌گیرد که توضیح آن در زیر آمده است.

این سامانه فازی از نوع ممدانی^۲ است و دو ورودی (شکل(۱)) به صورت زیر دارد (توجه کنید که اندیس t به

¹ Feeling temperature

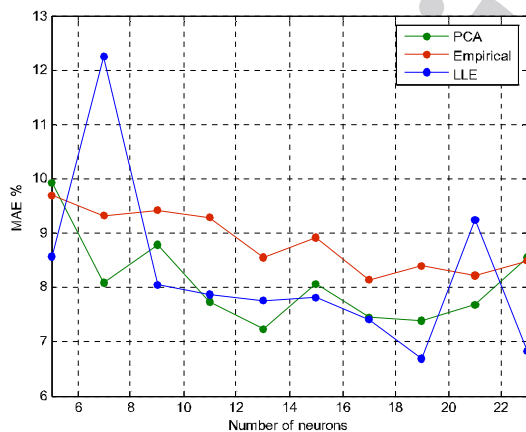
² Mamdani-type Fuzzy system

روزانه برای سال‌های ۸۴ و ۸۵، از یازدهم اردیبهشت تا نهم شهریور (ماه‌های می تا آگوست) است. داده‌های آزمون، هیجدهم خرداد تا هفدهم تیر سال ۸۶ (۸ ژوئن تا هشت ژوئیه) است. پیش‌بینی بار در منطقه مازندران در این دوره زمانی پیچیده است؛ چون که نوسانات درجه حرارت به شدت بر منحنی بار تأثیر می‌گذارد. علاوه بر این، در تاریخ ۲۸ خرداد (۱۸ ژوئن) یک تعطیلات قرار دارد که موقعیت خوبی برای آزمایش روش پیشنهادی ما است.

در ابتدا ساختار موتور پیش‌بینی در بخش ۱-۵ توضیح داده شده است. در بخش ۲-۵، سه روش انتخاب ورودی به سامانه پیش‌بینی اعمال شده و نتایج آنها مقایسه می‌شود. مقایسه شبکه پیش‌بینی پیشنهادی و محک‌های شناخته شده (بنچمارک‌ها)، در بخش ۵-۲-۵ داده شده و اهمیت آماری روش پیشنهادی در بخش ۵-۲-۶ مورد مطالعه قرار گرفته است.

۱-۵- موتور پیش‌بینی

برای تعیین ساختار شبکه RBF، ابتدا تعداد نورون‌های لایه میانی را تعیین می‌کنیم. برای این کار خطای پیش‌بینی روی داده‌های آزمون بر حسب تعداد نورون‌های لایه میانی در شکل زیر رسم شده است. در نتیجه، تعداد نورون‌های لایه میانی برای روش‌های تجربی، PCA و LLE به ترتیب ۱۷ و ۱۳ و ۱۹ انتخاب می‌شود.



(شکل-۵): میانگین قدرمطلق خطای پیش‌بینی بر حسب تعداد

نورون‌های لایه میانی

(Figure-5): The mean absolute error of prediction versus number of neurons

مراکز این نورون‌ها را نیز می‌توان به وسیله روش خوشه‌بندی K-means تعیین کرد. واریانس هر تابع گوسی را می‌توان از رابطه (۹) تخمین زد. پس از تنظیم مراکز، وزن لایه خروجی RBF با الگوریتم کمینه مربعات تخمین زده

قوانین این سامانه فازی، بر اساس دانشی که در مورد نحوه تأثیرگذاری دما بر بار منطقه مازندران داریم، تعیین می‌شود. این قوانین در جدول زیر آمده است:

(جدول-۱): قوانین سامانه فازی

(Table-1): Fuzzy rule Base

RULES	E_L	E_T	α
R1	L	L	M
R2	L	M	M
R3	L	H	H
R4	M	L	L
R5	M	M	M
R6	M	H	M
R7	H	L	L
R8	H	M	L
R9	H	H	M

حال با توجه به این که درجه عضویت E_L و E_T به ترتیب μ_{E_L} و μ_{E_T} است، شدت آتش یا به عبارتی ارزش هر قانون، با اپراتور min فازی، از رابطه زیر به دست می‌آید:

$$\mu_i = \text{Min}(\mu_{E_L}(i), \mu_{E_T}(i)) \quad (12)$$

در عبارت بالا i به شماره قانون اشاره می‌کند و در هر قانون $\mu_{E_L}(i)$ ، $\mu_{E_T}(i)$ چنانچه در شکل‌های (۲) و (۳) نشان داده شده است، نمایانگر درجه عضویت L، M و یا H باشند.

در نهایت خروجی سامانه فازی که ضریب تصحیح بار (W) است، با عمل عکس فازی‌سازی (روش مرکز میانگین) بر اساس رابطه زیر به دست می‌آید:

$$W = \frac{\sum_{i=1}^9 C_{\alpha_i} \mu_i}{\sum_{i=1}^9 \mu_i} \quad (13)$$

در رابطه بالا C_{α_i} مرکز تابع عضویت خروجی مربوط به i امین قانون فازی است. بار پیش‌بینی شده توسط شبکه RBF با ضریب W تصحیح می‌شود:

$$\hat{x}(t) = (1+W) \cdot \hat{u}(t) \quad (14)$$

در رابطه بالا، $\hat{x}(t)$ بار پیش‌بینی شده برای ساعت t است.

۵- شبیه‌سازی

هدف از این شبیه‌سازی، پیش‌بینی بار الکتریکی یک روز قبل، در منطقه مازندران است. مجموعه داده‌های آموزش برای شبکه RBF-Fuzzy شامل بار روزانه و بیشینه دمای

Defuzzification

شبه‌سازی نمی‌افزاید؛ همچنین انتخاب تعداد کمتر از سه تابع عضویت، کارایی مدل را کاهش می‌دهد.

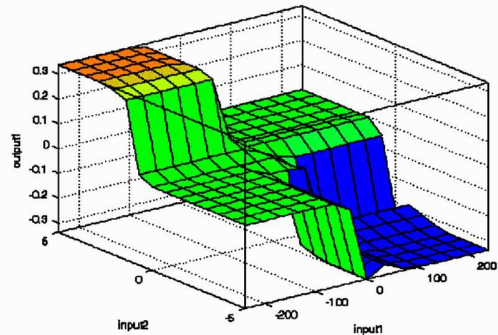
۲-۵- روش‌های انتخاب ورودی

در این بخش، روش انتخاب ورودی پیشنهاد شده (LLE) با دو روش مهم دیگر ورودی (انتخاب ورودی تجربی، PCA) مقایسه می‌شود. در هر مورد، ورودی‌های حاصل به سامانه پیش‌بینی RBF-Fuzzy تغذیه می‌شوند. خطاهای پیش‌بینی محاسبه شده و میانگین قدر مطلق درصد خطاها (MAPE) در سه روش مقایسه می‌شود. در این بخش، موتور پیش‌بینی (Fuzzy-RBF) نیز با برخی از ابزارهای پیش‌بینی معیار، مقایسه می‌شود. علاوه بر این، اهمیت آماری نتایج با استفاده از روش T-Test نشان داده شده است.

۱-۲-۵- روش انتخاب ورودی تجربی

در انتخاب ورودی تجربی، بار روز گذشته (۲۴ ورودی) و یک هفته قبل (۲۴ ورودی) به‌عنوان ورودی سامانه پیش‌بینی در نظر گرفته شده است. بار روز گذشته به‌علت نزدیکی به روز پیش‌بینی و به‌عنوان نزدیک‌ترین داده در دسترس، همبستگی زیادی با بار روز پیش‌بینی دارد و همچنین بار هفته گذشته به‌علت یکسان بودن کد روز هفته دارای همبستگی زیادی با بار روز پیش‌بینی است. همچنین پیشینه دمای هفته گذشته و پیشینه دمای تخمین زده شده برای روز پیش‌بینی، ورودی‌های سامانه تصحیح فازی هستند. نتایج پیش‌بینی در شکل (۸) نشان داده شده است.

می‌شود. خروجی شبکه RBF توسط سامانه فازی استفاده می‌شود. مقادیر ثابت توابع عضویت ورودی و خروجی در جدول (۲) آورده شده است. سطح تصمیم‌گیری سه‌بعدی سامانه تصحیح فازی (خروجی در برابر دو ورودی) در شکل (۶) نشان داده شده است. مقادیر مختلف ضریب تصحیح بار (W) در مقادیر مختلف ورودی، به‌دست آورده می‌شود.



(شکل-۶): سطح تصمیم‌گیری سه‌بعدی سیستم فازی (Figure-6): 3D decision surface of the fuzzy system

توابع عضویت فازی با توجه به محدوده داده‌های ورودی (E_L و E_T) و خروجی (α) و دامنه تغییرات آنها تعیین شده است. به این ترتیب که ابتدا با توجه به محدوده تغییرات، بازه‌های a_1 تا a_8 و b_1 تا b_4 تعیین شده و سپس هر یک از این محدوده‌ها با سه تابع عضویت فازی با برجسب‌های کم، متوسط و زیاد پوشش داده شده است. دلیل انتخاب سه تابع عضویت این است که انتخاب تعداد بیشتر تابع عضویت به کارایی مدل بر اساس نتایج

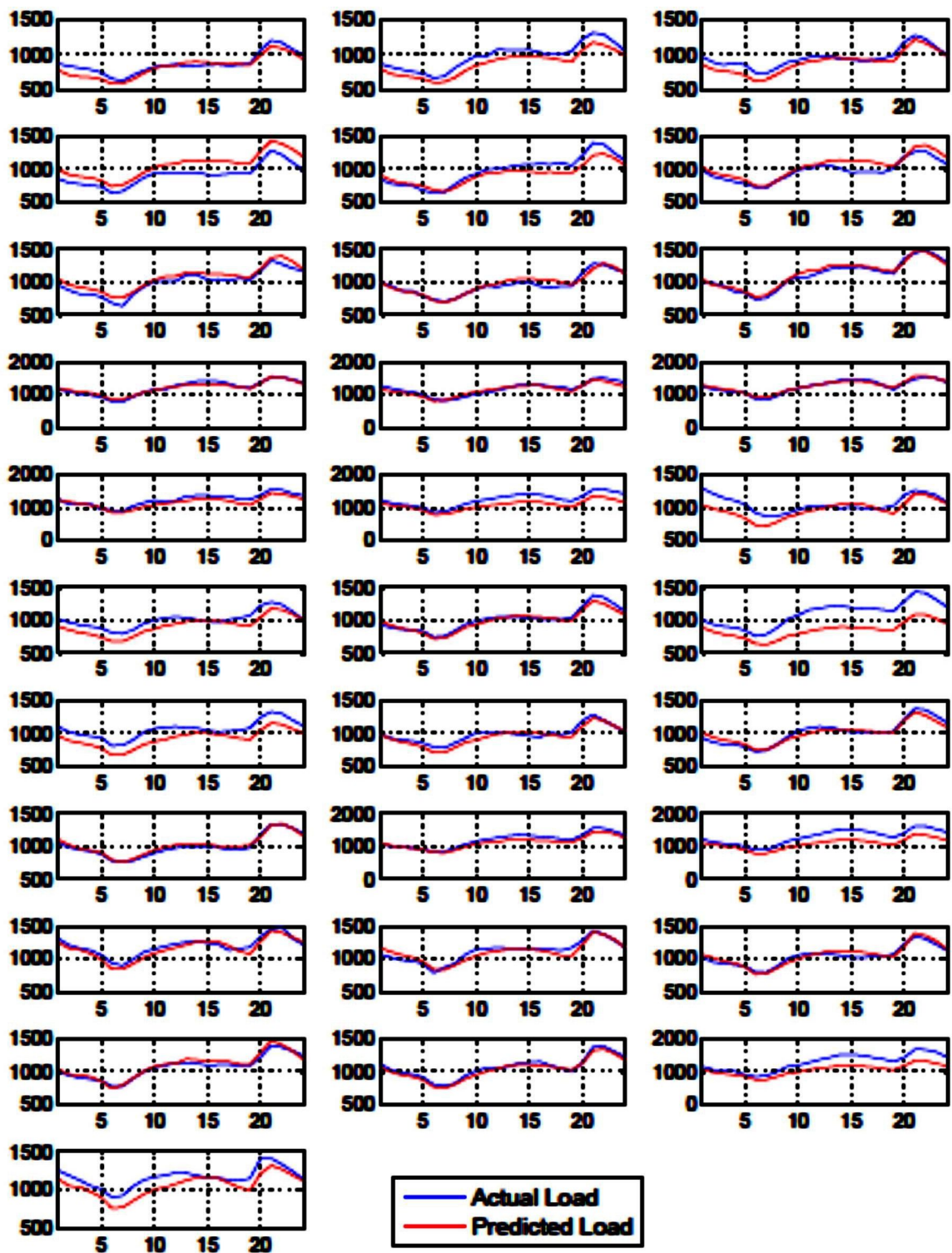
(جدول-۲): مراکز توابع عضویت ورودی‌ها و خروجی (Table-2): Input and output membership functions' centers

Parameter	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	b_1	b_2	b_3	b_4
Value	-100	-50	50	100	-3	-1	1	3	-0.5	-0.5	0.5	0.5

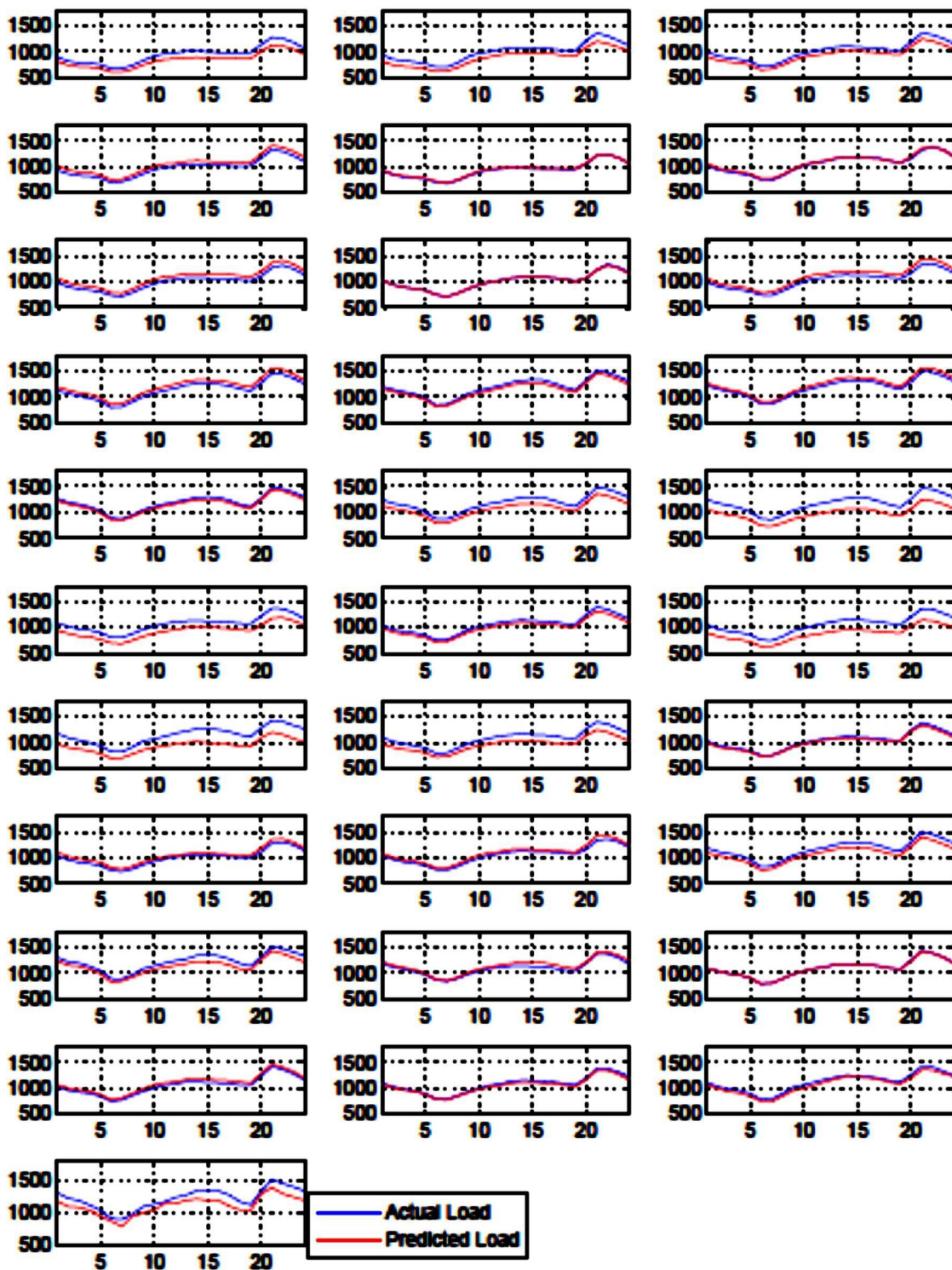
۲-۲-۵- انتخاب ورودی با روش PCA

۳-۲-۵- انتخاب ورودی با روش LLE
انتخاب ورودی توسط LLE مشابه روش PCA است. ابعاد فضای ورودی برای داده‌های آموزش و آزمون به ترتیب ۱۶۸×۲۴۶ و ۱۶۸×۳۱ است. مقدار d و k را به ترتیب برابر پانزده و پنجاه تنظیم می‌کنیم. بنابراین تعداد ورودی‌ها از ۱۶۸ تا پانزده کاهش می‌یابد. نتایج پیش‌بینی در شکل (۹) نشان داده شده است.

برای انتخاب ورودی با استفاده از روش PCA، ویژگی‌ها از بار هفت روز گذشته که شامل ۱۶۸ ویژگی هستند، استخراج شده است. با توجه به تعداد روزهای داده‌های آموزش و داده‌های آزمون، ابعاد ماتریس R در رابطه (۱) به ترتیب ۱۶۸×۲۴۶ و ۱۶۸×۳۱ برای داده‌های آموزش و داده‌های آزمون خواهد بود. پس از محاسبه PCA، ۱۵ بردار ویژه نخست را براساس رابطه (۲) انتخاب می‌کنیم. به این ترتیب تعداد ورودی‌ها به پانزده کاهش می‌یابد. نتایج پیش‌بینی در شکل (۸) نشان داده شده است.

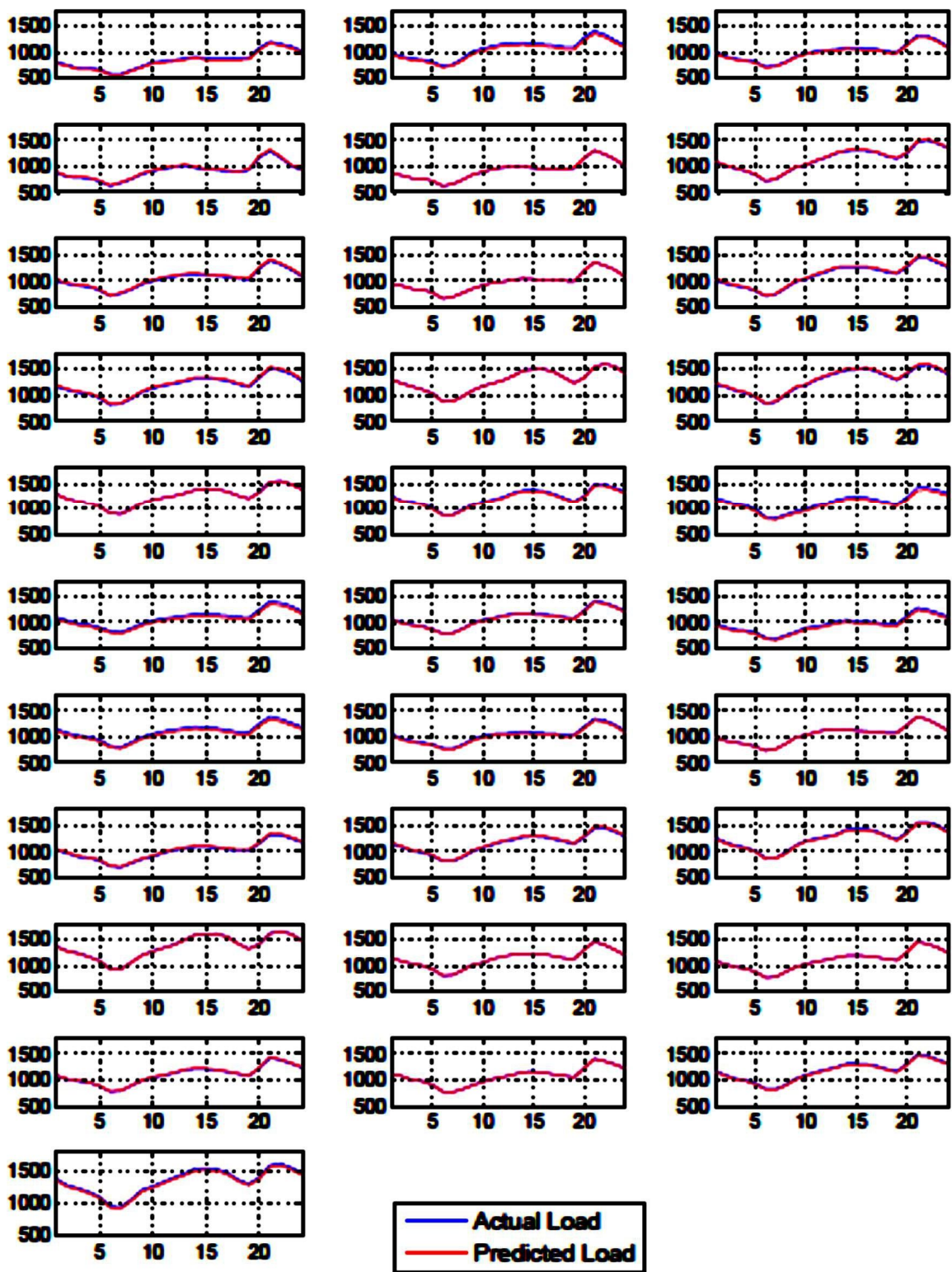


(شکل-۷): مقایسه بین منحنی بار واقعی و منحنی بار پیش‌بینی شده (مقدار بار روزانه بر حسب ساعت) برای دوره یک‌ماهه، زمانی که انتخاب ورودی به صورت تجربی انجام می‌شود و پیش‌بینی توسط شبکه Fuzzy-RBF انجام می‌شود
 (Figure-7): comparison between actual load curves and predicted load curves when input selection is carried out empirically and prediction is carried out by RBF network-fuzzy system.



(شکل-۸): مقایسه بین منحنی بار واقعی و منحنی بار پیش بینی شده (مقدار بار روزانه بر حسب ساعت) برای دوره یک ماهه، بر اساس انتخاب ورودی به روش PCA و پیش بینی توسط شبکه Fuzzy-RBF

(Figure-8): comparison between the actual load curves and the predicted load curves based on PCA input selection and prediction by RBF-fuzzy system.



(شکل-۹): مقایسه منحنی بار واقعی و منحنی بار پیش‌بینی شده (مقدار بار روزانه بر حسب ساعت) برای دوره یک‌ماهه، بر اساس انتخاب ورودی به کمک LLE و پیش‌بینی با شبکه Fuzzy-RBF

(Figure-9): comparing actual load curves and predicted load curves based on LLE input selection and RBF-fuzzy prediction system.

پیشنهادی در نظر گرفته شده است، که عبارتند از شبکه‌های عصبی مصنوعی، سری‌های زمانی و روش‌های نوروفازی. برای شبکه‌های عصبی مصنوعی، یک شبکه عصبی با دو لایه و ۲۴ نورون در لایه میانی انتخاب شده است. ورودی‌ها تاریخچه بار و بیشینه دمای روز هفته گذشته، روز قبل و بیشینه دمای تخمین‌زده‌شده روز پیش‌بینی است. از میان روش‌های مختلف سری زمانی، جهت سادگی و مقابله با ورودی خارجی (دما در مطالعه ما)، مدل ARX انتخاب و همچنین ANFIS به‌عنوان یک روش معمول نوروفازی برای پیش‌بینی بار استفاده شده است. در ANFIS، کدهای روز هفته، تاریخچه بار و تاریخچه درجه حرارت، به‌عنوان ورودی شبکه در نظر گرفته شده و مقایسه جریمه در این روش‌ها در جدول (۳) آمده است.

(جدول-۳): مقایسه عملکرد روش‌های پیشنهادی
(Table-3): Comparing the performance of proposed approaches

	Average MAPE	Maximum MAPE
Fuzzy-RBF	4.12%	9.15%
RBF	5.95%	10.18%
ANN	5.869%	12.91%
ARX	4.85%	10.03%
ANFIS	5.60%	13.94%

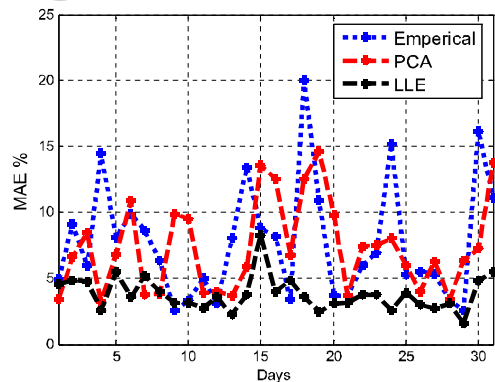
نتایج نشان می‌دهد که شبکه RBF-Fuzzy نسبت به سایر روش‌ها، MAPE پایین‌تری دارد. همچنین روش‌های ARX، ANN و ANFIS خطاهای کمتری نسبت به روش RBF (با استفاده از درجه حرارت به‌عنوان ورودی) کسب کرده‌اند. این نتایج تأیید می‌کنند که تأثیر دما بر نوسانات بار مازندران، بسیار زیاد است.

۵-۵- برتری آماری

برای سنجش اهمیت آماری روش پیشنهادی، روش T-Test مورد استفاده قرار گرفته است. برای تعیین اینکه آیا تفاوت بین دو میانگین، بیشتر از حد مورد انتظار از تصادف و شانس است، آزمون T می‌تواند به‌عنوان یک روش آماری، کارساز باشد. آزمون T، مقایسه میانگین دو گروه است، که به تفاوت تعداد و واریانس اعضای دو گروه توجه دارد. فرضیه آماری برای آزمون T به‌عنوان فرضیه نول در مورد تفاوت‌ها بیان می‌شود. برای آزمون فرضیه، مقدار محاسبه‌شده با مقدار بحرانی برای سطح اهمیت انتخابی، مقایسه می‌شود. اگر مقدار T محاسبه‌شده بزرگ‌تر یا مساوی مقدار بحرانی T

۳-۵- نتایج مقایسه‌ای

مقایسه MAPE در روش LLE، PCA و روش تجربی در شکل (۱۰) نشان داده شده است. همان‌طور که در شکل (۱۰) دیده می‌شود، بیشترین خطاهای پیش‌بینی برای انتخاب ورودی تجربی در روزهای ۴، ۱۴، ۱۸ و ۲۴ دوره پیش‌بینی رخ داده است. میانگین خطای پیش‌بینی در این دوره یک‌ماهه ۶/۷۳ درصد است. بیشینه خطاهای پیش‌بینی برای روش PCA در روزهای ۳ و ۶ دوره پیش‌بینی، رخ داده است. میانگین قدر مطلق خطای پیش‌بینی برای PCA در دوره زمانی یک‌ماهه ۵،۵۲ درصد است. این مقایسه نشان می‌دهد که خطای پیش‌بینی، هنگام اعمال PCA به فضای ورودی، کاهش می‌یابد. با استفاده از روش LLE بیشترین خطای پیش‌بینی در روز بیست‌ونجم از دوره پیش‌بینی رخ داده است. خطای متوسط در این دوره زمانی یک ماهه ۴،۱۲ درصد است. بنابراین، هنگام استفاده از روش LLE برای کاهش ابعاد فضای ورودی، خطای پیش‌بینی کمتر از روش PCA تجربی است. علاوه‌براین، مشاهده می‌شود که در تعطیلات ۲۸ خرداد (۱۸ ژوئن)، روش LLE نسبت به PCA و روش انتخاب ورودی تجربی، عملکرد بهتری از نظر کاهش خطای پیش‌بینی داشته است.



(شکل-۱۰): مقایسه میانگین قدر مطلق خطای پیش‌بینی در هر روز از دوره پیش‌بینی، زمانی که انتخاب ورودی به‌صورت تجربی، PCA و LLE انجام می‌شود.

(Figure-10): Comparing average absolute prediction error in each day of prediction, when input selection is carried out empirically, based on PCA and LLE methods.

۴-۵- مقایسه بین موتور پیش‌بینی و روش‌های معیار

در این بخش چندین مطالعه تطبیقی برای ارزیابی کارایی روش پیشنهادی ارائه شده است. برای این منظور، سه خانواده اصلی روش‌های پیش‌بینی، برای مقایسه با روش

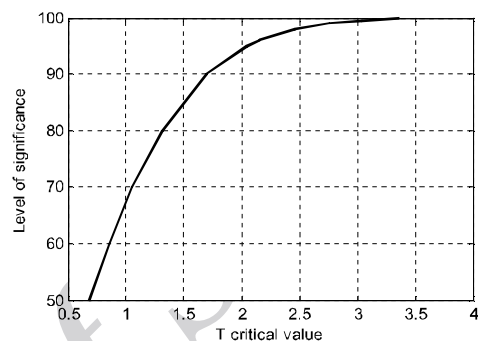
ورودی پیشنهاد شده با یک موتور پیش‌بینی ترکیبی همراه شده و عملکرد سامانه با استفاده از یک مطالعه جامع در مورد بار منطقه مازندران، ارزیابی شده است. روش LLE با روش PCA و روش انتخاب ورودی تجربی و موتور پیش‌بینی نیز با چند ابزار شناخته‌شده پیش‌بینی، مقایسه شده و نشان داده شده که موتور پیشنهادی از روش‌های شناخته‌شده مورد مقایسه قرار گرفته، عملکرد بهتری داشته است؛ همچنین، روش LLE به روش PCA و روش انتخاب ورودی تجربی، از نظر خطای پیش‌بینی، برتری دارد.

7- References

۷- مراجع

- [1] D. Liang, M. Zhichun, "Short-term load forecasting based on fuzzy neural network," *Journal of University of Science and Technology Beijing*, 4, pp. 46-49, 1997.
- [2] K.H Kim, H.S Youn, Y.C. Kang, "Short-term Load Forecasting for Special Days in anomalous Load Conditions Using Neural Network and Fuzzy Inference Method," *IEEE Transactions on Power Systems*, vol.15, pp. 559-569, 2000.
- [3] W. Charytoniuk, M. S. Chen, "Neural Network design for Short-Term Load Forecasting," *Proceedings of International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, London, pp. 4-7, 2000.
- [4] Z. Tao, Z. Dengfu, Z. Lin, W. Xifan, X. Daozhi, "Short-Term Load Forecasting Using Radial Basis Function Networks and Expert system," *Journal of XT'AN JIAOTONG University*, vol.35, pp. 331-334, 2001.
- [5] Z. Xin, C. Tian-Lun, "Nonlinear Time Series Forecast Using Radial Basis Function Neural Network," *Commun. Theor. Phys*, vol.40, pp.165-168, 2003.
- [6] V. S. Kodogiannis, E. M. Anagnostakis, "Soft computing based techniques for short-term load forecasting," *Fuzzy Sets and Systems* 128(3), pp. 413-426, 2002.
- [7] R.R.B. de Aquino, et.al, "Combined Artificial Neural Network and Adaptive Neuro-Fuzzy Inference System for Improving a Short-Term Electric Load Forecasting," *Lecture Notes in Computer Science*, 4669, 779-788, 2007.
- [8] Musa, Abdallah Bashir. "A comparison of ℓ_1 -regularization, PCA, KPCA and ICA for dimensionality reduction in logistic regression." *International Journal of Machine Learning and Cybernetics* 5.6, pp.861-873, 2014.
- [9] L. Cayton, *Algorithms for manifold learning*, University of California, San Diego, Tech. Report, 2005.

باشد، فرضیه نول رد شده و نتیجه گرفته می‌شود که میزان تفاوت حاصل‌شده، بیشتر از تفاوت مورد انتظار از تصادف و شانس است. اگر مقدار T محاسبه‌شده کوچکتر از مقدار بحرانی T باشد، فرضیه نول رد نشده و نتیجه می‌شود که این مقدار اختلاف، می‌تواند نتیجه شانس و تصادف باشد [21]. مقادیر بحرانی T برای مقادیر مختلف سطوح اهمیت در شکل (۱۱) نشان داده شده است. برای این آزمون، ۳۱ روز بازه مورد مطالعه، استفاده شده است. سطح اهمیت روش پیشنهادی Fuzzy-RBF از نظر MAPE با سایر روش‌ها مقایسه شده است. سطوح اهمیت در جدول (۴) آمده است.



(شکل-۱۱): نمودار بحرانی مقادیر T در سطوح مختلف اهمیت (Figure-11): The critical diagram of "T" values in different levels of significance.

(جدول-۴): سطح اهمیت Fuzzy-RBF در مقایسه با روشهای

دیگر از نقطه نظر MAPE

(Table-4): Level of significance of Fuzzy-RBF in comparison with other methods (In sense of MAPE)

	RBF	ANN	ARX	Δ NFIS
Fuzzy-RBF	0.96	0.94	0.906	0.954

۶- نتیجه‌گیری

ورودی‌های موتور پیش‌بینی، پایه تخمین است و موتور پیش‌بینی، ورودی‌ها را به خروجی‌های از پیش تعیین‌شده، تصویر می‌کند. وابستگی متقابل ویژگی‌ها در فضای ورودی، منجر به افزایش ابعاد ورودی، بدون افزایش اطلاعات مؤثر، می‌شود، که منجر به عدم کارایی سامانه پیش‌بینی می‌شود. در فضای ورودی، ممکن است، رابطه غیرخطی بین ویژگی‌ها وجود داشته باشد و تبدیل‌های خطی مانند PCA نمی‌توانند چنین وابستگی‌های غیر خطی را پیدا کنند. در این مقاله نشان داده شده است که روش‌های یادگیری غیرخطی خم مانند LLE، فضای گسترده‌تر را برای استخراج ویژگی‌ها جستجو می‌کنند و بنابراین می‌توانند ورودی‌های مناسب‌تری را برای پیش‌بینی بار پیدا کنند. روش انتخاب

- load and neural networks," *Journal of Modeling in Engineering*, vol.15 (16), 2007.
- [۲۳] سید شنوا سید جلال، قاسمی علی، شایقی حسین و نوشیار مهدی. ارایه یک مدل ترکیبی در پیش‌بینی بار در بازار برق تجدید ساختار یافته. نشریه علمی پژوهشی کیفیت و بهره‌وری صنعت برق ایران، سال دوم شماره سوم، بهار و تابستان ۹۲
- [23] S.sh. Seyed Jalal, et al, "A Composite Model for Load Forecasting in the Restructured Electricity Market," *Journal of Quality and Productivity in Iran Electric Industry*, vol.2(3), 2013.
- [۲۴] مشاری امیر، ابراهیمی اکبر، صدری سعید، ابراهیمی محمد. پالایش داده‌های آموزشی شبکه عصبی و بررسی تأثیر آن در کاهش خطای پیش‌بینی کوتاه مدت بار سیستم‌های قدرت. نشریه استقلال. سال ۲۸. شماره ۲. اسفند ۸۸
- [24] A. Moshari, et al., "Purification of Neural Network Train Data and its Effect on Reducing Short-Term Prediction Errors of Power Systems," *Esteghlal Magazine*. vol.28 (2), 2009.
- [۲۵] شایقی حسین، قاسمی علی. مدل‌سازی سیستم چند ورودی چند خروجی برای پیش‌بینی همزمان قیمت و بار در شبکه هوشمند با اعمال مدیریت بار. نشریه هوش محاسباتی در مهندسی برق. سال ششم، شماره چهارم، زمستان ۹۴.
- [25] H. Shayeghi, et al, "Multi-Input Multi-output System Modeling for simultaneous prediction of price and load in the smart grid by applying load management," *Journal of Computational Intelligence in Electrical Engineering*, Vol. 6 (4), 2015.
- [۲۶] کریمی مازیار، کریمی حسین، غلامی مصطفی، خطیب زاده هادی، مسلمی نیکی. اولویت‌بندی روزهای مشابه جهت پیش‌بینی بار کوتاه مدت شبکه ایران با در نظر گیری دما و بخش‌بندی سیستم قدرت. مجله انجمن مهندسی برق و الکترونیک ایران، سال چهاردهم، شماره سوم، پاییز ۹۶
- [26] M. Karimi , et al., "Prioritizing the same days to predict the short-term load of Iran's network by considering the temperature and power system segmentation," *Journal of Iranian Association of Electrical and Electronics Engineers*. Vol. 14 (3), 2016.
- [۲۷] قانعی یخدان حسین. روشی جدید برای اختفای خطا در فریم‌های ویدئو با استفاده از شبکه عصبی RBF.
- [10] I. Borg and P. Groenen, *Modern Multidimensional Scaling: Theory and Applications*. New York: Springer-Verlag, 1997.
- [11] J. B. Kruskal, "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis," *Psychometrika*, vol.29, pp. 1-27, 1964.
- [12] Ji, Rongrong, et al, "Towards Optimal Manifold Hashing via Discrete Locally Linear Embedding," *IEEE Transactions on Image Processing*, 2017.
- [13] Liu, Xin, et al, "Locally linear embedding (LLE) for MRI based Alzheimer's disease classification," *Neuroimage*, vol. 83, pp.148-157.2013.
- [14] J. Yang, B. Ming Xiang, and Y. Zhang, "Multi-manifold Discriminant Isomap for visualization and classification," *Pattern Recognition*, vol. 55, pp. 215-230, 2016.
- [15] M. Belkin ,and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Compute*, vol. 15(6), pp. 1373-1396, 2003.
- [16] Ye. Qiang, and W. Zhi, "Discrete hessian eigenmaps method for dimensionality reduction," *Journal of Computational and Applied Mathematics*, vol.278, pp.197-212, 2015.
- [17] Su, Zuqiang, et al., "Fault diagnosis method using supervised extended local tangent space alignment for dimension reduction," *Measurement*, vol. 62, pp.1-14, 2015.
- [18] L. Haghverdi, F. Buettner, and J. Fabian, "Diffusion maps for high-dimensional single-cell analysis of differentiation data," *Bioinformatics*, vol. 31.18, pp.2989-2998, 2015.
- [19] Lunga, Dalton, et al. "Manifold-learning-based feature extraction for classification of hyperspectral data: A review of advances in manifold learning," *IEEE Signal Processing Magazine*, vol. 31.1, pp.55-66, 2014.
- [20] J. Wang, Z. Zhang, and H. Zha, "Adaptive manifold learning," *In Advances in Neural Information Processing Systems*, 2004.
- [21] W. H Press, et al, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, pp. 616, 1992.
- [۲۲] فریبرز نیا روح الله، امجدی نیما، پیش‌بینی بار کوتاه‌مدت با استفاده از تجزیه سری زمانی بار و شبکه عصبی، مجله مدل‌سازی در مهندسی، دوره ۲، شماره ۱۶، بهار ۸۷
- [22] R. Fariborznia, N. Amjadi, "short term load forecast by using time series decomposition of

فصلنامه علمی - پژوهشی پردازش علائم و داده‌ها،
جلد دهم، شماره ۱، آذر ۹۲

نشانی رایانامه ایشان عبارت است از:
h-kamalinejad@trec.co.ir



بابک نجار اعرابی دانش‌آموخته دکترای
تخصصی رشته مهندسی برق با گرایش
پردازش سیگنال و تصویر از دانشگاه
Texas A&M آمریکا است. در حال

حاضر ایشان استاد گروه کنترل و نیز
هوش ماشین و رباتیک، در دانشکده برق و کامپیوتر دانشگاه
تهران و پژوهش‌گر غیرمقیم پژوهشکده علوم شناختی در
پژوهشگاه دانش‌های بنیادی است. زمینه‌های پژوهشی مورد
علاقه ایشان شامل بازشناخت الگو، بینایی ماشین، یادگیری
ماشین، تصمیم‌گیری در حضور عدم قطعیت، سامانه‌های
نوروفازی، تشخیص خطا، پیش‌بینی و مدل‌سازی سامانه‌ها
است. وی عضو مؤسس انجمن ایرانیان علوم شناختی و
تکنولوژی و عضو هیأت مدیره چندین انجمن علمی است.
نشانی رایانامه ایشان عبارت است از:

araabi@ut.ac.ir

[27] H. Ghanei yakhdan, "A new method to hide errors in video frames using RBF neural network," *Journal of Signal and Data Processing*, vol.10 (1), 2013.

[۲۸] گوهریان نازنین، مقیمی سحر، غلامی مصطفی، کلانی هادی. استفاده از ترکیب الگوریتم ژنتیک و شبکه‌های عصبی مصنوعی برای پیش‌بینی نیروی گاز گرفتن از روی سیگنال الکترومایوگرام. فصل‌نامه علمی-پژوهشی پردازش علائم و داده‌ها، جلد چهاردهم، شماره ۱، تیر ماه ۹۶.

[28] N. Goharian, et al., "Use of combination of genetic algorithm and artificial neural networks to predict the bite force from an electromyogram signal," *Journal of Signal and Data Processing*, Vol.14 (1), 2017.



حامد کبریائی دانش‌آموخته مقطع دکترا

در رشته مهندسی برق کنترل از دانشگاه تهران و در حال حاضر استادیار گروه کنترل در دانشکده برق و کامپیوتر دانشگاه تهران است. زمینه‌های پژوهشی

مورد علاقه ایشان شامل تئوری بازی‌ها، بهینه‌سازی و کنترل تصادفی و کاربرد آن در مخابرات بی‌سیم و شبکه، و شبکه‌های الکتریکی هوشمند است. وی عضو IEEE و کمیته کنترل IEEE بوده و مقالاتی را در چندین مجله مجموعه IEEE و مجلات دیگر داوری کرده‌اند.

نشانی رایانامه ایشان عبارت است از:

kebriaei@ut.ac.ir



حوراء کمالی نژاد دانش‌آموخته

کارشناسی ارشد رشته مهندسی برق کنترل از دانشگاه تربیت مدرس است. در حال حاضر ایشان کارشناس ارشد مطالعات فنی بازار برق در دفتر بازار برق

شرکت برق منطقه‌ای تهران است. زمینه‌های پژوهشی مورد علاقه ایشان مدیریت بار مشترکین صنعتی، پیش‌بینی بار الکتریکی کوتاه‌مدت، پیش‌بینی قیمت در بازار برق، کاربرد تئوری بازی‌ها در بازار برق و محیط‌های رقابتی و شبکه‌های الکتریکی هوشمند است.