

FORECASTING ENERGY CONSUMPTION USING FUZZY TRANSFORM AND LOCAL LINEAR NEURO FUZZY MODELS

Hossein Iranmanesh¹, Majid Abdollahzade² and Arash Miranian³

¹Department of Industrial Engineering, “University of Tehran” & “Institute for International Energy Studies”, Tehran, Iran
h.iranmanesh@ut.ac.ir

²Department of Mechanical Engineering, “K.N.Toosi University of Technology” & “Institute for International Energy Studies”, Tehran, Iran
m.abdollahzade@gmail.com

³School of Electrical and Computer Engineering, “University of Tehran” & “Institute for International Energy Studies”, Tehran, Iran
ar.miranian@gmail.com

ABSTRACT

This paper proposes a hybrid approach based on local linear neuro fuzzy (LLNF) model and fuzzy transform (F-transform), termed FT-LLNF, for prediction of energy consumption. LLNF models are powerful in modelling and forecasting highly nonlinear and complex time series. Starting from an optimal linear least square model, they add nonlinear neurons to the initial model as long as the model's accuracy is improved. Trained by local linear model tree learning (LOLIMOT) algorithm, the LLNF models provide maximum generalizability as well as the outstanding performance. Besides, the recently introduced technique of fuzzy transform (F-transform) is employed as a time series pre-processing method. The technique of F-transform, established based on the concept of fuzzy partitions, eliminates noisy variations of the original time series and results in a well-behaved series which can be predicted with higher accuracy. The proposed hybrid method of FT-LLNF is applied to prediction of energy consumption in the United States and Canada. The prediction results and comparison to optimized multi-layer perceptron (MLP) models and the LLNF itself, revealed the promising performance of the proposed approach for energy consumption prediction and its potential usage for real world applications.

KEYWORDS

LLNF, LOLIMOT, F-transform, energy consumption, forecasting

1. INTRODUCTION

The increasing trend in energy consumption world-wide necessitates appropriate investments on energy sectors by governments. For instance, consumption of petroleum in the United States, as one of the world's largest energy consumers, has increased by 20.4% from 1990 to 2005. U.S. natural gas consumption has also experienced 16.32% increase within the same period [1]. Hence, providing accurate forecasts of energy consumptions is of crucial importance for policy-makers in large-scale decision makings, such as investment planning for generation and distribution of energy.

Different methods, including time series and computational intelligence (CI) -based approaches have been proposed for energy consumption prediction, to date. Auto-regressive integrated moving average (ARIMA) model, as one of the most popular time series methods, has been used for energy consumption prediction to a great extent. For instance, Haris and Liu proposed ARIMA and transfer function models for prediction of electricity consumption [2]. ARIMA and seasonal ARIMA (SARIMA) were used by Ediger and Akar to estimate the future primary energy consumption of Turkey from 2005 to 2020 [3]. Linear regression models have also been proposed for energy consumption prediction [4]. Other time series approaches such as Grey models and Markov models have been developed for forecasting energy consumption [5]-[7].

Numerous researches have been carried out in the past decade on developing CI-based methods for energy consumption prediction. In contrast to time series approaches which are inherently linear, CI-based methods can effectively model the nonlinear behaviour of time series. Fuzzy logic and neural networks (NN) are two main CI-based techniques which have found many applications in function approximation, simulation, modelling and prediction [8], [9]. Energy consumption modelling and prediction are also among the applications of the CI-based approaches. Long-term prediction of gasoline demand in several countries by multi-layer perceptron (MLP) network has been carried out by Azadeh et al. [10]. In another study, Azadeh et al. proposed a fuzzy regression algorithm for oil demand estimation of the U.S., Canada, Japan and Australia [11]. A hybrid technique of well-known Takagi-Sugeno fuzzy inference system and fuzzy regression has been proposed for prediction of short-term electric demand variations by Shakouri et al. [12]. In their study, Shakouri et al. introduced a type III TSK fuzzy inference machine combined with a set of linear and nonlinear fuzzy regressors in the consequent part to model effects of the climate change on the electricity demand. Yokoyama et al. concentrated on neural networks to predict energy demands [13]. They used an optimization technique, termed Modal Trimming Method to optimize model's parameters and then forecasted cooling demand in buildings. Neuro-fuzzy models, which are combination of fuzzy logic and neural networks, have also been proposed for energy consumption prediction. For instance, Chen proposed a fuzzy-neural approach for long-term prediction of electric energy in Taiwan [14]. In this study, Chan proposed multiple experts which construct their own fuzzy back propagation networks from various viewpoints to forecast the long-term electric energy consumption. Chen used fuzzy intersection to aggregate these long-term load forecasts and then constructed a radial basis function network to defuzzify the aggregation result and to generate a representative/crisp values. As another example, short-term prediction of natural gas demand has been carried out using Adaptive neuro-fuzzy inference system (ANFIS) in [15]. Azadeh et al. in [15] used ANFIS approach with pre-processing and post-processing concepts to improve accuracy of predictions. Nostrati et al. also proposed a neuro-fuzzy model for long-term electrical load forecasting. [16]. Various traditional and CI-based models for energy demand forecasting have been also reviewed in [17].

In this paper, we employ the technique of fuzzy transform and LLNF model for prediction of energy consumption. The recently introduced concept of fuzzy transform will be used for data pre-processing and denoising the consumption series in order to obtain a well-behaved time series. Then the filtered time series is used by the LLNF model to forecast future values of consumption. The developed FT-LLNF approach will be applied to prediction of annual oil consumption in the U.S. and Canada as well as monthly consumption of natural gas in the U.S. The results of prediction demonstrate the effectiveness of the proposed method for energy consumption prediction.

2. LOCAL LINEAR NEURO FUZZY MODEL

The local linear neuro fuzzy approach based on the incremental tree based learning algorithm starts from an initial optimal linear model and increases the complexity of the model as long as improvements occurs. The local linear model tree (LOLIMOT) learning algorithm, employed in this paper, applies axis-orthogonal splits to divide the original input domain into sub-domains,

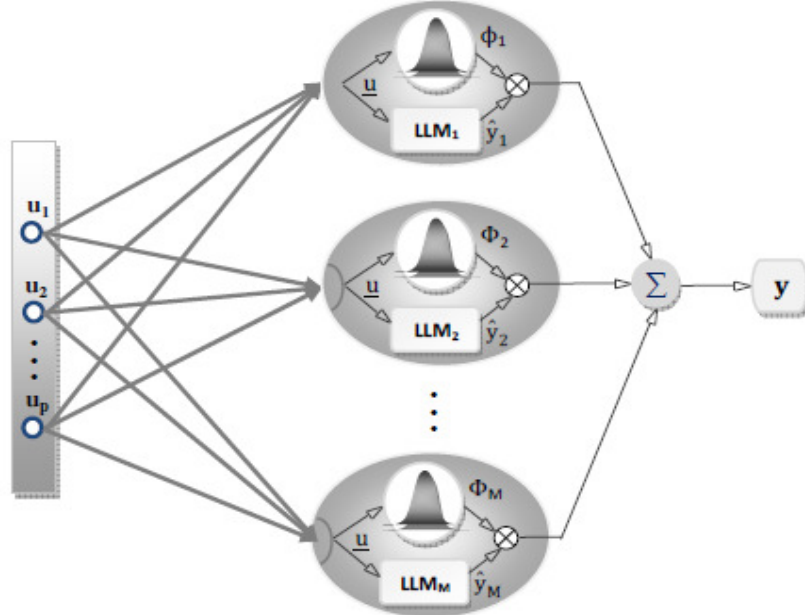


Fig. 1. Structure of LLNF

each identified by a local linear model (LLM) and its associated validity function. Thus, it is inferred that the LLNF models use the divide-and-conquer approach to solve a complex modelling problem by decomposing it into smaller and simpler sub-problems [18]. The general structure of the LLNF model for a p -dimensional input space and M local linear models is illustrated in fig. 1.

The global output of the LLNF model can be stated as follows,

$$\hat{y} = \sum_{i=1}^M \hat{y}_i \phi_i(\underline{u}) \quad (1)$$

where, y_i and ϕ_j are the output and normalized validity function of the LLM_j , respectively and $\underline{u} = [u_1 \ u_2 \ \dots \ u_p]^T$ is the input vector.

The local output of each LLM is found form the linear estimation, below

$$\hat{y}_i = \theta_j^T \underline{u}, j=1, \dots \quad (2)$$

where, θ_j is the vector of parameters for LLM_j .

The LLNF model offers a transparent representation of the nonlinear system. In other words, the LLNF model divides a complex model into a set of sub-models which perform linearly and independently.

Two types of parameters, i.e. rule consequent and rule premise parameters must be identified in the LLNF model. The rule consequent parameters (θ_{ij}), associated with local linear models, are estimated by linear least-square optimization procedure. Both global and local least square estimations can be applied. In this paper, the former is employed since it considers the interaction

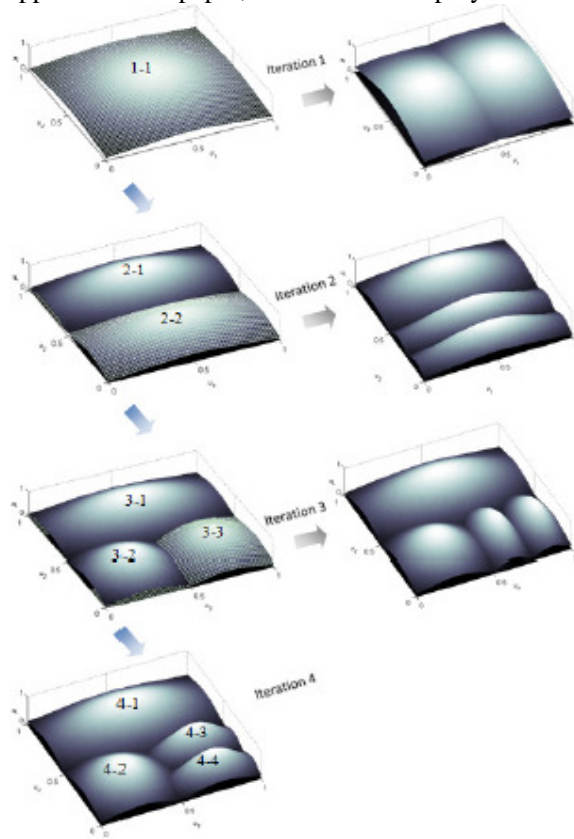


Fig. 2. Operation of LOLIMOT in the first four iterations in a two-dimensional input space

between LLMs to a greater extent and therefore show better performance [18].

For an LLNF model with M neurons and p inputs, the vector of linear parameters contains $n = M(p+1)$ elements,

$$\underline{\theta} = [\theta_{10} \theta_{11} \dots \theta_{1p} \theta_{20} \theta_{21} \dots \theta_{2p} \theta_{M0} \theta_{M1} \dots \theta_{Mp}]^T \quad (3)$$

The corresponding regression matrix for N measured data samples is,

$$\underline{X} = [\underline{X}_1 \ \underline{X}_2 \ \dots \ \underline{X}_M] \quad (4)$$

where regression sub-matrix \underline{X}_i takes the following form,

$$\begin{bmatrix} \phi_i(\underline{u}(1)) & u_1(1)\phi_i(\underline{u}(1)) & \dots & u_p(1)\phi_i(\underline{u}(1)) \\ \phi_i(\underline{u}(2)) & u_1(2)\phi_i(\underline{u}(2)) & \dots & u_p(2)\phi_i(\underline{u}(2)) \\ \vdots & \vdots & & \vdots \\ \phi_i(\underline{u}(N)) & u_1(N)\phi_i(\underline{u}(N)) & \dots & u_p(N)\phi_i(\underline{u}(N)) \end{bmatrix} \quad (5)$$

Hence,

$$\hat{\underline{y}} = \underline{X}\hat{\underline{\theta}}; \quad \hat{\underline{\theta}} = (\underline{X}^T \underline{X})^{-1} \underline{X}^T \underline{y} \quad (6)$$

where, $\underline{y} = [y(1) y(2) \dots y(N)]^T$ contains the measured outputs.

The LOLIMOT algorithm is used for estimation of validity functions parameters. This algorithm is fast in convergence and computationally efficient and therefore preferred over other optimization methods such as genetic algorithms and simulated annealing. LOLIMOT algorithm utilizes multivariate normalized axis-orthogonal Gaussian membership functions, as stated below.

$$\mu_i(\underline{u}) = \exp\left(-\frac{1}{2}\left(\frac{(u_1 - c_{i1})^2}{\sigma_{i1}^2} + \dots + \frac{(u_p - c_{ip})^2}{\sigma_{ip}^2}\right)\right) \quad (7)$$

$$= \exp\left(-\frac{1}{2}\frac{(u_1 - c_{i1})^2}{\sigma_{i1}^2}\right) \dots \exp\left(-\frac{1}{2}\frac{(u_1 - c_{i1})^2}{\sigma_{i1}^2}\right)$$

$$\phi_i(\underline{u}) = \frac{\mu_i(\underline{u})}{\sum_{j=1}^M \mu_j(\underline{u})} \quad (8)$$

where c_{ij} and σ_{ij} represent centre coordinate and standard deviation of normalized Gaussian validity function associated with its local linear model.

In LOLIMOT algorithm, the input space is divided into hyper-rectangles by axis-orthogonal cuts based on a tree structure. Each hyper-rectangle represents an LLM. In the original LOLIMOT algorithm, the LLM with worst performance is divided into two halves at each iteration. Then Gaussian membership functions are placed at the centres of the hyper-rectangles and standard deviations are selected proportional to the extension of hyper-rectangles (usually 1/3 of hyper-rectangle's extension). Starting from an initial optimal linear model, the LOLIMOT algorithm adds nonlinear neurons provided that the model's performance is enhanced. Therefore, a model with the highest generalization is achieved.

A graphical representation of partitioning a two-dimensional input space by LOLMOT up to the first four iterations is illustrated by fig. 2.

3. FUZZY TRANSFORM

The new technique of fuzzy transform (F-transform) has been developed in recent years for function approximation applications [19]. However, image compression, denoising and time series analysis are other applications of the F-transform [20]-[21]. In this paper we employ F-transform due to its denoising properties for pre-processing of the input data in prediction of energy consumption time series. In this section, the concept of fuzzy partition in $[a,b]$ is defined and then the direct and inverse F-transform are introduced. It must be noted that here in this we

focus only on discrete F-transform and the term discrete is omitted through the rest of the paper for convenience.

Definition 1 [19]: Let $x_1 < \dots < x_n$ be fixed nodes within $[a, b]$, such that $x_1 = a, x_n = b$ and $n \geq 2$. We say that fuzzy sets A_1, \dots, A_n identified with their membership functions $A_1(x), \dots, A_n(x)$ defined on $[a, b]$, form a fuzzy partition of $[a, b]$ if they fulfill the following conditions for $k = 1, \dots, n$:

- (1) $A_k : [a, b] \rightarrow [0, 1], A_k(x_k) = 1$;
- (2) $A_k(x) = 0$ if $x \notin (x_{k-1}, x_{k+1})$ where for the uniformity of denotation, we put $x_0 = a$ and $x_{n+1} = b$;
- (3) $A_k(x)$ is continuous;
- (4) $A_k(x), k = 2, \dots, n$, strictly increases on $[x_{k-1}, x_k]$ and $A_k(x), k = 1, \dots, n-1$, strictly decreases on $[x_k, x_{k+1}]$;
- (5) for all $x \in [a, b], \sum_{k=1}^n A_k(x) = 1$

The membership functions A_1, \dots, A_n are called basic functions.

It must be noted that the basic functions are determined by the set of nodes $x_1 < \dots < x_n$ and properties 1-5, and their shape is not pre-specified. Therefore, the shape of basic functions can be determined according to additional requirements. Fig. 3 show an example of fuzzy partitions and by triangular membership functions in [1, 4]. The following formulae describe the formal representation of such triangular membership functions in $[x_1, x_n]$,

$$A_1(x) = \begin{cases} 1 - \frac{(x-x_1)}{h_1}, & x \in [x_1, x_2] \\ 0 & \text{otherwise} \end{cases}$$

$$A_k(x) = \begin{cases} \frac{(x-x_{k-1})}{h_k}, & x \in [x_{k-1}, x_k] \\ 1 - \frac{(x-x_k)}{h_k}, & x \in [x_k, x_{k+1}] \\ 0 & \text{otherwise} \end{cases}$$

$$A_n(x) = \begin{cases} \frac{(x-x_{n-1})}{h_{n-1}}, & x \in [x_{n-1}, x_n] \\ 0 & \text{otherwise} \end{cases}$$

where, $k = 1, \dots, n-1$ and $h_k = x_{k+1} - x_k$.

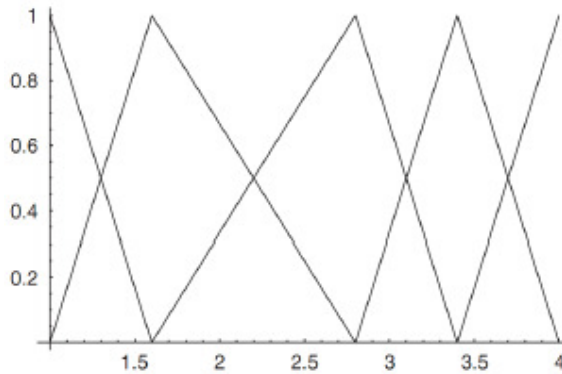


Fig. 3. A fuzzy partition with triangular membership functions

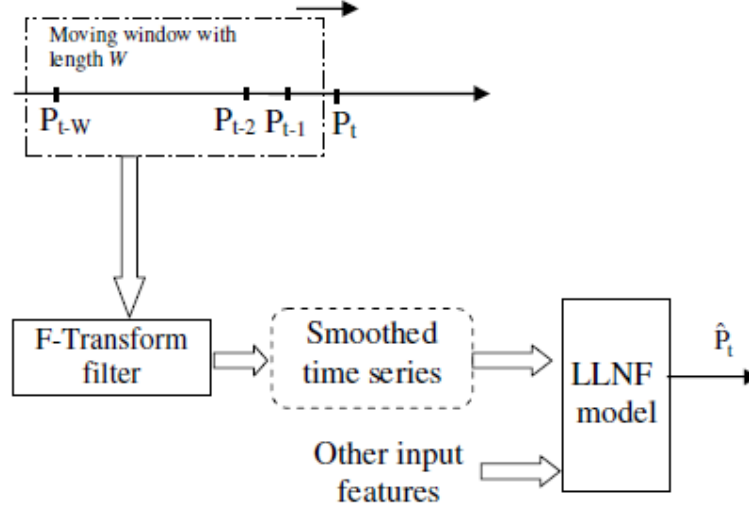


Fig. 4. Proposed forecast framework. Prediction of time series at period t , using a window of length W .

3.1 Discrete F-transform

The mathematical description of discrete F-transform and its inverse is presented in this section. The F-transform transforms an original function into an n -dimensional vector. The inverse F-transform constructs an approximation of the original function using the transformed n -dimensional vector. In fact, the F-transform presents a unique and simple representation of the original function which can be employed in complex computations [20]. The following definition introduces the F-transform.

Definition 2 [19]: Let A_1, \dots, A_n be basic functions which form a fuzzy partition of $[a, b]$ and f be a function given at nodes $p_1, \dots, p_l \in [a, b]$. We say that the n -tuple of real numbers $[F_1, \dots, F_n]$ is the F-transform of function f given by,

$$F_k = \frac{\sum_{j=1}^l f(p_j) A_k(p_j)}{\sum_{j=1}^l A_k(p_j)}, \quad k = 1, \dots, n \quad (9)$$

It's been shown in [19] that the components of F-transform are the weighted mean values of the original function and the weights are given by the basic functions.

Definition 3 [19]. Let A_1, \dots, A_n be basic functions which form a fuzzy partition of $[a, b]$ and f be a function given at nodes $p_1, \dots, p_l \in [a, b]$. Let $F_n(f) = [F_1, \dots, F_n]$ be the F-transform of function f with respect to A_1, \dots, A_n . Then the function

$$f_{F,n}(p_j) = \sum_{k=1}^n F_k A_k(p_j), \quad j = 1, \dots, l \quad (10)$$

is called the inverse F-transform of function f .

4. PROPOSED FORECAST FRAMEWORK

In this section, the procedure of time series denoising by means of F-transform is illustrated. Based on the proposed approach, the energy consumption series before being fed into the hybrid forecaster is processed by the F-transform. Such a procedure removes the high frequency components of the time series and produces a smooth version of the series. This is depicted in fig. 4. According to this fig., for prediction of time series at instant t , a window of time series with length W , prior to point t are filtered using F-transform. Therefore a smoothed and a residual series are resulted. The residual series contains high frequency variations of the original time series (and is considered as noise in this paper). The smoothed series which contain main information of the original time series is used as forecaster input. It's worth noting that a separate model can be developed for prediction of the residual series and then the final prediction can be produced by summing up predicted values of the residual and smoothed series. This will be considered in future works.

5. ENERGY CONSUMPTION PREDICTION

In this section the annual oil consumption of the United States and Canada and monthly gas consumption in the U.S. will be forecasted using the proposed FT-LLNF model. These data are collected from World Bank Development Indicator datasets and the U.S. energy information administration website [22], [23]. For evaluation of the performance of the FT-LLNF model, the Mean absolute percentage error (MAPE) and absolute percentage error (APE) are computed,

$$MAPE = \frac{100}{T} \sum_{t=1}^T \frac{|y_t - \hat{y}_t|}{y_t} \quad (15)$$

$$APE = \frac{|y_t - \hat{y}_t|}{y_t} \times 100 \quad (16)$$

where, y_t and \hat{y}_t are the actual and predicted consumptions at period t , respectively.

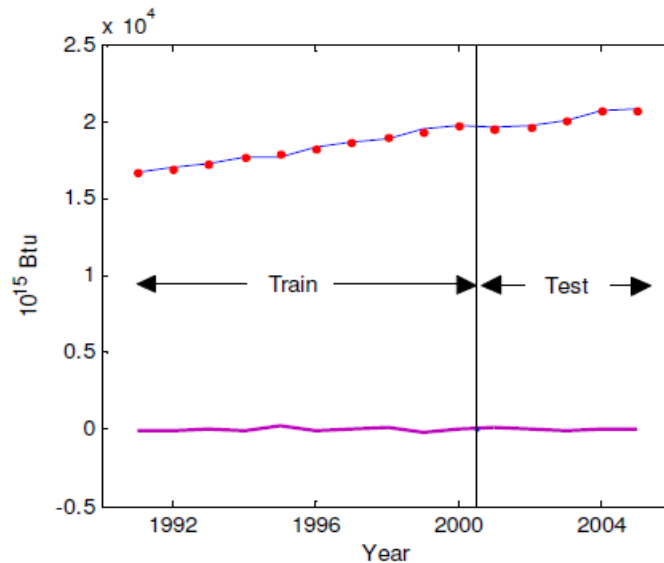


Fig. 5. Actual and predicted U.S. oil consumption. Blue: Actual, red dots: predictions, thick purple: error

Table1. Comprison of MAPE between different methods. The case of U.S. oil consumption

Method	Training	Test
MLP	1.65%	1.91%
LLNF	0.34%	0.54%
FT-LLNF	0.23%	0.35%

5.1. Prediction of Annual oil Consumption in the U.S. and Canada

Table 2. Actual and predicted values for U.S. oil consumption

Year	Actual	Forecast	APE
2001	19648.71	19542.98	0.54%
2002	19761.3	19685.53	0.38%
2003	20033.5	20060.39	0.13%
2004	20731.15	20668.15	0.30%
2005	20802.16	20724.01	0.37%

Our first case study focuses on prediction of annual oil consumption in the U.S. and Canada as two large consumer of crude oil in the world. For this purpose, crude oil consumption data from 1990 to 2005 are considered. The data points from 1990 to 2000 are selected for training the FT-LLNF model and the remaining 5 data samples are used for testing model performance. The standard input variables used in this study to construct the forecast model are annual population, cost of crude oil import, gross domestic product (GDP), annual oil production and annual oil consumption of the last period and the model's output is the annual oil consumption. Hence a multivariate time series prediction is performed. It must be noted that only oil consumption series are filtered using F-transform. Furthermore, test points are not subject to filtering. Consequently, for filtering consumption series using F-transform, 15 basic functions were used to form the fuzzy partition.

The predicted and actual oil consumptions of the U.S., for both training and test points are shown in fig. 5. Obviously the proposed FT-LLNF model has tracked the consumption series with remarkable accuracy. Furthermore, a numerical comparison between an optimized MLP network (constructed and trained by authors for the purpose of comparison) and the LLNF model is presented in table 1. The advantage of the proposed method over MLP is clear based on the numerical comparison in table 1. Furthermore, the effectiveness of the F-transform filter can be understood from this table. Interestingly, by integration of the F-transform filter into the hybrid model, the MAPE for test data has been reduced from 0.54% to 0.35% (almost 35% improvement). For a more detailed analysis on performance of the proposed method, the actual and forecasted values for test points are tabulated in table 2.

Similar predictions are performed for annual oil consumption in Canada. The results of these predictions are illustrated by fig. 6. Outstanding performance of the FT-LLNF model is demonstrated by this fig., again. The model has noticeable performance for training and test data as well as low prediction errors. Comparing the training and test MAPE of the FT-LLNF model with an optimized MLP network and the LLNF model, as presented in table 3, confirms the superiority of the proposed method and the noteworthy effect of the data pre-processing using F-

Table 3. Comparison of MAPE between different methods. The case of Canada oil consumption

Method	Training	Test
MLP	2.2%	3.1%
LLNF	0.88%	1.95%
FT-LLNF	0.63%	1.83%

Table 4. Actual and predicted values for Canada oil consumption

Year	Actual	Forecast	APE
2001	2056.84	2057.56	0.035%
2002	2078.35	2137.67	2.85%
2003	2207.15	2141.40	2.98%
2004	2299.67	2228.76	3.08%
2005	2296.86	2307.78	0.47%

transform. Furthermore, the actual and predicted values of the test data and the relative errors are shown in table 4.

5.2. Prediction of Monthly Natural Gas Consumption in the U.S.

In previous case studies annual energy consumptions were carried out and presented. In this case study we intend to demonstrate model performance on prediction of short term energy consumption. Therefore, we consider monthly gas consumption in the U.S. For this purpose, U.S. monthly consumption of natural gas from Jan. 1973 to Dec. 2007 are collected from [22]. In contrast to previous case studies, univariate time series prediction is carried out here; i.e., the input of hybrid forecaster is only the previous values of natural gas consumption. The data from Jan. 1973 to Dec 2009 are selected as training set. Hence, the natural gas consumption from Jan. 2008 to Dec. 2010 (i.e. three years) will be forecasted as test data.

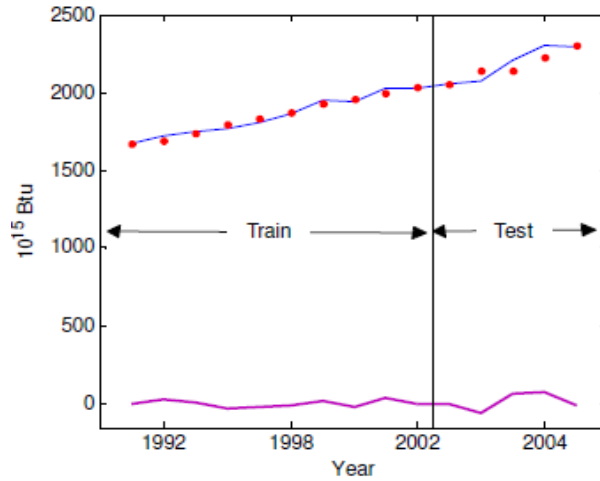


Fig. 6. Actual and predicted Canada oil consumption. Blue: Actual, red dots: predictions, thick purple: error

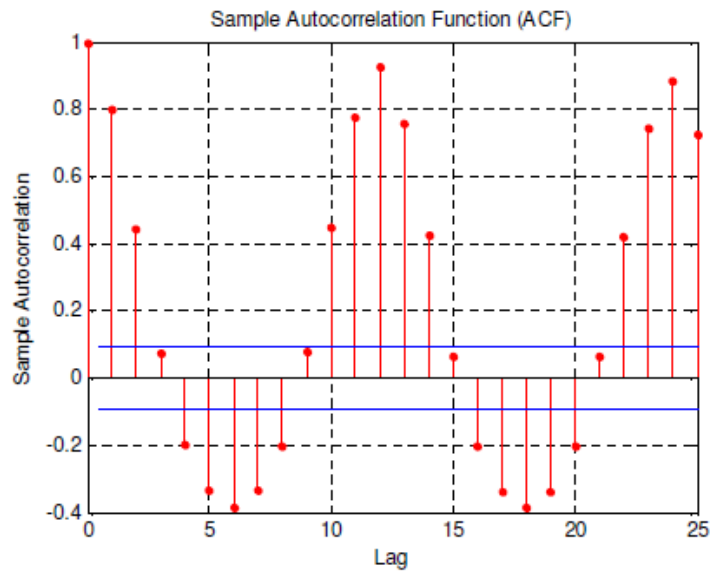


Fig. 7. Autocorrelation factor for U.S. monthly natural gas time series

In order to select proper input features, first an autocorrelation analysis is performed. The time lags with largest autocorrelation factors (ACF) will be selected as input features of the forecaster. The result of the autocorrelation analysis on training series is depicted in fig. 7. Interesting findings can be extracted from fig. 7. First, time lags 1 and 2 are highly informative, as they show the trend components of the consumption series. Besides, time lags 12, 13, 24 and 24 are other important features which represent the cyclic components of the natural gas consumption series.

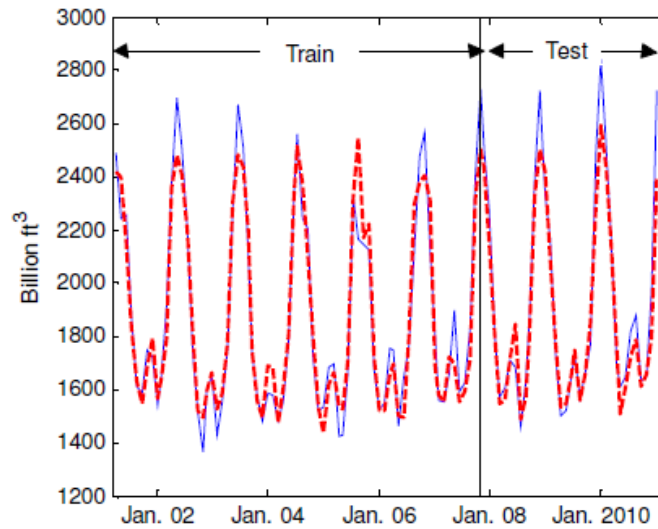


Fig. 8. Actual and predicted U.S. monthly natural gas consumption. Blue: Actual, dashed red: prediction

Table 5. Comparison of MAPE between different methods. The case of U.S. natural gas consumption

Method	Training	Test
MLP	5.5%	6.7%
LLNF	3.5%	4.1%
FT-LLNF	2.8%	3.8%

Hence, consumption at {1, 2, 12, 13, 24 and 25} months ago are employed for prediction of the gas consumption at the current month. The model is then constructed using the selected 6 input features.

The actual and predicted consumptions for training and test data are shown in fig. 8. A comparison with an MLP network and the LLNF model is also presented in table 5. Fig. 8 and table 4 reveal the noteworthy modelling and prediction capabilities of the proposed TF-LLNF method and its superior performance.

5.3. Performance comparison

This sub-section provides an overview of the results of the three case studies. Fig. 9 shows the MAPE of the MLP, LLNF and FT-LLNF for each case study. Clearly, the proposed approach exhibits better performance in all forecasts. Superiority of the FT-LLNF approach over the LLNF model demonstrates the favourable filtering effect of the F-transform on forecasting accuracy.

Another comparative study is carried out using MAPE ratio. This study is performed by dividing the test MAPE of the proposed approach to the test MAPE of the compared methods (i.e. MLP and LLNF). The values of MAPE ratio, computed for all case studies, are summarized in table 6. There are far differences between MLP and the proposed methods in different case studies,

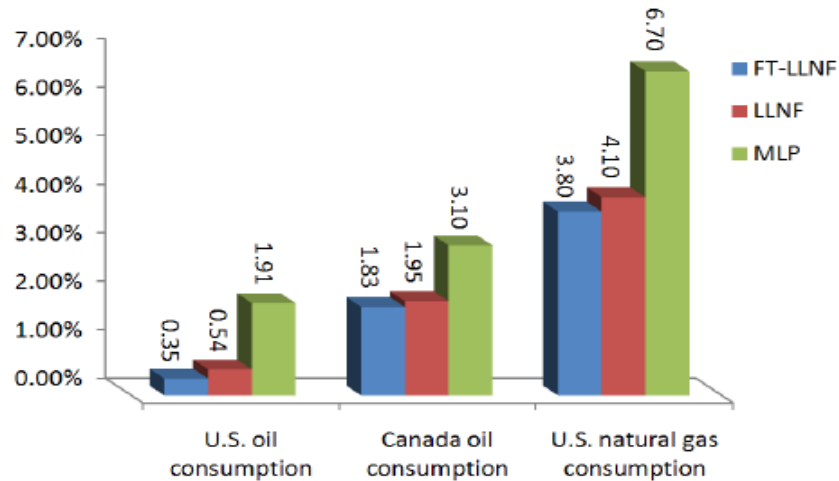


Fig. 9. Performance comparison of different models in three case studies

Table 6. Test MAPE ratio for all case studies

Method	MAPE ratio		
	U.S. oil consumption	Case study 2	Case study 3
MLP	0.18	0.59	0.56
LLNF	0.64	0.93	0.92
FT-LLNF	1	1	1

showing that the MLP model has not accurately modelled the nonlinearities of the energy consumption series, compared to the two other approaches. LLNF and FT-LLNF models showed comparable performance in case studies two and three. However the difference in their performance is noticeable for the first case study. It is concluded here that proper adjustment of the F-transform filter (e.g. shape and number of the basic functions), can considerably affect the performance accuracy of the forecast model.

6. CONCLUSION

This paper proposed a forecast method based on local linear neuro fuzzy models and fuzzy transform for energy consumption prediction. Combination of LLNF model, with distinguished modelling and forecast abilities and the recently introduced technique of F-transform, with noticeable noise removal properties, resulted in an effective forecaster which was used for prediction of energy consumption. The proposed FT-LLNF model was evaluated through different case studies. Prediction of annual and monthly energy consumption in the U.S. and Canada demonstrated accurate performance of the FT-LLNF model. Furthermore, multivariate and univariate predictions were also considered. Besides, comparison to an optimized MLP network and the LLNF model revealed superior performance of the proposed model. Through comparisons with LLNF model, the favourable data processing properties of the F-transform was recognised. Interestingly, considerable improvement in prediction accuracy through integration of F-transform into the forecast model was identified.

REFERENCES

- [1] International Energy Agency (IEA) [online] Available: www.iea.org
- [2] Harris, J.L. and Liu, L.M., 2002. Dynamic structural analysis and forecasting of residential electricity consumption. *International Journal of Forecasting*, 9(4), 437-455.
- [3] Edigera, V.S. and Akarb, S., 2007. ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy Policy*, 35(3), 1701-1708.
- [4] Bianco, V., Manca, O. and Nardinia, S., 2009. Electricity consumption forecasting in Italy using linear regression models. *Energy*, 34(9), 1413-1421.
- [5] Kumar, U., Jain, V.K., 2010. Time series models (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India. *Energy*, 35(4), 1709-1716.
- [6] Lee, Y.S. and Tong, L.I., 2011. Forecasting energy consumption using a grey model improved by incorporating genetic programming. *Energy Conversion and Management*, 52(1), 147-152.
- [7] Zhou, P., Ang, B.W. and Poh, K.L., 2006. A trigonometric grey prediction approach to forecasting electricity demand. *Energy*, 31(14), 2839-2847.
- [8] Kulkarni, V.R., Mujawar, S. and Apte, S., 2010. Hash Function Implementation Using Artificial Neural Network. *IJSC*, 1(1), 1-8.
- [9] Obi, J.C. and Imainvan, A.A., 2011. Decision Support System for the Intelligent Identification of Alzheimer using Neuro Fuzzy logic. *IJICIC*, 2(2), 25-38.

- [10] Azadeh, A., Araba, R. and Behfarda, S., 2011. An adaptive intelligent algorithm for forecasting long term gasoline demand estimation: The cases of USA, Canada, Japan, Kuwait and Iran. *Expert Systems with Applications*, 37(12), 7427-7437.
- [11] Azadeh, A., Khakestani, M. and Saberi, M., 2009. A flexible fuzzy regression algorithm for forecasting oil consumption estimation. *Energy Policy*, 37(12), 5567-5579.
- [12] Shakouri, H., Nadimia, R. and Ghaderia, F., 2008. A hybrid TSK-FR model to study short-term variations of the electricity demand versus the temperature changes. *Expert Systems with Applications* 36(2), Part 1, 1765-1772.
- [13] Yokoyama, R., Wakui, T. and Satake, R., 2009. Prediction of energy demands using neural network with model identification by global optimization. *Energy Conversion and Management*, 50(2), 319-327.
- [14] Chen, T., 2011. A collaborative fuzzy-neural approach for long-term load forecasting in Taiwan. *Computers & Industrial Engineering*, doi:10.1016/j.cie.2011.06.003.
- [15] Azadeh, A., Asadzadeh, S.M. and Ghanbari, A., 2010. An adaptive network-based fuzzy inference system for short-term natural gas demand estimation: Uncertain and complex environments. *Energy Policy*, 38(3), 1529-1536.
- [16] Nosrati Maraloo, M., Koushki, A.R., Lucas, C. and Kalhor, A. Long term electrical load forecasting via a neuro-fuzzy model. *14th International CSI Computer Conference, CSICC 2009*, art. no. 5349440, 35-40.
- [17] Suganthi, L. and Samuel, A., 2011. Energy models for demand forecasting- A review. *Renewable and Sustainable Energy Reviews*.
- [18] Nells, O., 2001. *Nonlinear System Identification*, Springer, Berlin, Germany.
- [19] Perfilieva, I., 2006. Fuzzy transforms: theory and applications. *Fuzzy Sets and Systems*, 157(8), 993-1023.
- [20] Di Martino, F., Loia, V. and Sessa, S., 2010. A segmentation method for images compressed by fuzzy transforms. *Fuzzy Sets and Systems*, 161(1), 56-74.
- [21] Di Martino, F., Loia, V. and Sessa, S., 2010. Fuzzy transforms for compression and decompression of color videos. *Information Sciences*, 180(20), 3914-3931.
- [22] U.S. Energy Information Administration [online] available: <http://www.eia.gov/>.
- [23] World Bank, 2006. *World Development Indicators 20 06*. World Bank, Washington, DC.