Improving Daily Natural Gas Forecasting by Tracking and Combining Models

Ronald H. Brown, David J. Kaftan, Jarrett L. Smalley, Maral Fakoor, Sarah J. Graupman, Richard J. Povinelli, and George F. Corliss
GasDay$^{\text{TM}}$ Project, Marquette University, Milwaukee, WI

***Abstract*—Daily demand forecasting is a necessary process in the supply chain of natural gas. One of the largest challenges in demand forecasting is adapting to systematic changes in demand. While there are many types of mathematical models for forecasting, there is no perfect formula. Ensembling several models often results in a better forecast. A common method for ensembling component models is taking a weighted average of the model forecasts. Due to the challenge of adapting to changes in demand, it is important to track the weights associated with each component model in an ensemble. We have developed an ensembling method, called the Dynamic Post Processor (DPP). The DPP ensembles several forecasting models, while tuning the weights based on recent performance of the models. It also removes biases from the component models in order to track changing patterns in natural gas demand. The ensemble yields better forecasts than any of the individual component models and reduces the mean forecasting error caused by systematic changes.**

***Index Terms*—Ensemble, dynamic systems, neural network, linear regression, recursive least squares, natural gas, load forecasting, energy**

1. Introduction

Natural gas is used to cook, heat water, dry clothes, and heat homes. The amount of natural gas needed to heat homes is weather sensitive; therefore, the winter has highly variable demands. Energy utilities need to procure the appropriate amount of daily natural gas to meet customer demands. Inaccurate forecasts can lead to excessive costs, which are passed to the customers. Therefore, accurate forecasts are needed.

Linear regression models have been found to predict natural gas demand well when weather and historical gas flow are used as dependent variables [1]. Although gas demand tends to be linearly correlated with temperature, there are limits to linear regression models' abilities to forecast. Neural networks are able to represent many of the nonlinear relationships and characteristics that regression models cannot [2]. For example, when temperature rises towards "comfortable" levels, residential consumers tend to have their furnaces on less, and thus the gas consumption pattern changes. Neural networks are able capture this nonlinear behavior [2].

Research at Marquette University's GasDay Laboratory has found that no single model can track accurately the behavior of real time weather data and consumer usage trends. To account for linear and non-linear trends, both linear regression models and artificial neural networks are used. Combining both models increases forecasting accuracy. However, the weighting of the models may no longer accurately represent emerging trends as new data becomes available if weights are predetermined by training data. We propose a mechanism to tune the ensemble model weightings to capture emerging trends in the incoming data during the heating season.

A background for this work, including previous works and a mathematical background, is given in Section 2. In Section 3, our methods are described. In Section 4, results of our experiment are presented and discussed. A conclusion is presented in Section 5.

2. Background

A. Ensemble Models

Ensemble models are effective for a variety of reasons; each component model of the ensemble uses a specific set of information and tends to make conclusions that focus on certain aspects of the data. Thus, each model usually represents an incomplete model of the entire system [3]. Ensembling the results uses more aspects of the information. The combination of forecasts also helps compensate for biases in individual models [4]. Various techniques exist for ensembling component models, with different techniques each having their own advantages and disadvantages [5], [6]. These techniques differ in the way that different components are weighted and combined into the ensemble. Weigel et al. [7] has shown the considerable advantage of a weighted ensemble over any single component. The weighted ensemble reduced overconfidence and mean error. Ensemble models that dynamically change how component models are included in the model have been shown to be advantageous over those that do not change [8]–[10]. Ensembling methods have been used extensively in forecasting natural gas demand [11]–[15].

B. Component Models

There are two component models we use throughout this paper; a linear regression (LR) model and an artificial neural network (ANN).

1) Linear Regression Models: Historically, LR models [16] have been a popular method for time series prediction and have been used for energy forecasting [1], [17]. Using linear regression, the energy demand S for a single day is estimated by

This work was supported by the GasDay Laboratory at Marquette University.

$$S \approx \hat{S} = \beta_0 + \sum_{j=1}^{m} \beta_j x_j, \tag{1}$$

where x_j is the value of input j, and β_j is the parameter which specifies the relationship between input j and the output for each of the m inputs. For energy forecasting, β_0 usually represents base load, and each of the inputs are weather variables such as heating degree days. For more on how linear regression models are used in natural gas forecasting, see [1].

2) Artificial Neural Network Models: One of the weaknesses of LR models is the assumption that all the inputs used in the model have a linear relationship with the outputs. This is rarely true in real world systems and certainly not true in energy forecasting. For this reason, many energy forecasters use Artificial Neural Networks (ANNs) [18], as they are capable of modeling the nonlinear relationships between the inputs and outputs of a system [19]. The ANNs used in this paper are feed-forward neural networks which have been shown to be effective approximators for many systems [20].

C. Recursive Least Squares

Natural gas consumption model parameters are estimated through absolute error minimization between predicted and actual consumption. Since natural gas consumption provides a steady stream of data, the parameters that define a natural gas consumption model are out of date days after they are first estimated. To incorporate new data into the estimation of parameters, we use a *recursive identification* method.

Recursive Least Squares is the recursive identification method used in this paper. Recursive Least Squares was first introduced by Gauss [21], and later clarified by Plackett [22], as a method for updating estimations of unknown parameters given additional data to a model. It has since been used to that effect in numerous energy forecasting applications [11], [23], [24]. A full derivation of this method can be found in [25].

3. METHODS

Recall that the major goal of these methods is to reduce the error of daily natural gas forecasting over the course of a year. We build component models to map weather inputs to daily natural gas demand. Due to the dynamic characteristics of natural gas demand, we expect the component models to become less representative of the underlying relationships over the course of the year being forecast. To this end, we track characteristics of the system by shifting and scaling the outputs of the dynamic models. We refer to this shifting and scaling as *tuning*. In practice, we find that an error bias often remains after tuning. To compensate, we subtract the recent mean error. Finally, a weighted average of model outputs is taken based on the recent variance of each component's errors.

A. Creating the Component Models

A collection of component forecasting models is chosen for ensembling. The inputs to these component models are weather and calendar variables. The output of each of these models is the forecasted natural gas consumption for one operating area with the time horizon of the next day. We use linear regression and artificial neural networks. However, the method proposed in this paper is not exclusive to any particular model structure, nor any number of components. Therefore, the component models will be described as black boxes with the aforementioned inputs and outputs for the remainder of the Methods. Additionally, the component models are assumed to have been determined from a data set independent of the data set on which they are tested.

B. Evaluation of the Component Models

Each component model forecasts the consumption of natural gas for day k slightly prior to the start of that day. The component forecast demand is denoted by $\hat{\mathrm{c}}_k^j$ for day k and component model j. The error associated with each forecast can only be determined two days later due to the time it takes to record and publish actual flow. For example, the most recent error known on day k (e.g., Thursday) is from day $k-2$ (e.g., Tuesday).

C. Tuning Each Component Model

We expect the characteristics that define the demand in an operating area to change slowly over time. For example, if a new neighborhood is developed, the output of each model should be scaled and shifted up to match a growing demand. To generalize the tracking of dynamic daily gas demand characteristics to any class of forecasting model, we scale and shift the output of each component model rather than the parameters that define the component model. In practice, this means each component's forecast demand, $\hat{\mathrm{c}}_k^j$, is used as an input to a two parameter linear regression model. The tuning parameters θ_0^j and θ_1^j of the linear regression model correspond to the coefficients of the bias term and $\hat{\mathrm{c}}_k^j$ for the j^{th} component model, respectively. The output from the linear regression model is the tuned component forecast, $\tilde{\mathrm{c}}_k^j$.

D. Updating the Tuning Parameters

The updated tuning parameters are determined using a two-step ahead method using Recursive Least Squares. The tuning parameters are updated using the error from day $k-2$. The error generated for day $k-2$ does not reflect the current state of the system, as θ is updated on day $k-1$. An a posteriori forecast for day $k-2$, $\tilde{\mathrm{c}}_{k-2}^{j'}$, is therefore made using the θ updated on day $k-1$.

$$\tilde{c}_{k-2}^{j'} = \begin{bmatrix} 1 & \hat{\mathrm{c}}_{k-2}^j \end{bmatrix} \begin{bmatrix} \theta_0^j \\ \theta_1^j \end{bmatrix}. \tag{2}$$

The a posteriori error is

$$\tilde{e}_{k-2}^{j'} = \tilde{c}_{k-2}^{j'} - s_{k-2}, \tag{3}$$

where s_{k-2} is the measured natural gas consumption. $\tilde{\mathrm{e}}_{k-2}^{j'}$ is processed further because the measured natural gas consumption is subject to human and mechanical errors. Thresholds are set such that large errors are limited so that extreme errors (often caused by bad data) do not impact the weights. Small errors are treated as zero error and the tuning step is skipped. The small error threshold is set manually and denoted $minerr$. The large error threshold is calculated by

$$maxerr = maxerrfactor \cdot s_k + minerr, \tag{4}$$

where $maxerrfactor$ is set manually. With this limited error, the tuning parameters are updated using Recursive Least Squares. The tuning is forced to favor the output of the component model, rather than the bias term. This occurs according to Equation 5:

$$\theta = (1-\gamma)\theta + \gamma \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{5}$$

where γ is a manually set forgetting factor.

E. Calculating Recent Mean and Variance Error

We have found that our tuning through recursive least squares does not drive the mean error of the forecast to zero. For this reason, we subtract the recent mean error from our forecasts. Using the newly tuned and stabilized θ, the error from two days ago is recalculated and denoted by $\tilde{e}^{j''}_{k-2}$. This error is limited in the same manner as previous errors. The recent mean error, $\tilde{\mu}^j_{k-2}$, is updated

$$\tilde{\mu}^j_{k-2} = \alpha \cdot \tilde{\mu}^j_{k-3} + (1-\alpha) \cdot \tilde{e}^{j''}_{k-2}, \tag{6}$$

where α is a manually set forgetting factor. Subtracting the recent mean drives the mean error to zero (for an empirical example, see Figure 2). Intuitively, given zero mean error, it makes sense to weight the component models based on the variance of error. The recent error variance is updated,

$$\tilde{v}^j_{k-2} = \alpha \cdot \tilde{v}^j_{k-3} + (1-\alpha) \cdot (\tilde{e}^{j''}_{k-2} - \tilde{\mu}^j_{k-2})^2. \tag{7}$$

The recent mean is subtracted from the current forecast $\tilde{c}^j_k$, yielding the component's tuned forecast, $\bar{c}^j_k$.

F. Weighted Mean Ensemble

The components are ensembled using a weighted mean. Given n component models, the weight of each component model, j, is

$$w^j = \frac{\frac{1}{\sqrt{\tilde{v}^j}}}{\sum_{i=1}^{n} \frac{1}{\sqrt{\tilde{v}^i}}}, \tag{8}$$

and the weighted forecast of the ensemble is

$$\hat{s}_k = \sum_{j=1}^{n} w^j \bar{c}^j_k \tag{9}$$

This forecast is a weighted ensemble of the component models that includes adjustments to recent variances.

The method therefore reduces several types of errors. The squared error of the resulting forecast is reduced by tuning each of the components using Recursive Least Squares. The mean error is reduced by subtracting the recent mean error from each component. Finally, the components are ensembled by scaling the weights based on recent errors.

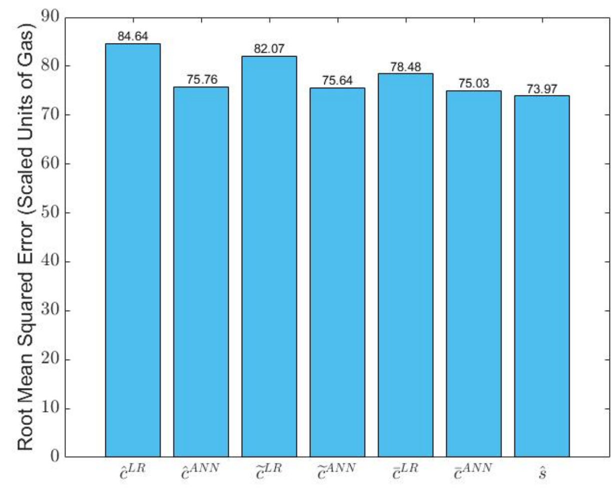

Figure 1. Root mean squared error of each stage of DPP for Linear Model and ANN components.

4. Results and Discussion

Two component models are trained to predict natural gas demand for the upcoming day. Daily cumulative flow data for an operating area in the midwest United States is used. The weather data - to be used as inputs to the component models - is obtained from the National Oceanic and Atmospheric Administration (NOAA). The model is trained on data from September of 1993 to August of 2015 and is tested on data from September of 2015 to August of 2016.

The performance of each component at each stage of the tuning is evaluated using root mean squared error (RMSE). RMSE is particularly useful in natural gas forecasting because it is especially sensitive to large errors, which are of great interest in practice. The RMSE improved after each stage in the tuning. Further, the RMSE of the ensemble is better than the RMSE of each of the components.

Much of the reason for tuning the component models is to remove the mean error. For this reason, we show how quickly the mean error of each component tends to zero. As seen in Figure 2, the mean error of the ensemble tends to zero over the period of the 365 days. The mean error of the raw component outputs does not approach zero. Therefore, the tuning and ensembling of the components filter out a significant amount of error. Figure 2 is also demonstrates how the ensemble can perform better than the best component that makes it up. The mean error of the ensemble is never the farthest from zero, but it is often times the closest to zero.

5. Conclusion

The proposed dynamic tracking and ensembling presented in this paper improves forecasts of natural gas consumption in a non-stationary system. We can therefore improve planning for natural gas demand even as the characteristics that drive demand change over time. Improved forecasts result in lower planning costs for utilities.

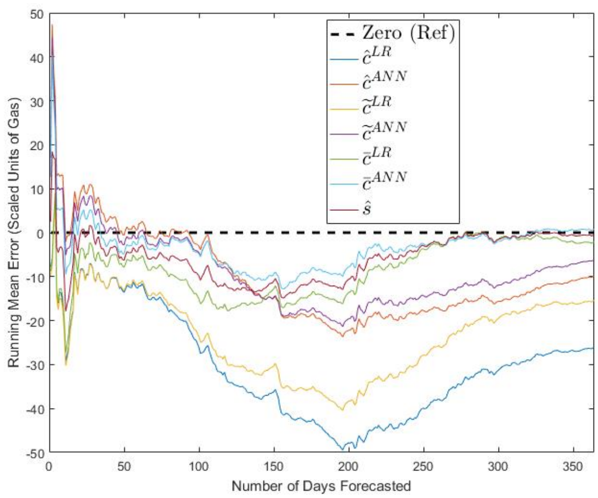

Figure 2. The running mean error of each stage of the DPP for Linear Model and ANN components. For each day k, the errors of days 0 through k are summed, then divided by k. The mean error of forecasts when the recent mean is subtracted out is driven to zero.

REFERENCES

[1] S. R. Vitullo, R. H. Brown, G. F. Corliss, and B. M. Marx, "Mathematical models for natural gas forecasting," *Canadian Applied Mathematics Quarterly*, vol. 17, no. 7, pp. 807–827, Jan. 2009.

[2] R. H. Brown and I. Matin, "Development of artificial neural network models to predict daily gas consumption," in *Industrial Electronics, Control, and Instrumentation, 1995., Proceedings of the 1995 IEEE IECON 21st International Conference on*, vol. 2. IEEE, 1995, pp. 1389–1394.

[3] M. Gilliland, U. Sglavo, and L. Tashman, *Business Forecasting: Practical Problems and Solutions*. John Wiley & Sons, 2016.

[4] P. Goodwin, "New evidence on the value of combining forecasts," *Foresight: The International Journal of Applied Forecasting*, no. 12, pp. 33–35, 2009.

[5] E. Bauer and R. Kohavi, "An empirical comparison of voting classification algorithms: Bagging, boosting, and variants," *Machine Learning*, vol. 36, no. 1-2, pp. 105–139, 1999.

[6] L. Rokach, "Ensemble-based classifiers," *Artificial Intelligence Review*, vol. 33, no. 1-2, pp. 1–39, 2010.

[7] A. P. Weigel, M. Liniger, and C. Appenzeller, "Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?" *Quarterly Journal of the Royal Meteorological Society*, vol. 134, no. 630, pp. 241–260, 2008.

[8] D. Jiménez, "Dynamically weighted ensemble neural networks for classification," in *Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence. The 1998 IEEE International Joint Conference on*, vol. 1. IEEE, 1998, pp. 753–756.

[9] A. H. Ko, R. Sabourin, and A. S. Britto Jr, "From dynamic classifier selection to dynamic ensemble selection," *Pattern Recognition*, vol. 41, no. 5, pp. 1718–1731, 2008.

[10] J.-h. Zhai, H.-y. Xu, and X.-z. Wang, "Dynamic ensemble extreme learning machine based on sample entropy," *Soft Computing*, vol. 16, no. 9, pp. 1493–1502, 2012.

[11] A. Khotanzad, H. Elragal, and T.-L. Lu, "Combination of artificial neural-network forecasters for prediction of natural gas consumption," *IEEE Transactions on Neural Networks*, vol. 11, no. 2, pp. 464–473, 2000.

[12] G. Xu and W. Wang, "Forecasting China's natural gas consumption based on a combination model," *Journal of Natural Gas Chemistry*, vol. 19, no. 5, pp. 493–496, 2010.

[13] H. Elragal, "Improving neural networks prediction using fuzzy-genetic model," in *Radio Science Conference, 2004. NRSC 2004. Proceedings of the Twenty-First National*. IEEE, 2004, pp. C12–1.

[14] N. H. Viet and J. Mandziuk, "Neural and fuzzy neural networks in prediction of natural gas consumption," *Neural, Parallel & Scientific Computations*, vol. 13, no. 3, pp. 265–286, 2005.

[15] R. Chen, J. Wang, L. Wang, N. Yu, and P. Zhang, "The forecasting of China natural gas consumption based on genetic algorithm," in *INC, IMS and IDC, 2009. NCM'09. Fifth International Joint Conference on*. IEEE, 2009, pp. 1436–1439.

[16] N. R. Draper, H. Smith, and E. Pownell, *Applied Regression Analysis*. Wiley New York, 1966, vol. 3.

[17] T. Haida and S. Muto, "Regression based peak load forecasting using a transformation technique," *IEEE Transactions on Power Systems*, vol. 9, no. 4, pp. 1788–1794, 1994.

[18] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," DTIC Document, Tech. Rep., 1985.

[19] D. C. Park, M. El-Sharkawi, R. Marks, L. Atlas, and M. Damborg, "Electric load forecasting using an artificial neural network," *IEEE Transactions on Power Systems*, vol. 6, no. 2, pp. 442–449, 1991.

[20] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.

[21] C.-F. Gauss, *Theoria combinationis observationum erroribus minimis obnoxiae.-Gottingae, Henricus Dieterich 1823*. Henricus Dieterich, 1823.

[22] R. L. Plackett, "Some theorems in least squares," *Biometrika*, vol. 37, no. 1/2, pp. 149–157, 1950.

[23] D.-C. Lee, K.-J. Lee, J.-K. Seok, and J.-W. Choi, "Online capacitance estimation of dc-link electrolytic capacitors for three-phase ac/dc/ac pwm converters using recursive least squares method," *IEEE Proceedings-Electric Power Applications*, vol. 152, no. 6, pp. 1503–1508, 2005.

[24] G. Mbamalu and M. El-Hawary, "Load forecasting via suboptimal seasonal autoregressive models and iteratively reweighted least squares estimation," *IEEE Transactions on Power Systems*, vol. 8, no. 1, pp. 343–348, 1993.

[25] L. Ljung, *System Identification*. Wiley Online Library, 1999.