The Impact of Additional Weather Inputs on Gas Load Forecasting

Bo Pang
Marquette University

THE IMPACT OF ADDITIONAL WEATHER INPUTS ON GAS LOAD FORECASTING

by

Bo Pang, B.S.

A Thesis Submitted to the Faculty of the
Graduate School, Marquette University,
in Partial Fulfillment of the Requirements for the
Degree of Master of Science

Milwaukee, Wisconsin

August 2012

ABSTRACT
THE IMPACT OF ADDITIONAL WEATHER INPUTS ON GAS LOAD FORECASTING

Bo Pang, B.S.

Marquette University, 2012

Natural gas utilities need to estimate their customers' gas demand accurately. This thesis develops a number of daily forecasting models for test the possibility to extend the weather inputs in the current method for three different operating areas. Our goal is to improve the accuracy of our forecast by extending the number of inputs used by the existing GasDay model.

We present a detailed explanation of the identification of the significance for each of the new weather input candidates. The significance of the new weather inputs was tested by statistical hypothesis testing, by forecasting performance testing, and by unusual day evaluation. We show that with some combinations of additional weather instruments, the accuracy of the forecast is improved.

For most gas utilities, the primary use of natural gas is for space heating, so temperature is a critical factor when we build forecast models. In this thesis, we develop a method to split the Heating Degree Day (HDD) term into smaller pieces and generate the forecast based on these small factors. We name the method that developed as Multiple Weather Station (MWS) model in Chapter 4. We show that the MWS model yields better results compared to the existing method.

ACKNOWLEDGEMENTS

Bo Pang, B.S.

I wish to express my gratitude to many people who have provided me with crucial help and support while carrying out the research for this thesis. Without the help and encouragement that my advisors, Dr. Ronald Brown and Dr. George Corliss, have provided, none of this would be possible. Dr. Brown, Dr. Corliss, Dr. Povinelli, and Dr. Vitullo, my committee members, have all spent countless hours in discussion of this project and provided me with ideas while keeping me on the right track.

I would like to extend a special thanks to my friends, Tsuginosuke Sakauchi, Yifan Li, Anisha DSilva, James Lubow, and Samson Kiware for supporting me throughout my graduate studies at Marquette University. Their thoughts, knowledge, and insights have added to my education.

I would also dedicate this work to my parents Mr. Guangwei Pang , Ms. Ting Yu, and my wife Rubing Han, who have supported me and offered encouragement throughout my life and education. I appreciate their unselfish love, support, and encouragement. Their confidence in me has helped me achieve this level of education as well as other successes in my life. This thesis would not be possible without them.

Finally, I would like to acknowledge the financial support provided by the GasDay Lab and also Marquette University's Department of Electrical and Computer Engineering.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction to Natural Gas Forecasting

1.1 Overview of the United States natural gas industry

Natural gas provides about 25% of the energy used in United States. According to the American Gas Association (AGA) and the U.S. Energy Information Administration, about 58 million American homes use natural gas because it is a clean and reliable energy source [3]. In recent decades, the use of natural gas increased rapidly in the field of electric power generation, cooling, and as a transportation fuel due to its environmental advantages of low emissions. According to the U.S. Department of Energy [41], natural gas consumption is expected to increase about 11% by 2030. Therefore, accurately forecasting natural gas consumption is as important as the prediction of any other kind of energy consumption.

Users of natural gas fall into five categories: residential, commercial, industrial, agricultural, and power generation purposes [1], [42] . Residential customers use natural gas in their homes to fuel furnaces and appliances such as stoves, clothes dryers, and water heaters. The use of natural gas for commercial customers can be retail space, office buildings, restaurants, hospitals, and hotels. Industrial customers use natural gas for heating processes, as fuel for the generation of steam, for foundries, and as a raw material from which their product is processed. Various agricultural processes use natural gas, such as canning tomatoes [9], drying corn, or powering irrigation pumps. Electric utilities and independent power producers use natural gas to generate electricity, and some industrial electricity is sold back to the grid.

Distribution plays an important role in the process of delivering natural gas to customers. Some large industrial, commercial, and electric generation customers receive natural gas directly from high capacity interstate and intrastate pipelines. Most other users receive natural gas from the local utilities, also called local distribution companies (LDCs). LDCs are regulated utilities involved in the delivery of natural gas to customers within a specific geographic region. There are three types of natural gas utilities: utilities owned by investors, utilities owned by local governments, and utilities owned privately [35].

1.2 The purpose of forecasting natural gas demand

Natural gas demand is different from natural gas consumption. Consumption, also known as load or send-out, is less than the demand when LDCs interrupt (shut off gas supply) their interruptible industrial customers because the actual demand is higher than the available supply. When no customer is interrupted, the demand is equal to the consumption.

We forecast natural gas consumption to predict the expected gas demand in a region. Accurate forecasts play an important role for LDCs as well as for their customers.

When the actual consumption exceeds the forecast, LDCs are forced to extract gas from storage, interrupt service to their customers, or purchase additional gas on a spot market. When the demand is lower than the forecast, utilities need to find a way to store the additional gas, spend more time to sell it, or leave it in the pipe and increase the penalties imposed by the pipeline companies. All these results lead to higher operational costs to LDCs. Hence, there is a need for accurate forecasting of natural gas demand.

Demand forecasts are classified depending on the forecasting period: long, intermediate, short, and very short term forecasts. Long term forecasts typically are used to forecast the gas demand for more than one year [10]. An intermediate forecast is made for a range between one month and one year. Short term forecasts are made for one day to one week. A very short term forecast predicts one hour to about 30 hours ahead [45]. For different forecasting periods, the methods and variables being used for forecasting are not the same. The work in this thesis focuses on short term forecasting, or specifically on daily forecasts for time horizons of one to about seven days.

1.3 Marquette’s GasDayLab

Marquette University's GasDay Lab is a research facility that has been developing and refining natural gas demand forecasting models since 1993. Developed by Dr. Ronald H. Brown and students at Marquette University, GasDay uses weather data, gas usage data, and domain knowledge to forecast natural gas flow accurately. Currently serving 26 utilities in 22 states, GasDay forecasts about one fifth of the nation's natural gas usage for residential, commercial, and industrial customers in more than 130 operating areas. The GasDay Lab has developed a series of mathematical algorithms and models and then implemented them in software built for each of the operating areas for LDCs to provide accurate demand forecasts.

1.4 Problem: LDCs need accurate demand forecasting

Natural gas demand depends primarily on temperature, weather conditions, day of the week, holidays, and sudden weather changes. In addition, economic factors (natural gas price, GDP, occupancy rate, and number of customers) affect long-term gas demand. However, the current GasDay forecasting models are primary temperature related. Research shows that there are variables in addition to temperature that can have both a direct and an indirect effect on forecasts of energy loads [11], [26], [27], [45].

The LDCs are distributed all over the country. Each of them has its own geographic features and local climate that may affect gas consumption. Hence, each utility has its own sensitivity to the temperature and other weather effects [40], [52]. The current forecasting method used by GasDay usually has forecasting errors for gas flow that is not temperature sensitive or has low temperature sensitivity [45]. In other words, the current models need to be updated with additional weather instruments[1]. On the other hand, unique inputs are needed to fit the demand of individual customers. The current GasDay models are designed primarily for the customers who use natural gas for heating purposes, whose gas consumption is highly correlated to the temperature. However, some LDCs use natural gas as the primary energy source for irrigation. Thus, it is almost

[1] In this thesis, weather instruments refer to independent variables that have potential impact on gas load such as temperature, dew point, precipitation, and so on.

impossible for the current GasDay models to provide a good forecast for such areas. According to our data sources, much more weather data (such as Precipitation, Cloud Cover, Dew Point, and so on) are available to use, providing the opportunity to estimate the impact of non-temperature variables on the consumption of natural gas. Our challenge is to identify and evaluate appropriate variable(s) that can help to obtain an accurate forecast.

Besides the problem mentioned above, the accuracy of the calculation for the Heating Degree Days (HDDs) also has hurt forecast accuracy. Inaccurate HDDs may lead to large errors in gas flow forecasts [22], especially in critical "shoulder months" in the spring and fall. In the current GasDay model, HDD is calculated as Equation 1.1 [5], [6], [16].

$$HDD_k = \mathrm{Max}(T_{ref} - T_k, 0) \ , \tag{1.1}$$

where T_k is the average temperature for the k^{th} day, and T_{ref} is the reference temperature, historically set to 65 ℉ or 18 ℃.

The LDCs are distributed all over the country. Some of them provide services for an especially large geographic area, which leads to an inaccurate estimation of HDDs using the current method. For such an area, the temperature data from a single weather station is not enough to represent all characteristics of the operating area, which is the reason that GasDay supports weighted multiple weather stations. The current GasDay method is to divide a single geographic area into multiple small operating areas based on the gas flow data reported by the LDCs [25]. However, occasionally, some LDCs only have single time series of daily gas flow data reported, which means their service areas are almost impossible to divide into multiple geographic regions. For these LDCs, GasDay forecasting errors are larger than if the region was subdivided.

1.5 Our solution: Additional Weather Inputs

The inputs of the current model cannot fit the characteristics of the all the regions we forecast. Thus, by considering more variables in the current model, we may be able to reduce the difference between the actual and forecast natural gas consumption. In Chapter 3, we propose to extend the weather inputs by including new weather instruments supported by our data resources. As a simplification of the current GasDay model, Equation 1.2 has the most significant independent variables (HDD65 and HDD55), which fits the actual flow with Adjusted-R^2 of more than 90%.

We define the variables of Equations 1.2:

S_k is the actual gas consumption for the k^{th} day;

$\hat{S}_k$ is the predicted gas consumption for the k^{th} day;

β_0 is the constant or intercept of the model;

β_1 is the coefficient of HDD65 with a 65 degree reference temperature;

β_2 is the coefficient of HDD55 with a 55 degree reference temperature;

A simple representation of the current GasDay model:

$$\hat{S}_k = \beta_0 + \beta_1 * HDD65_k + \beta_2 * HDD55_k \ . \tag{1.2}$$

The HDDs explain most of the variation of gas flow that is used for heating. To provide accurate forecasting results, however, the existing GasDay regression model has more variables embedded, such as the Sin/Cos (Day of Week), the Sin/Cos (Day of Year), the lagged gas flow, the lagged temperature, and other factors. To protect the confidential knowledge of GasDay, the actual GasDay production model is not published in this paper. Alternatively, we use Equation 1.3 to represent the existing GasDay model as a reference model for comparison:

$$\hat{S}_k = \beta_0 + \sum_{n=1}^{N} \beta_n * GDinput_n \,, \tag{1.3}$$

where *GD input* represents the actual input variables in the existing GasDay models. In Chapter 3, we will discuss the possibility of improve our forecasting by adding more weather inputs as Additional Weather Input (AWI). Therefore, we compare Equation 1.3 with the model:

$$\hat{S}_k = \beta_0 + \sum_{n=1}^{N} \beta_n * GDinput_n + \beta_{N+1} * (New\ weather\ input)\,. \tag{1.4}$$

Our goal is to find the most valuable input candidate(s) that can be used in the future work of GasDay.

Above, we have briefly introduced our proposed solution for adding new external inputs into the existing GasDay regression model. To solve the problem caused by combining the real weather stations into a virtual weather station, in Chapter 4, we propose to extend the weather inputs to support multiple weather stations co-existing in one model (shown in Equation 1.5). In other words, both the reference model (Equation

1.4) and the Multiple Weather Station (MWS) model (Equation 1.5) are based on the same real weather stations. The critical difference is that the real weather stations are combined into one virtual weather station in the reference mode, but not in the MWS model. Again, we compare the proposed solution with the reference model which is shown in Equation 1.4. $HDD65_{(virtual)}$, $HDD55_{(virtual)}$, and $CDD55_{(virtual)}$ denote the weighted average of HDD65, HDD55, and CDD65 among multiple weather stations. A detailed explanation of these variables is given in Chapter 4.

Reference model:

$$\hat{S}_k = \beta_0 + \beta_1 * HDD65_{(virtual)k} + \beta_2 * HDD55_{(virtual)k} + \beta_2 * CDD65_{(virtual)k} + \beta_n * GDinput\,, \quad (1.5)$$

where the *GDinput* represents all the other independent variable that are in the existing GasDay model.

MWS model (Multiple Weather Stations):

$$\hat{S}_k = \beta_0 + \beta_1 * HDD65_{k(WSi)} + \beta_2 * HDD55_{k(WSi)} + \beta_3 * CDD65_{k(WSi)} + ... + \beta_n * GDinput\,, \ (1.6)$$

where WS_i is the i^{th} weather station of this operating area; $i = 1$ to 6, which is the actual number of weather stations that the LDC used in this operating area. As Equation 1.5 shows, the temperature is represented by each real weather station independently instead of by an approximation of the virtual weather station.

1.6 Performance Criteria

Performance criteria are needed to measure whether our models perform better than the existing GasDay model. The outputs of our models are the estimated daily consumption of natural gas [10], [13], [28]. The error in the prediction on the k^{th} day is

$$e_k = \hat{S}_k - S_k \quad . \tag{1.7}$$

The accuracy of our model is assessed based on the following measures of forecast accuracy: N is the total number of days in the set to be analyzed.

1. Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{(\sum_{k=1}^{N} e_k^2)}{N}} \tag{1.8}$$

2. Mean Absolute Percent Error (MAPE):

$$MAPE = \left(\frac{1}{N} \sum_{k=1}^{N} \left| \frac{e_k}{S_k} \right| \right) * 100\% \tag{1.9}$$

3. Weighted Mean Absolute Percent Error (wMAPE):

$$wMAPE = \left(\frac{\frac{1}{N} \sum_{k=1}^{N} \left(\frac{|e_k|}{S_k} * S_k \right)}{\frac{1}{N} \sum_{k=1}^{N} S_k} \right) * 100\% = \left(\frac{\sum_{k=1}^{N} |e_k|}{\sum_{k=1}^{N} S_k} \right) * 100\% \tag{1.10}$$

A complete discussion of the results can be found in Chapters 3 and 4.

1.7 Unusual days evaluation

For any kind of forecasting, there may be some unusual events that occur in underlying data, which is more difficult to forecast than the overall usual events. To GasDay, forecasting on the days where the unusual event occurs are more important than forecasting on the normal days.

“Unusual Day” is a term that used in the GasDay lab to represents the days on which an unusual event occurs. These unusual events include sudden temperature increases/decreases, high humidity/low humidity, extremely cold, and so on. Based on these unusual events, our unusual days are: coldest day, colder (warmer) than normal days, windiest heating day, colder (warmer) today than yesterday, the first cold (warm) days, high (low) humidity heating days, and sunny (cloudy) heating days [46]. The current GasDay models may have larger error when the unusual events occur. Hence, evaluating the model performance on unusual days has become a significant part of our testing process. The next section briefly reviews the outline and the overall structure of this thesis.

1.8 Thesis Outline

This thesis begins with background of natural gas industry and a problem statement in Chapter 1. Chapter 2 includes a literature survey of forecasting approaches and an introduction to the new weather instruments that we use in following chapters. Chapter 3 gives statistical hypothesis tests, forecasting performance testing, and the unusual day evaluation on new weather instruments and the analysis of the significance of all the selected inputs. In Chapter 4, we discuss the possibility of improving our forecasting accuracy in terms of expanding HDDs and CDDs by the number of real weather stations. Finally, Chapter 5 offers conclusions and some suggestions for the current GasDay research. Some opportunities for further improvement are stated at the end of Chapter 5 as well.

CHAPTER 2

Survey of Forecasting Techniques and New Weather Instruments

This chapter provides a summary of the literature of linear regression and statistical hypothesis testing. Then we introduce the properties of the new weather instruments candidates and our expectation of the impact of the new variables on the forecasts for natural gas consumption.

2.1 Guidelines for forecasting

Forecasting has a long history in the development of human society as a process of forecasting an expected result of some future event. The ancient Egyptians predicted harvests from the level reached by the Nile River in the flood season [38]. In the 17th century, William Petty discovered a seven-year business cycle which suggested a basis for systematic economic forecasts [24]. In the United States, the forecasting industry developed around 1910-1930, and it has become an important part of our lives [24]. In the development of forecasting, principles and guidelines for forecasting are summarized from both experts' conclusions and empirical studies. J. Armstrong [4] offers some guidelines for forecasting, based on the work of 40 leading experts who have reviewed the published research involving thousands of studies. With the experts' guidelines of forecasting, forecasters can construct their forecasting methods and models more effectively and with fewer mistakes.

Some of Armstrong's guidelines for forecasting are:

1. Use a simple model. Models should have as few factors as possible before becoming complicated. Complicated models tend to run into more model specification problems[2]. In this thesis, our solutions are based on linear regression models, a basic approach that is widely used in the domain of forecasting.

2. Have an expectation of the objective. Being clear about the goal we are going to achieve with the models can help us choose an appropriate method.

3. Use domain knowledge and theory to choose the correct variables. It is important to have the correct inputs; otherwise the model will not perform as expected. In Chapter 3, we select our weather instrument candidates with the domain knowledge that is given at the end of this chapter.

4. Use as much data as possible. Insufficient data may lead to an ineffective estimation. Data used in this thesis are daily time series weather and gas flow data that are longer than five years. For each training data set and testing data set, we have many observations for each variable.

---

[2] "Model specification" refers to the initial steps of selecting an appropriate function and choosing variables. Model misspecification can result in biased and inconsistent estimates of the coefficients [21].

5. Start a model from initial model specifications. Use a general-to-specific approach. Always define a general model and then make it specific. The extending of inputs is based on the simplified model (Equation 1.4) of the current GasDay model. We are going to use this general model as a benchmark to explore more possibilities.

6. Use a single model equation when it is possible. Multiple equation models and very complicated models may lead to inaccurately estimated coefficients and inconsistent results. All models built in this thesis are from single model equations.

7. Test and correct for model misspecification. For example, some input factors might need to be removed or added to the initial model to correct the model's misspecification. In Chapter 3, we apply this technique to test the significance of selected weather instruments frequently.

8. Rerun the original model and examine new performance against a predetermined benchmark. In this thesis, we compare new candidate models to reference models at the end of both Chapters 3 and 4. A summary of the performance comparison of the solution methods and current methods is given in Chapter 5.

As we stated above, the research of this thesis follows Armstrong's guidelines for forecasting. Beside the guidelines of forecasting, he also offers some approaches that can help forecasters select appropriate forecasting method(s), which are given in the next section.

2.2 Approaches for selecting methods

According to Armstrong's guideline number two, selecting an appropriate forecasting method is a critical step toward obtaining a good result. However, there is no single correct forecasting method to use [4]. Selecting a method should be based on the objectives and conditions of the forecast. Armstrong [4] suggests six approaches for selecting methods: Selecting by convenience, market popularity, structured judgment, statistical criteria, relative track records, and guidelines from prior research. In this thesis, we discuss three of Armstrong's approaches that may be valuable to improve current GasDay modeling and that are used in this thesis.

1. Market popularity

Market popularity involves determining what methods are used by other people or organizations. Armstrong suggests two assumptions: i) "Over time, people figure out which methods work best"; and ii) "What is best for others will be best for you." Work in this thesis is based on the linear regression method and statistical analysis, which are two methods used widely in the area of forecasting [18]. Both of them have proven to be very efficient approaches in prior research in the GasDay Lab. The work of this thesis is based on a linear regression model of daily natural gas consumption and the statistical analysis applied in Chapter 3.

2. Statistical criteria

Forecasters often use statistical criteria to select methods. This approach is useful to help forecasters determine whether they should keep or drop variables and whether they should use a particular method. Statisticians and econometricians rely heavily upon whether a method meets statistical criteria, such as distribution of errors, statistical significance of relationships, or the Durbin-Watson statistic [4], to help a forecaster judge the significance of variables. Even though the decision might be arbitrary, the combined use of statistical criteria and domain knowledge leads to results with a higher degree of confidence in practice [4]. The main approach of selecting the inputs to be used Chapter 3 is based on statistical analysis and hypothesis testing.

3. Guidelines from prior research

"Drawing upon extensive research, we developed guidelines to help practitioners decide which methods are appropriate for their situations". J. Scott Armstrong [4]. Extensive research has developed principles for forecasting. Through guidelines, such as those given by Armstrong, one can select methods more likely to perform well in one's own application. Based on previous research of the GasDay Lab, guidelines have been developed to help researchers decide which methods to use in their research. This thesis is based on prior research [22], [32], and [45] of the GasDay Lab, which guides our own method selection and development.

The other three approaches are selecting by convenience, structured judgment, and relative track records. For a thorough interpretation of these approaches, refer to Armstrong's guidelines for selecting methods [4]. According the literature of forecasting, it is not hard to find that linear regression is a very common and effective technique. Many researchers and practitioners of forecasting are very familiar with this approach. To help new researchers understand this technique, we give a brief discussion of linear regression models in the next section.

2.3 Linear regression model

Linear regression is an approach to modeling the relationship between a dependent variable and one or more independent variables. In a linear regression model, unknown parameters are estimated from the data using a linear function. Often, linear regression is used to test the relationship between a dependent variable Y and values of independent variables X_k (Equation 2.1). Usually, a random error term u is added to the regression model since there are always some random affects that are impossible to forecast. To apply statistical methods for evaluation, the error term u is generally assumed to be normally distributed, uncorrelated, zero mean, and with constant variance [21]. As Equation 2.1 shows, if a value of independent variable X_k is given, the fitted model can be used to estimate the value of the dependent variable Y.

$$Y_k = \beta_0 + \beta_1 X_k + u \,, \tag{2.1}$$

where β_0 is the intercept or constant term, and β_1 is the slope coefficient.

In the case of natural gas forecasting, we can use a simple reference regression model:

$$\hat{S}_k = \beta_0 + \beta_1 * HDD65_k \,. \tag{2.2}$$

β_0 is a constant coefficient known as base load. The base load is everthing not dependent on today's air temperature. β_1 is the coefficient of the heating load which is represented by HDD with reference temperature 65 degrees. The heating load is the energy consumption per unit time that is supplied to maintain a specified temperature [14]. In the GasDay Lab, we are concerned mainly with the forecasting accuracy of the heating load plus the base load.

2.4 Problems of linear regression models

Although linear regression is a good method for forecasting [4], it has some potential problems, which can cause a regression model to give biased and inconsistent results [21], [48]. The problems are often caused by using inaccurate data, using the wrong factors, replacing missing data, or adjusting for seasonality [21]. The problems listed here are not the main concern of this thesis, but they are worth noting if one is going to use linear regression as a forecasting technique.

Heteroscedasticity is one of the common problems that are found in regression. Heteroscedasticity means that the variance of the error changes with time [48]. Heteroscedasticity can lead to inaccurate coefficient estimation and may bias the results. There are several approaches to diagnose if a model suffers from heteroscedasticity. The White test [21] is one of the most common methods used. For a discussion of heteroscedasticity, the readers may refer to [36].

In a regression model, if two or more variables are highly correlated, the model may suffers from multicollinearity [21], which is often detected by applying a Variance Inflation Factor (VIF) test to the regression model. Once the test detects that two variables are highly correlated, multicollinearity can be fixed by removing one or more of the correlated input variables from the model.

Another problem of regression modeling is autocorrelation. It is caused by the forecasting error being related to the measurement at a previous time [21], [36]. This can cause a bias in the data set and problems with statistical tests. To diagnose autocorrelation, one can use the Durbin-Watson test [21]. For further information about autocorrelation and Durbin-Watson test, refer to [21].

With the mathematical tool of linear regression modeling, we are able to forecast natural gas consumption. However, when we consider the models with several weather instruments as independent variables, how should we decide which ones are the right variables to use for forecasting? In the next section, we introduce statistical hypothesis testing, which can help forecasters make decisions.

2.5 Statistical hypothesis testing

Statistical hypothesis testing is a well-known method to help make decisions in the presence of uncertainty with given data. Hypothesis testing can be used either for a controlled experiment or for an observational study [15]. In statistical hypothesis testing, a result can be either statistically significant or insignificant, given the level of significance and the degrees of freedom. A statistically significant result indicates an event that is unlikely to occur purely by chance. In contrast, a statistically insignificant result indicates the event that is likely to occur by chance [31]. In the early 20th century, Ronald Fisher[3] was the first one to uses the phrase "Test of significance." "Critical tests of this kind may be called tests of significance, and when such tests are available, we may discover whether a second sample is or is not significantly different from the first [19]." The primary use of hypothesis testing is to decide whether a pre-determined result contains enough evidence to cast doubt on conventional wisdom. If the statistical result exceeds the critical value[4], the null hypothesis is rejected in favor of the alternative hypothesis; we are prone to believe that the alternative hypothesis is a better representation of the truth. In contrast, if the statistical result is lower than the critical value, we do not have enough evidence to reject the null hypothesis, and we agree with the statement of the null hypothesis.

In the literature, statistical hypothesis testing plays a fundamental role [14], [18], [21]. The usual steps of a hypothesis test are:

1. State a null hypothesis (H_0) and an alternative hypothesis (H_1). Usually, the null hypothesis should be chosen in a way that it allows us to conclude whether the alternative hypothesis can either be accepted or stays undecided as it was before the test.

2. Consider the statistical assumptions being made about the sample. For example, there

[3] English statistician, evolutionary biologist, geneticist and eugenicist, who was described by Anders Hald as “A genius who almost single-handedly created the foundations for modern statistical science”.
[4] In statistics, a critical value usually is a cutoff value that determines the boundary between the samples that leads us to reject the null hypothesis and the samples that do not lead to rejecting the null hypothesis [21].

may be assumptions about the statistical independence or about the form of the distributions of the observations. This is important since invalid assumptions may mislead the results.

3. Decide the appropriate statistical test(s) and state the relevant test(s).

4. Derive the distribution of the test statistic under the null hypothesis from the assumptions. For example, the test statistics may follow a normal distribution.

5. The distribution of the test statistic partitions the possible values into those for which the null hypothesis is rejected, the critical region, and those for which it is not.

6. Based on the observations, calculate the value of chosen statistical test.

7. Compare the value of the test statistic to the given critical value with certain degrees of freedom and significance level. The chosen significance level and the calculation of degrees of freedom are discussed in Chapter 3.

8. Decide whether to reject the null hypothesis in favor of the alternative. We should reject the null hypothesis if the calculated test statistic value exceeds the critical value; otherwise we will not reject the null hypothesis.

In this thesis, our null hypothesis is that a coefficient is not statistically significant. Therefore, when we reject the null hypothesis, in our case, we are prone to keep a variable in our experimental model(s); otherwise we are more likely to drop it. In Chapter 3, we give a concrete discussion of the statistical hypothesis testing that follows the above steps. In the next section, we introduce the weather instruments in which we are interested and our expectation of their effect on the consumption of natural gas based on domain knowledge.

2.6 Introduction of new weather instruments

According to our data source, we list some weather instruments that are not currently used in the GasDay model. A brief background introduction of the weather instruments is given in this section, and the discussion for the significance for each variable is in Chapter 3.

1. Precipitation

To guarantee plants would be harvested on time, some farmers in the southern and western areas of the US use natural gas to power water pumps for irrigation when necessary. Therefore, Precipitation may lead to a decrease of natural gas use. Otherwise, farmers consume more gas to protect plants from drought. An accurate demand forecast of natural gas for LDCs serving an irrigation area is needed. Based on the US Department of Agriculture (USDA)'s Farm and Ranch Irrigation Survey [44], about 56 million acres of US farmland were irrigated with pumps powered by liquid fuels, natural gas, and electricity, costing a total of 1.55 billion dollars. Electricity was the principal power source for these pumps, costing 63.5% of the total to irrigate 43 million acres at an average cost of $39.50 per acre [44]. However, diesel fuel and natural gas were used to power pumps in many areas, about 12 million and 5 million acres, respectively [44]. Accurate forecasting of energy demand in agriculture can save significant operational cost for the LDCs and for the farmers as well. For the non-agricultural customers, even though there is no direct meteorological evidence pointing to an effect of Precipitation on daily gas consumption in previous studies, Precipitation may be an indicator of gas flow changes. Thus, it is worth testing the impact of Precipitation on natural gas consumption.

2. Cloud Cover

Cloud Cover has a huge impact on the temperature, which makes it a potential factor for forecasting natural gas consumption. Cloud Cover may be important because it blocks the sun's heat energy reaching the surface of the earth. Sunlight energy is absorbed by the earth's surface and then is emitted back into the air [12], [37]. Heavy Cloud Cover reflects solar energy back into space or absorbs it. With less sunlight reaching the surface of the earth, the temperature rises slowly. In other words, at the same outside temperature, the inside temperature is actually lower on a cloudy day since the Cloud Cover reflects the solar energy that could warm the buildings. At night, the heat absorbed by earth during the day time continues to be emitted from the surface to the air. If there is no Cloud Cover, this heat rises, leaving the surface cold [12]. Especially on clear nights, we may experience those extremely cold temperatures. However, if we have Cloud Cover at night, the clouds acts like a blanket which keeps the heat between the clouds and the earth surface, so the temperature is warmer the next morning [12].

3. Wind Speed and Wind Direction

The cooling effect of wind is a critical factor on the local temperature. More heat is lost from a building when the wind is blowing hard than on a calm day with the same temperature. The current GasDay model captures the wind effect by Heating Degree Day Wind adjusted (HDDW) (Equation 2.3),

$$HDDW = \begin{cases} HDD\ \dfrac{72 + Wind\ speed}{80}, & Wind\ Speed \geq 8 \\ HDD\ \dfrac{152 + Wind\ Speed}{160}, & Wind\ Speed < 8. \end{cases} \tag{2.3}$$

The HDDW takes the Wind Speed into account. However, another feature of wind, direction, is not represented in the current GasDay model. Because of the effect of wind, cities that are bounded by a large body of water, either a lake or an ocean, are likely to have a different weather condition when the Wind Direction changes. The wind usually blows from the land to water in winter (called the dry phase [52], because it carries cool, dry air), and from water to the land in summer (called the wet phase, because it carries warm, moist air), causing a drastic change of the local climate [52]. We also notice that it is common to build buildings with more windows on the south side and fewer on the north side to take the advantage of passive solar heating. However, when the wind blows on the south side of the houses, windows are likely to exchange more heat from the inside to the outside than the protection of walls. Also, residences often grow trees on the north side of their houses as a natural protection from the cold north wind, but leave the house unprotected at the south to benefit from solar radiation. Therefore, the Wind Direction may be a potential factor that impacts the exchange of thermal energies. In this study, when assessing the effect of Wind Directions, we cannot ignore the effect of the Wind Speed, so we use the product of Wind Direction and Wind Speed as a comprehensive factor to be tested in this paper.

In this study, Wind Direction is reported as degrees from 0 to 360 in the raw data file. It is pointing in the direction that the wind is blowing from. For instance, if the Wind Direction equals 0 degrees, the wind is blowing from north to south. If the Wind Direction equals 90 degrees, the wind is blowing from the east to the west. For this variable, 0 degrees and 360 degrees are two boundary conditions, but it does not necessary mean that 360 degrees of Wind Direction has a larger impact on gas consumption compared to 0 degrees. Hence, we do not use degrees directly, but apply a fuzzy logic method [7], [47] to describe Wind Direction as shown in Figure 2.1. In Figure 2.1, each color represents a Wind Direction. Red is for north, yellow for east, blue for south, and green for west. The horizontal axis represents the actual Wind Direction that is reported, and the vertical line (from 0 to 1) is the logic value range. Each function maps the Wind Direction to a logic value in between 0 and 1. For example, as the vertical line in the figure shows, the 138° of Wind Direction can be transformed to an approximate combination of 0.47 of east wind and 0.53 of south wind, holding the Wind Speed to be constant.

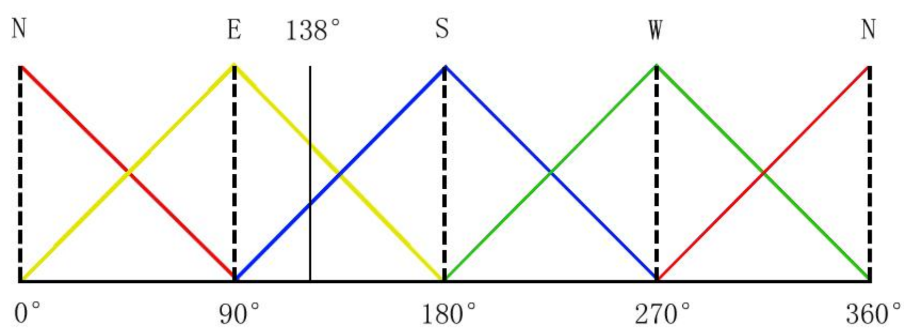

Figure 2.1 Wind Direction is transformed from degrees to numbers using a fuzzy logic method

To better capture the cooling effect of the wind, we consider the product of Wind

Direction and Wind Speed as one variable, which may be a more accurate way to test the significance of the effect of wind on gas consumption.

4. Dew Point and relative humidity

Dew Point is an indicator of saturation temperature, which is associated with relative humidity. At a given barometric pressure, when the Dew Point is close to the air temperature, a high relative humidity is indicated, as shown in Figure 2.2.

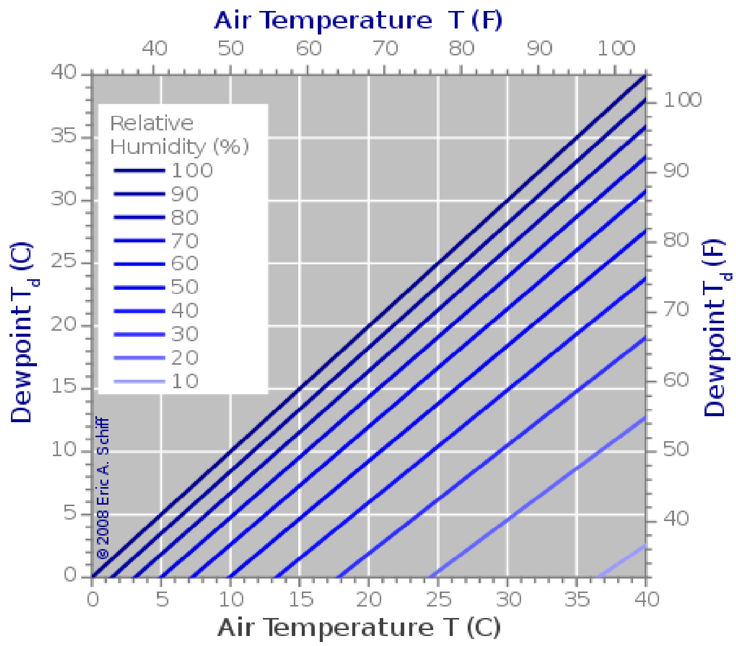

Figure 2.2 Dew Point vs. air temperature at varying relative humidity [49]

A relative humidity of 100% indicates that the Dew Point is equal to the current temperature, and the air is fully saturated with water [34]. If we hold the Dew Point constant and increase the temperature, the relative humidity decreases. As rule of thumb, temperature, Dew Point, and relative humidity are approximately related as [30]

$$Relative\,Humidity = 100 - \frac{25}{9} * (T^{^oF} - T_{dewp}^{^oF}) \,. \tag{2.4}$$

Equation 2.4 indicates that we can expect to catch the humidity by incorporating Dew Point into the existing GasDay model. This approach is accurate to within about ± 1 ℃ as long as the relative humidity is above 50%.

CHAPTER 3

Model with Additional Weather Inputs

Chapter 2 introduced several new weather instruments that may impact daily gas consumption. In this chapter, we discuss the statistical significance of each variable for daily natural gas demand forecasting. To satisfy the need of LDCs, the hypothesis tests in this chapter focus on two operating areas containing two different types of end users of natural gas.

3.1 Data description

In this chapter, we use time series data extending from January of 1996, to March of 2012 for one operating area, and to August of 2011 for another area. We keep the last year of available data as our testing set. All the data from 1996 to the beginning of the test set is our training set. The data come from two primary sources, the GasDay Lab at Marquette University and the National Oceanic and Atmospheric Administration (NOAA) [34]. Weather data is reported in hourly intervals, and it is aggregated to daily intervals to match the daily gas flow that we are going to forecast. In this chapter, we focus on two operating areas: M is a large city in the Midwestern United States; residential customers are the main contributions to consumption in this area. The natural gas delivered by the LDC in this operating area is primarily used for heating purposes. Operating area D is a small city in the southwestern United States, and agricultural processes are the primary consumers to their natural gas. The natural gas delivered by the LDC in this operating area is used primary to power irrigation. Due to confidentiality agreements between LDCs and Marquette University, the exact states and company names are withheld from this paper. The gas consumption data used in this paper have been scaled by an undisclosed scale factor to protect the confidential data.

3.2 Operating area M

New weather instrument candidates are listed in Table 3.1. The single input variables (Precipitation, Cloud Cover, and Dew Point) are used for the purpose of testing the effect of base load. The cross terms are for testing the impact on the heating load.

Table 3.1 New weather instrument candidates

New weather instruments	Unit
Precipitation	Inches/day
Cloud Cover	Percent
Dew Point	Degree
(Precipitation)*HDD65	(Inches/day)*Heating degree
(Cloud Cover)* HDD65	Percent* Heating degree
(Dew Point)* HDD65	Degree*Heating degree
Wind Direction* HDD65	Degree*Heating degree

For each of the new weather instruments we listed above, we applied statistical hypothesis testing, forecast performance testing, and the unusual day [46] evaluation. We discuss the significance of each of these variables for the two operating areas in the following sections. Figure 3.1 shows the steps of the tests we use in this chapter.

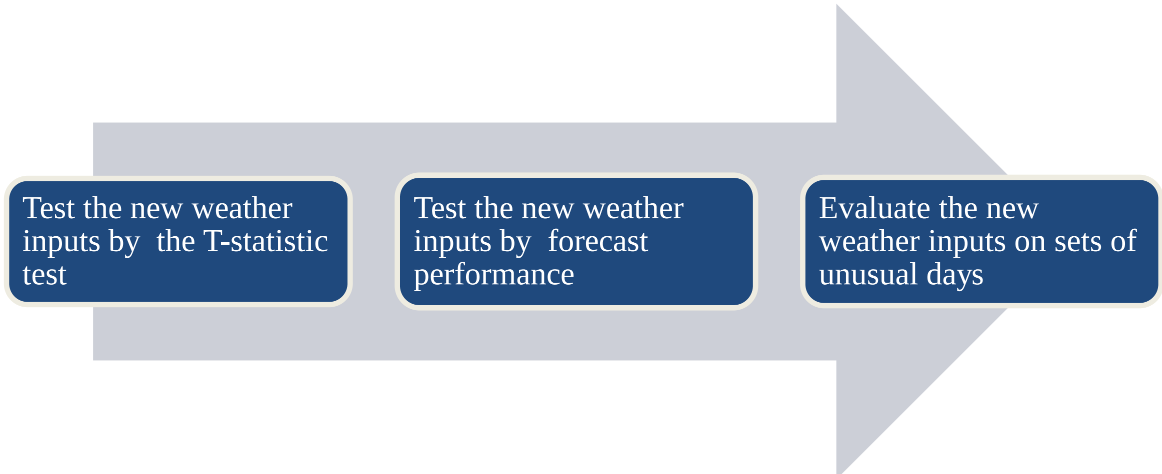

Figure 3.1 Steps of testing

For the statistical hypothesis test, we add all the weather instruments at one time and trim off the variables that are statistically insignificant, which is called a filter method in [39]. In contrast, a wrapper method [39] suggests adding the new inputs into the model one by one. This method is applied during the forecasting performance testing and during the unusual day evaluation.

3.2.1 T-statistic testing for operating are M

Table 3.2 contains the result of T-statistic testing for operating area M. All hypotheses testing in this chapter is done by one-tailed tests at the 5% level of significance (the red cutoff line in Table 3.2), which is the level of significance that Fisher suggested as a limit in judgment [20]. In this case, we are concerned about whether a variable has a positive or a negative impact on gas consumption. Thus, we apply a one tail test. In Table 3.2, at degrees of freedom[5] above 120, the critical value of the T-statistic is 1.645. For each independent variable, if its T-statistic value falls between 1.645 and -1.645, we fail to reject the null hypothesis, and the variable is considered to be statistically insignificant. If the T-statistic value falls outside of this range, we reject the null hypothesis in favor of the alternative hypothesis. The null hypothesis we state in this chapter is that the coefficient is not statistically significant. The alternative hypothesis is that the coefficient is statistically significant.

Table 3.2 T-statistics for operating area M

Independent variable	T-statistic
HDD55	36.33
West Wind*HDD65	17.11
North Wind*HDD65	15.24
Dew Point*HDD65	-13.52
East Wind *HDD65	11.43
Cloud Cover*HDD65	10.73
HDD65	9.17
South Wind*HDD65	8.57
Dew Point	5.83
Precipitation*HDD65	3.25
Cloud Cover	-1.96
Precipitation	0.29

Table 3.2 shows the T-statistic for the new input variables. The variables are sorted by the absolute values of T-statistics. Both HDD65 and HDD55 are statistically significant at any traditional level of significance. This result reinforced our expectation

[5] For the T-statistic, the degrees of freedom equals to the number of observation minus the number of independent variables [21]. In our case, we have more than 5000 observations of daily data.

that, for this operating area, gas consumption is primarily temperature related. In addition, the four variables that represent the wind effect are statistically significant at the 5% level of significance. This implies that the cooling effect of the wind is a significant factor for daily natural gas demand.

Cloud Cover and Dew Point are statistically significant at the 5% level of significance. The minus sign of the coefficients indicates a negative impact of this variable on gas consumption. A low T-statistic (0.29) of Precipitation does not allow us to reject null hypothesis, so we say that this variable is not statistically significant at the 5% level of significance, which indicates that the impact of Precipitation on the base load is not statistically different from zero. No cross term is rejected at the same level of significance. Therefore, Precipitation, Cloud Cover, and Dew Point can help to model the variation of the heating load.

Table 3.3 T-statistic for operating area M after dropped the insignificant variable

Independent variable	**T-statistic**
HDD55	36.64
West Wind*HDD65	16.85
North Wind*HDD65	14.78
Dew Point*HDD65	-13.22
East Wind *HDD65	10.90
Cloud Cover*HDD65	10.88
HDD65	9.06
South Wind*HDD65	8.25
Dew Point	5.71
Precipitation*HDD65	4.10
Cloud Cover	-1.88

From Table 3.2 to Table 3.3, we dropped the insignificant variable Precipitation from our input list. In Table 3.3, none of the variables are insignificant at the 5% level of significance. We name the model that combines by the existing GasDay model and the weather inputs that are statistically significant as the “M model”. For this area, our forecast runs from March/11/2011 to March/10/2012. In this chapter, for both operating areas M and D, we compared the performance of the models according to the performance criteria stated in Chapter 1.

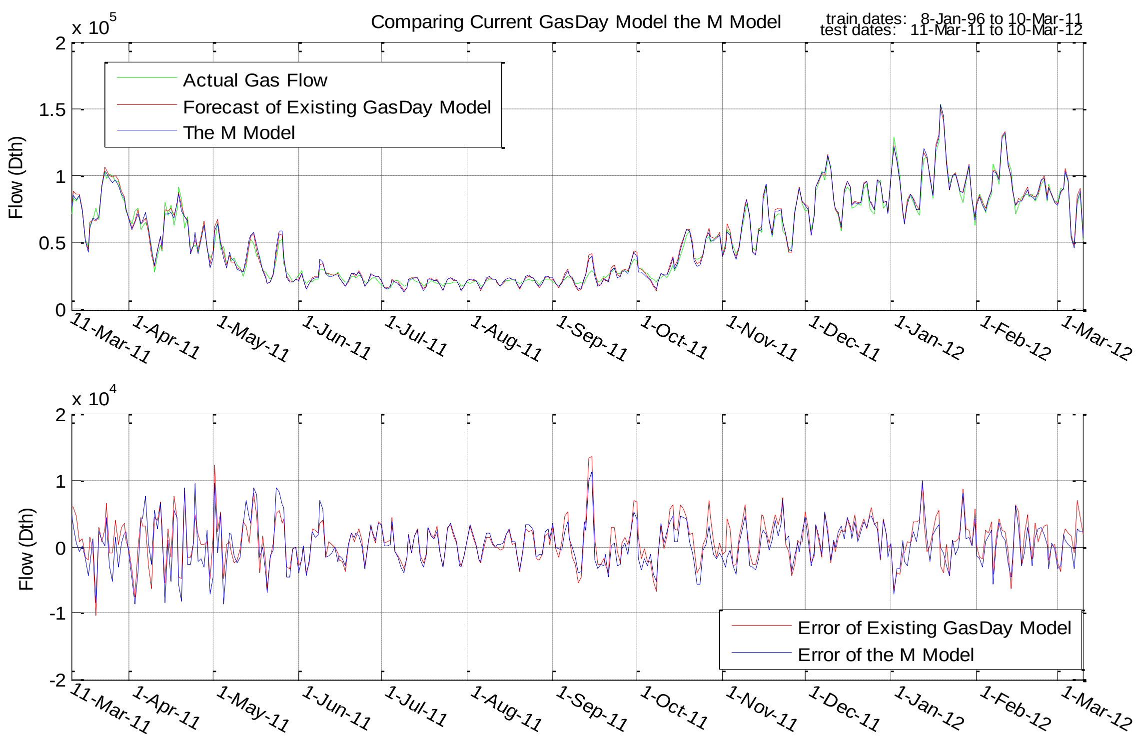

Figure 3.2 The existing GasDay model vs. the M model

Figure 3.2 shows the actual, forecasted GasDay, and forecasted M model gas demand. Additionally we show the residual errors for the GasDay model and the M model. Figure 3.3 shows the RMSE and MAPE measures. Overall, the M model performed close to but not as well as the current GasDay model. For example, MAPE is 6.63% (current GasDay model) vs. 6.74% (M model) on average. For the heating months (especially October, November, and February), the M model outperforms the existing GasDay model by about 500 to 1000 DTh, in term of RMSE. Over the shoulder months and the summer, the performance of our M model is not as effective as the current GasDay model.

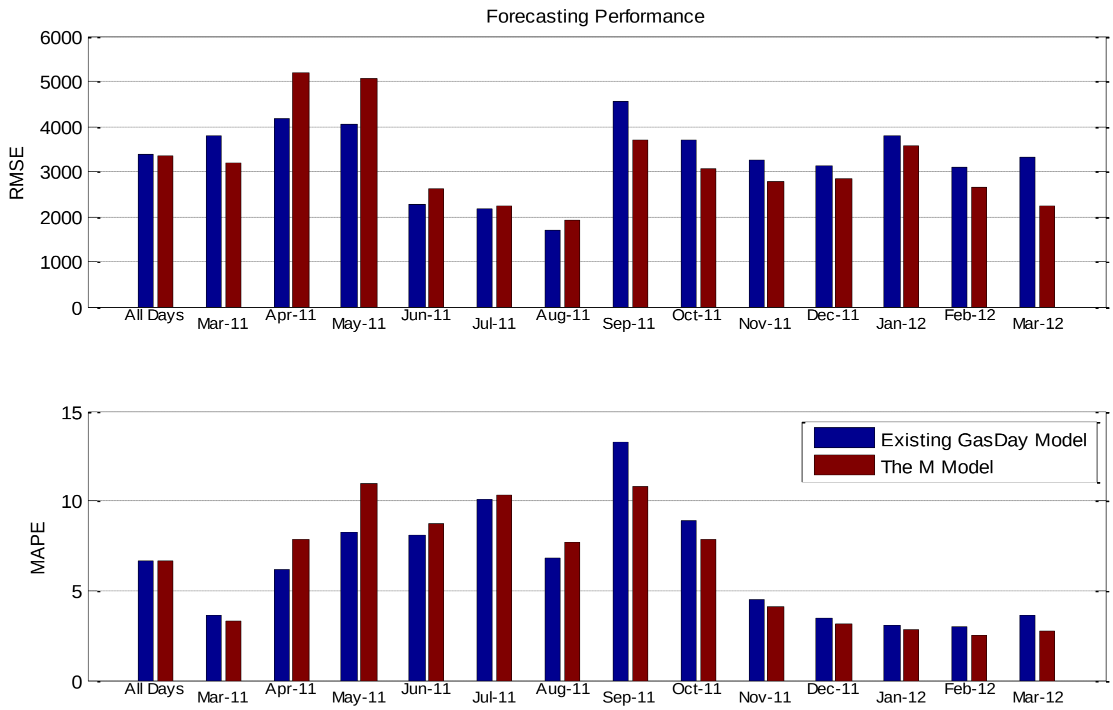

Figure 3.3 Forecasting performance of the M model based on the hypothesis test

For operating area M, we have a preliminary judgment for each weather instrument and for its potential impact on daily gas consumption according to the T-statistical test. In the next section, we will evaluate the significance of new weather inputs by a forecasting performance test.

3.2.2 Forecast performance test for operating area M

Based on the hypothesis testing stated in Section 3.2.1, we identified a set of input candidates according to the T-statistics. In this section, we re-evaluate the input candidates by forecasting performance. Based on the current GasDay model, we add one weather instrument to our existing model at a time and compare the forecasting performance of the existing model against the new model. For example, for Cloud Cover, we compare the performance of the current GasDay model vs. the current GasDay model with Cloud Cover added as additional weather variable. The variables that we test here are the same as we discussed in the previous section. Since we evaluate each weather instrument independently, the orders which we test the variables does not affect our forecasting accuracy. The training set and forecasting set are the same as in Section 3.2.1.

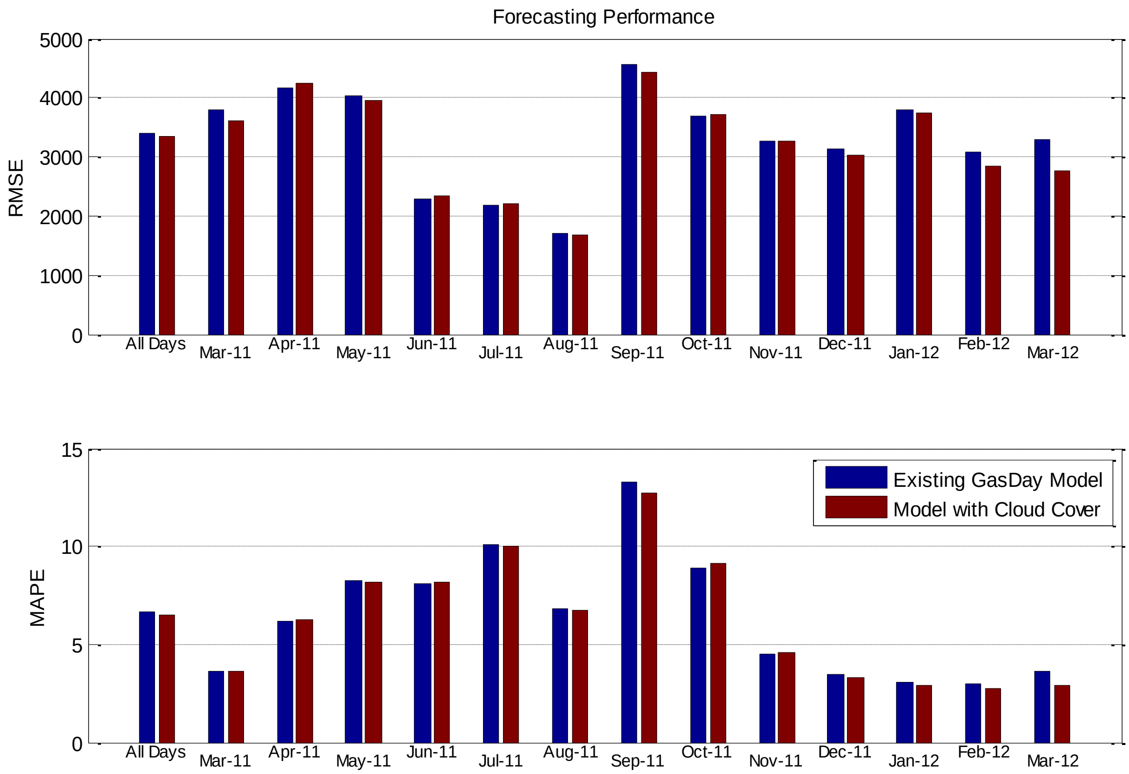

Figure 3.4 Adding Cloud Cover to the existing GasDay model

Figure 3.4 shows the impact of adding Cloud Cover into the existing GasDay model. When we say that we add Cloud Cover as an input variable, mathematically we add Cloud Cover times HDD65 as an indicator of the impact of Cloud Cover on the heating load. The same idea applies to the remaining weather input candidates. By plotting the forecasting error measures in Figure 3.4, we discover that Cloud Cover has a positive impact on improving the forecasting accuracy. The overall performance is very close, with MAPE of 5.04% of the existing GasDay model compared to 4.94% for the existing GasDay model augmented with the Cloud Cover. Significant improvement can be found at February and March, without any deterioration of performance over the testing period. The result matches the T-statistic that Cloud Cover is a significant variable for modeling this operating area.

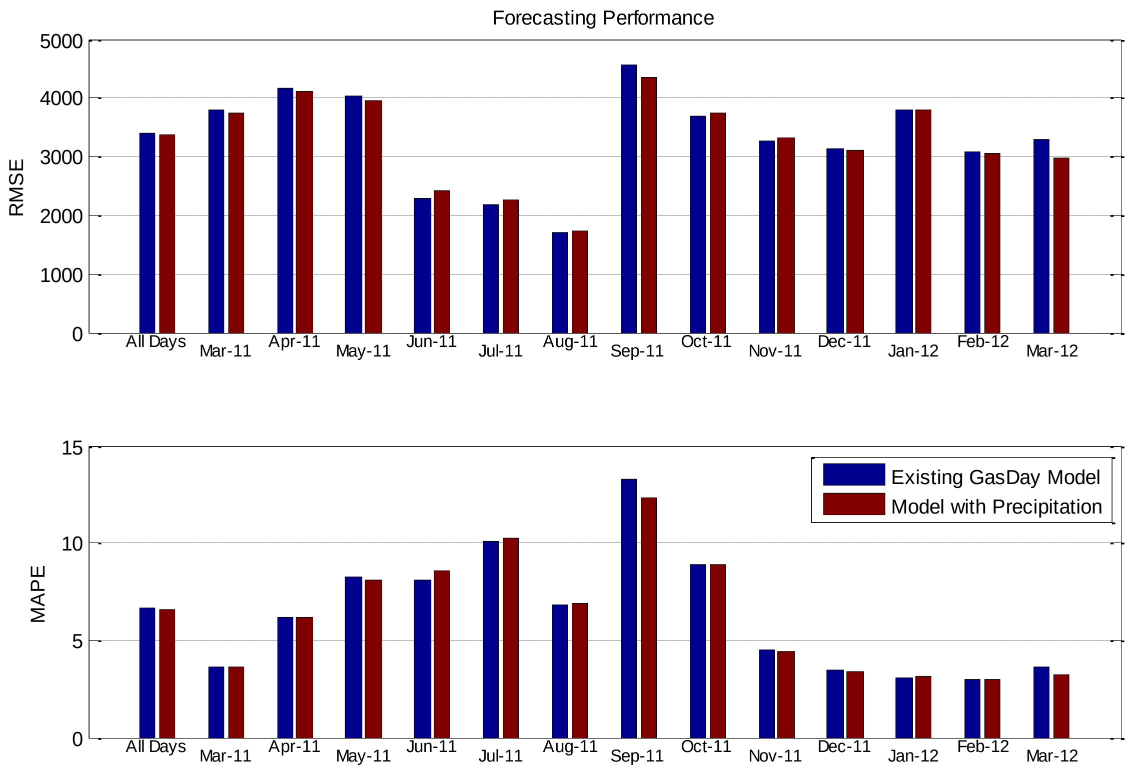

Figure 3.5 Adding Precipitation to the existing GasDay model

Figure 3.5 shows the forecasting performance achieved by adding Precipitation into the existing GasDay model. The results are close for model with Cloud Cover added. Precipitation contributes some marginal improvement overall without losing accuracy during any of forecasting months. Apparent improvement is found in March of 2012, when we lowered the MAPE by about 0.6%, based on a MAPE of 4.07% for the current GasDay model.

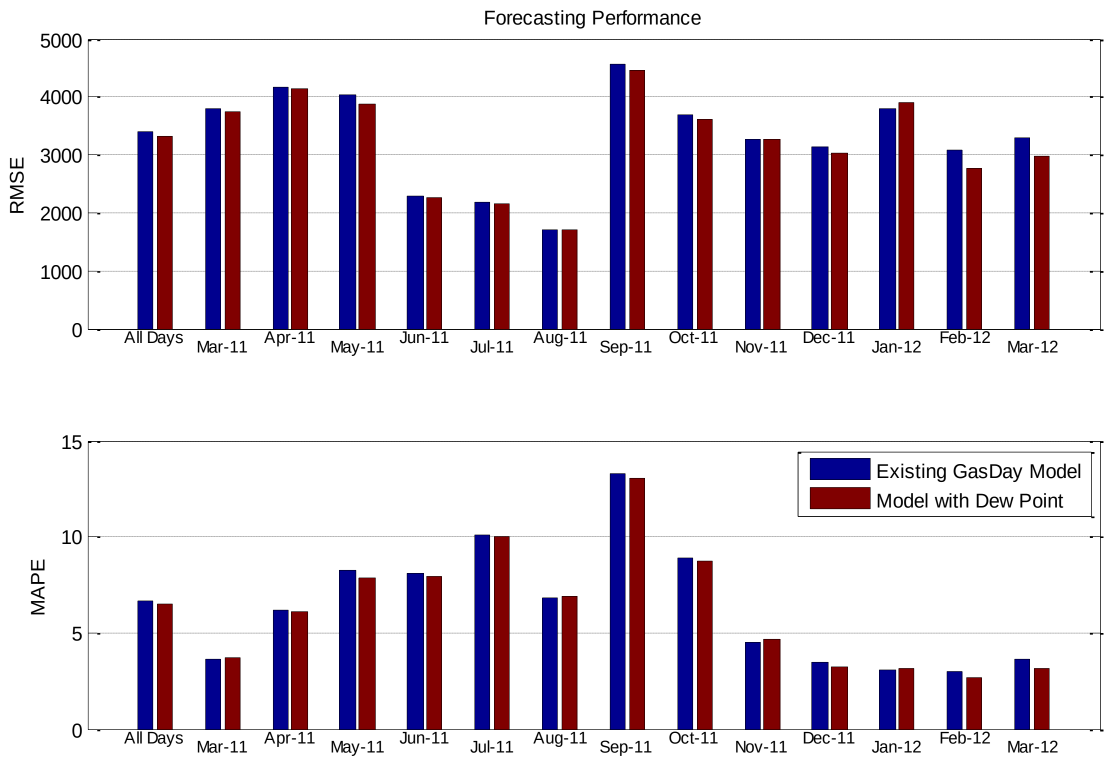

Figure 3.6 Adding Dew Point to the existing GasDay model

Figure 3.6 shows the forecasting performance achieved by the current GasDay model with Dew Point added. Improvements are found during April, May, September, and December of 2011 as well as March 2012. On average, we lowered the MAPE by about 1.1%. Overall, Dew Point provides marginal improvement over the testing set. It makes a positive contribution to the forecast performance as the T-statistic suggests.

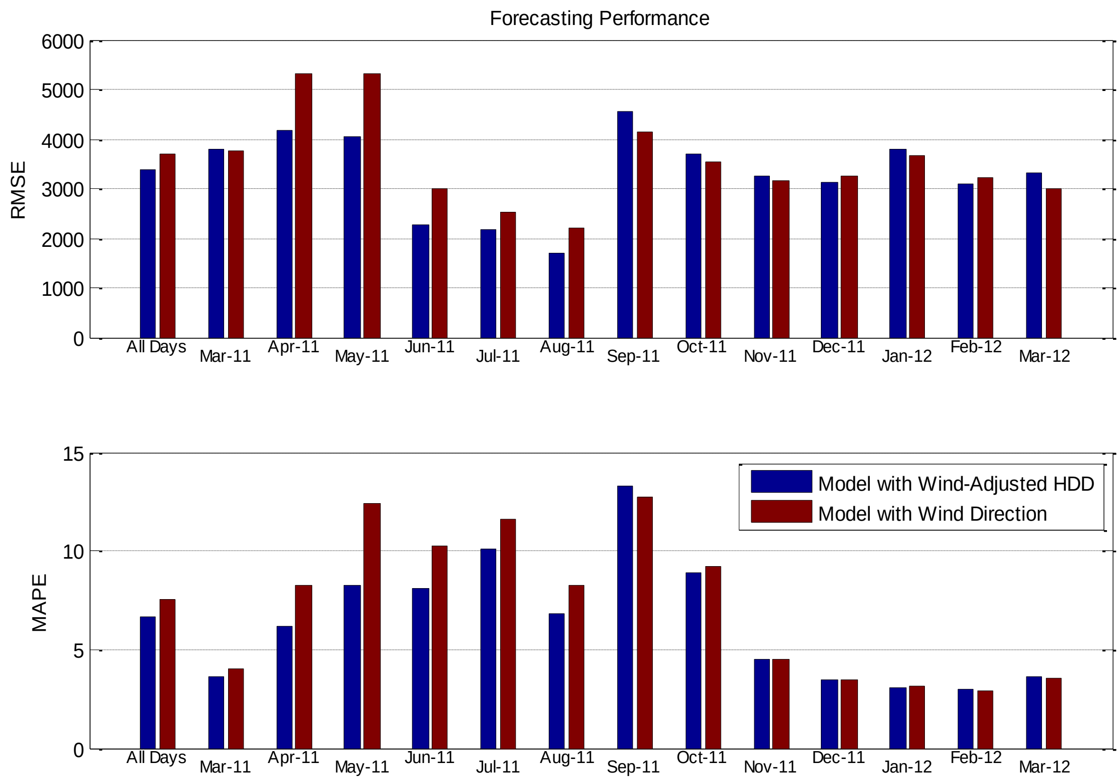

Figure 3.7 Adding Wind Direction vs. the existing GasDay model

Figure 3.7 shows the forecasting performance achieved by adding the wind effect to the existing GasDay model as we proposed in Chapter 2. The current GasDay model represents the wind by wind-adjusted HDDs. In our method, we separate the Wind Speed from the HDDs, and incorporate the Wind Directions as a new vector into the model. In our method, the variables that represent wind effect include: NorthWind*HDD65, EastWind*HDD65, SouthWind*HDD65, and WestWind*HDD65. We expect to improve our forecasting accuracy by a new way of capturing wind information. However, the new method is not as accurate as the existing method. We lost our accuracy by 2% of MAPE and about 400 DTh of RMSE. Using the new method to represent the wind effect occasionally can improve the forecast for a few months. However, it has a relatively large negative impact on the accuracy of the current GasDay model for the shoulder months during this testing set.

Since the Wind Direction variables are not performing as well as we expected, we do not recommend using the alternative method to replace the existing GasDay method of representing wind direction. Figure 3.8 shows the forecast performance achieved by adding Cloud Cover, Precipitation, and Dew Point (represented by “C.P.D” in Figure 3.8) simultaneously to the current GasDay model. The added weather inputs lowered the RMSE from 3400 DTh to about 3230 DTh and MAPE from 6.63% to 6.43% overall.

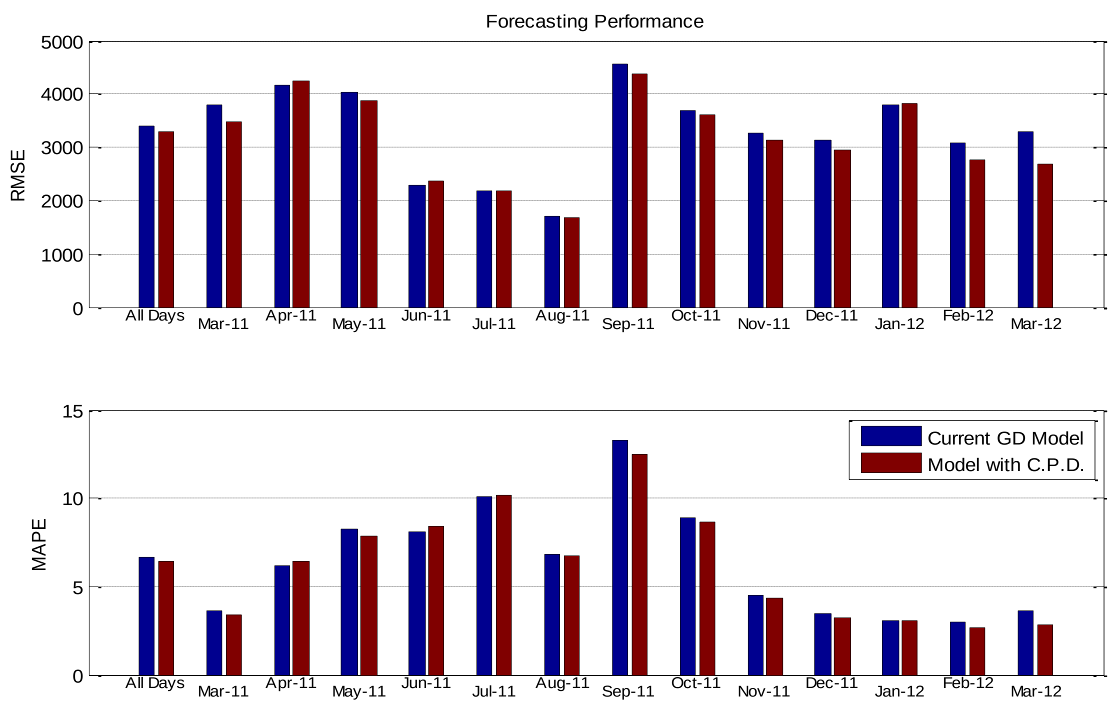

Figure 3.8 Adding Cloud Cover, Precipitation, and Dew Point (C.P.D.) to the existing GasDay model

Based on the forecasting performance tests, we conclude that Cloud Cover, Precipitation, and Dew Point can improve forecasting accuracy both individually and jointly. Wind Direction does not help to lower the error as expected.

According to our analysis, some of our additional weather inputs improved our existing model's accuracy, but others did not improve forecasts. For the convenience of readers, we provide a summary of our results of this section in Table 3.4. In this table, each row represents a model that is constructed from the existing GasDay model and the variables that are listed in the first column. The second and third columns represent the overall RMSE and MAPE values, which are the same values as the bar "All Days" shown in Figures 3.4 to 3.8. The fourth and fifth columns are the actual improvement based on the current GasDay model (RMSE 3400 DTh, MAPE 6.63%) during the same testing period. Cloud Cover, Precipitation, and Dew Point have improvement on both RMSE and MAPE independently and jointly. Wind Direction, however, is not as helpful as we expected.

Even though the results are very close, a student t-test [50] shows that each model, except the model with Wind Direction, has statistical significant lower residuals against the existing GasDay model at the 5% level of significance. And overall, an f-test [51] suggests that at the same level of significance, model adding with Cloud Cover, Precipitation, Dew Point, and model adding with joint of these three inputs fit the training set significantly better than the current GasDay model. The f-test is invalid when evaluate the model add Wind Direction, because the reference model is not 'nested' within the model with Wind Direction [51]. As for the goodness of fit, similar information can be obtained from the adjusted R^2. In this thesis, we are more focusing on the comparison of forecasting performance.

Table 3.4 Summary of forecasting performance

Current GasDay Model +	RMSE (DTh)	MAPE (%)	RMSE Improved (%)	MAPE Improved (%)
Cloud Cover	3300	6.53	2.5	1.5
Precipitation	3360	6.56	1.2	1.1
Dew Point	3320	6.50	2.4	2.0
Wind Direction	3690	7.55	-8.5	-13.8
C.P.D.	3230	6.43	5.0	3.0

In the next section, we will evaluate the performance of new weather input candidates on unusual days. Any new weather input that can significantly (not only statistically) improve the forecasting on the unusual days should also be considered as a strong candidate to be added to the current GasDay model.

3.2.3 Evaluate the new weather inputs by the forecasts performance on the unusual days

Previously, we looked at the forecasting performance for an entire test period. With natural gas demand forecasts, we also are interested in the forecasting accuracy during the unusual events. Accurately estimating values for the unusual days is very important to GasDay. We will compare the performance of the models on 12 categories of unusual days according to the performance criteria stated in Chapter 1.

In Figure 3.9, each group of bars represents one category of unusual day. Based on the unusual events, our unusual days are: coldest day, colder (warmer) than normal days, windiest heating day, colder (warmer) today than yesterday, the first cold (warm) days, high (low) humidity heating days, and sunny (cloudy) heating days. We apply the same idea of adding one weather input to the model at a time as in Section 3.21, but we are more focused on the unusual day rather than a times series of the entire testing data set. Each colors of bar represents a model. For example, the dark blue bar (first column) represents the performance of the existing GasDay model, while a light blue bar represents a model that adds Cloud Cover to the current GasDay model.

Based on the unusual day evaluation, our weather input candidates perform differently on different types of unusual days. Wind Direction drops the RMSE by about 100 DTh on the coldest days and on the first warm day, but for the other types of unusual days, Wind Direction does not help. For the colder than normal and warmer than warmer days, Cloud Cover, Precipitation, and Dew Point all provide slight improvements. The models that include all three variables perform the best. For the windiest heating days, the Wind Direction variable does not improve our forecast, but the Cloud Cover can help drop MAPE by 0.61%. For the days that are colder (warmer) than the day before, the model that includes Cloud Cover, Dew Point, and Precipitation performs the best. The model including Precipitation is most accurate on the first cold (warm) days. Precipitation is also helpful on the high (low) humidity heating days, but the model including Dew Point is most accurate on the unusual days that are humidity related. Cloud Cover was expected to be helpful on the sunny (cloudy) heating days. In our test, no variable improved forecasting on the cloudy heating days significantly. However, Cloud Cover offers significant improvement at the sunny heating days.

Considering the T-statistics, forecasting performance over time, and the forecasting on the unusual days, we suggest adding Cloud Cover, Precipitation, and Dew Point into the existing GasDay regression model. We should expect accuracy improvement by adding them either jointly or independently. However, we do not recommend using our method to represent the wind effect as a replacement of the current wind-adjusted HDDs. Alternatively, GasDay might adopta rule-based weight ensemble technique for unusual days, as a means to forecasts on unusual days without loosing accuracy on normal days.

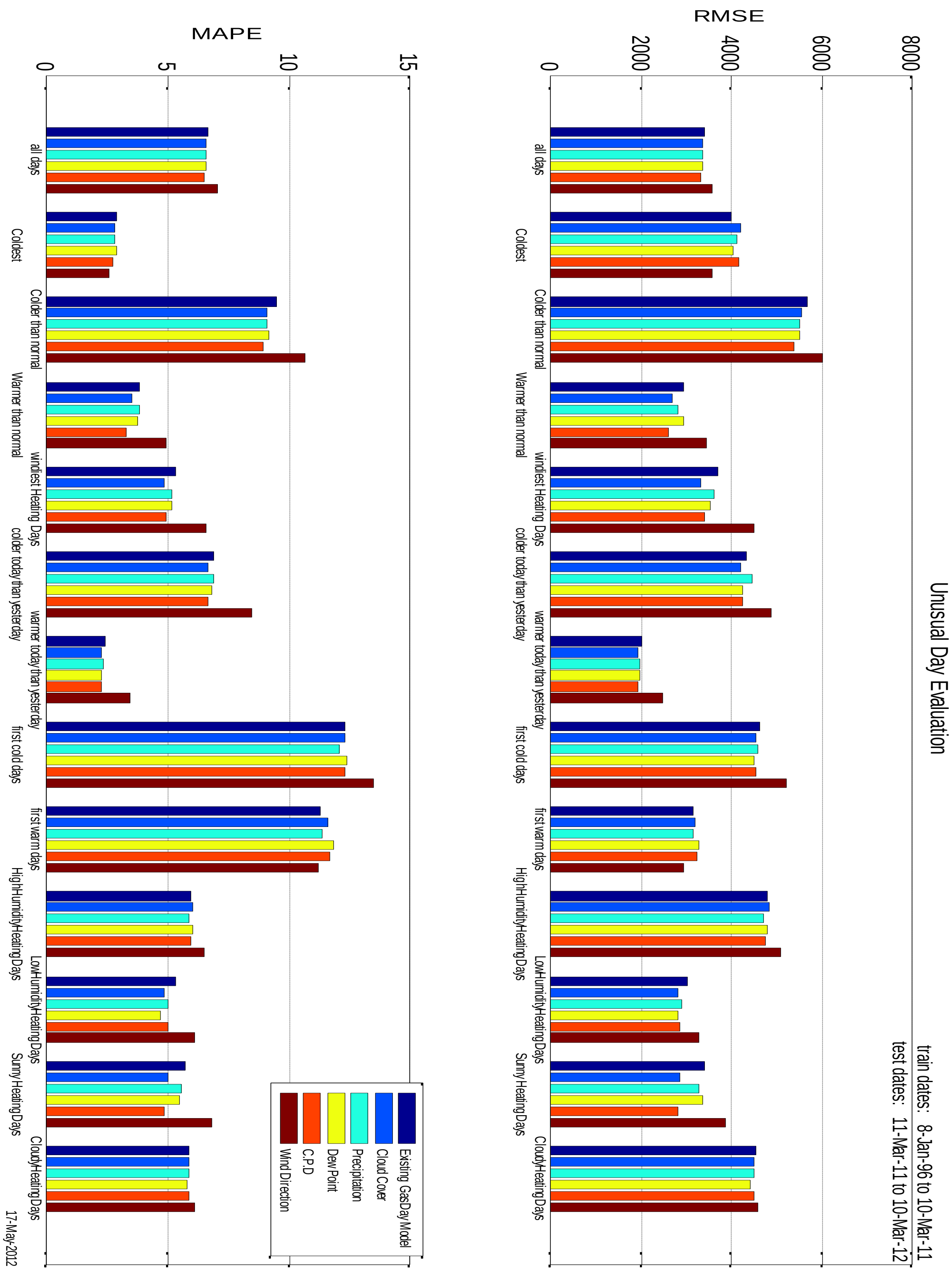

Figure 3.9 Unusual day evaluation of operating area M

3.3 Operating area D

In the previous section, we discussed the new weather inputs for operating area M. In this section, we use the same approach to build models for operating area D. Our goal is to find significant inputs for this area and then obtain an improved forecast. Unlike operating area M, operating area D is in the southwestern U.S., surrounded by desert (Figure 3.11 gives an example of geographic features of such area). Natural gas in this area is used primarily as an energy source to pump water to irrigate crops.

Figure 3.10 A typical irrigation region in a dessert area [33]

During the growing season, when the weather is sunny and relatively dry, farmers might consume more natural gas to power water pumps. However, when it is raining and relatively damp, farmers do not need to irrigate. Therefore, for this operating area, we conjecture that the daily gas consumption is related to Precipitation. For this area, our experience is that the traditional model has unexpected large errors since both HDD65 and HDD55 are almost zero in the summer, which is more than half of the testing set. However, it is in this interval that the gas flow has the largest variation, which needs to be captured accurately. The current GasDay model for this area does not include the HDDs terms. However, the CDD65 is still valid since it represents the cooling consumption that usually common in southern areas. Another significant input variable in the current GasDay production model is second order lagged gas flow, S_{k-2}, representing the gas flow reported two days ago. Since S_{k-2} is the latest available gas flow, one day ago flow is not included in our model. We also embedded other variables in our current production model, for example, HDD of the previous day and day of week/year factors.

3.3.1 T-statistic testing for operating area D

For this operating area, we start our hypothesis testing by adding Cloud Cover, Precipitation, and Dew Point to the existing GasDay model. The wind effect has been emphasized in the northern operating areas (such as area M) to better represent the impact of wind chill on the heating load. For operating area D, gas demand is primarily for irrigation and cooling, so we do not test the wind effect here. Table 3.5 shows the T-statistic for the new weather instruments. The variables are sorted by descending order of absolute values of the T-statistics. All variables are found to be statistically significant.

Table 3.5 T-statistics for operating area D

Independent variable	T-statistic
CDD65	15.35
Dew Point	-6.13
Cloud Cover	4.42
Precipitation	-2.10

According to Table 3.5, Cloud Cover is statistically significant at the 5% level. Precipitation and Dew Point are also found to be significant at the same level. Both Precipitation and Dew Point are found to negatively impact the gas demand of this operating area. CDD65 is not one of the new input instruments. We listed the T-statistic of CDD here to test whether the cooling consumption has a significant impact on gas demand over the summer. The high T-statistic tells us this is a critical variable. Another input that is not shown in this list is the two days ago gas flow. This variable has a T-statistic higher than 50 based on the tests of both the author and GasDay. This lagged gas flow term plays a very important role in the gas forecasting of this area.

By adding the three new weather inputs in the existing GasDay model, we arrive at a “D model” for operating area D. The performance of the current GasDay model vs. the D model is shown in Figure 3.11.

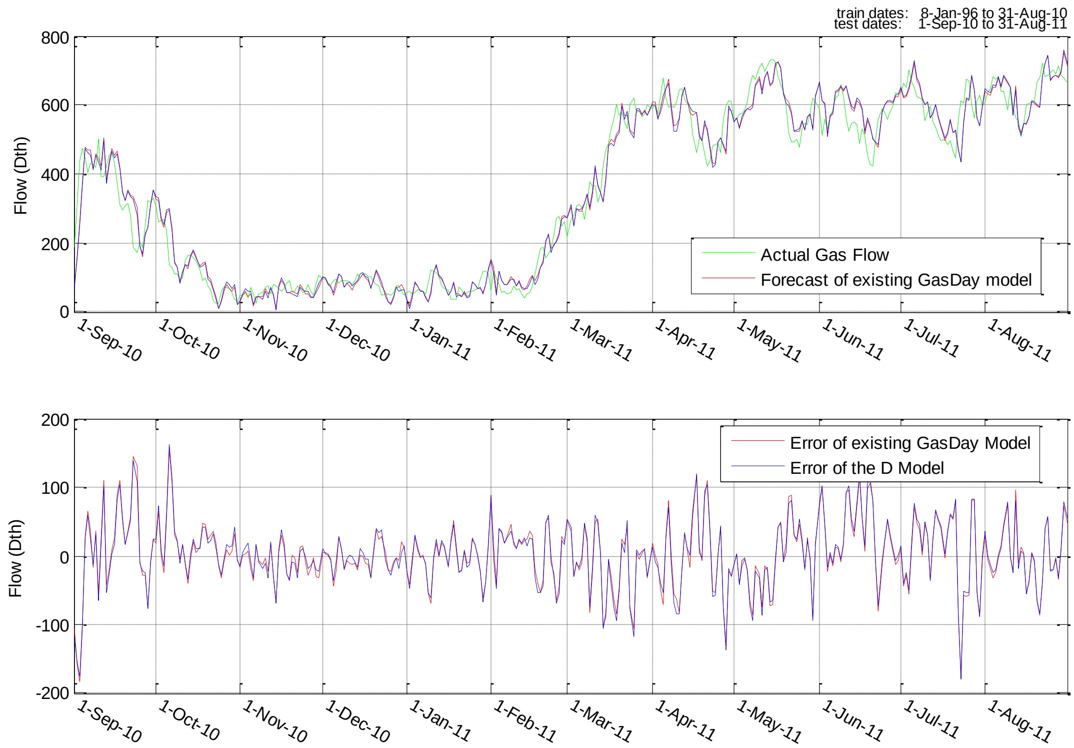

Figure 3.11 The existing GasDay model vs. the D model

Figure 3.11 gives an overview of the gas demand of operating area D over a one year testing period. The D model refers to the model built by adding the statistically significant weather inputs to the existing GasDay model. Unlike area M, the peak flow occurs during the summer months instead of during the winter months. Figure 3.12 shows the RMSE and MAPE measures. Overall, the D model performed close to, but not as well as, the current GasDay model. There is no significant improvement in any of the testing months.

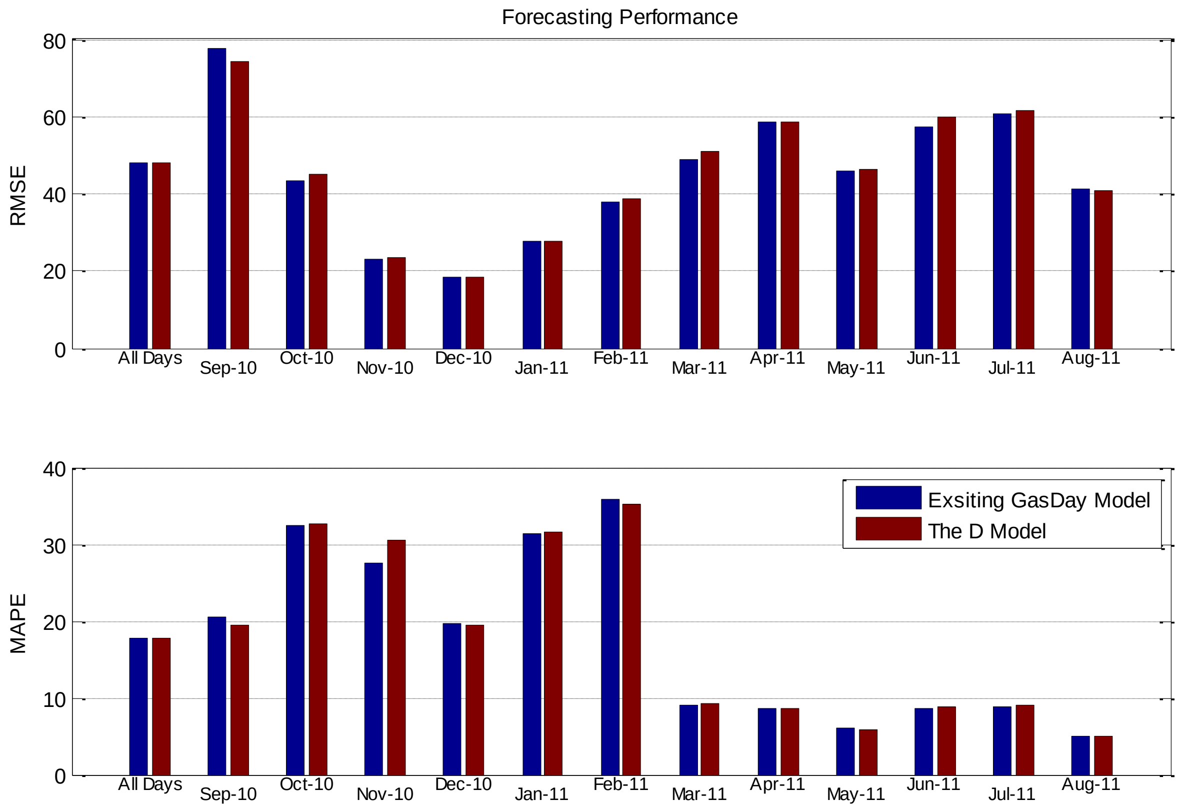

Figure 3.12 Forecasting performance of a D model based on T-statistic

So far, we have analyzed the operating area D from a statistical point of view. The hypothesis testing gives us a preliminary judgment for each weather instrument and for its potential impact on daily gas consumption. In the next section, we will evaluate the significance of new weather inputs using a forecasting performance test.

3.3.2 Forecast performance test for operating area D

Based on the hypothesis testing stated in Section 3.3.1, we identified a set of input candidates according to the T-statistics. In this section, we re-evaluate the input candidates using forecasting performance. Readers can refer to Section 3.2.2 for the details of this method. The variables that we test here are the same as we discussed in

Section 3.3.1. Since we evaluate each weather input independently, the order of testing variables does not affect our results. The training set and testing set are the same as in Section 3.3.1.

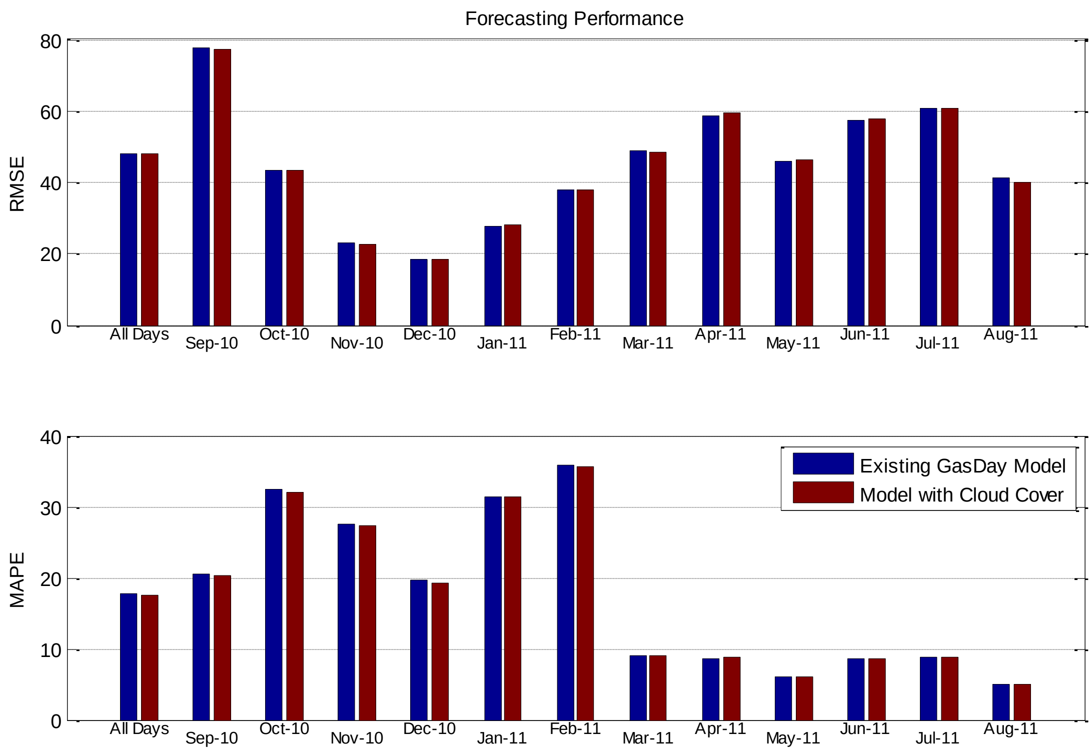

Figure 3.13 Adding Cloud Cover to the existing GasDay model

Figure 3.13 shows the impact of adding Cloud Cover into the existing GasDay model. The overall performance is almost identical.

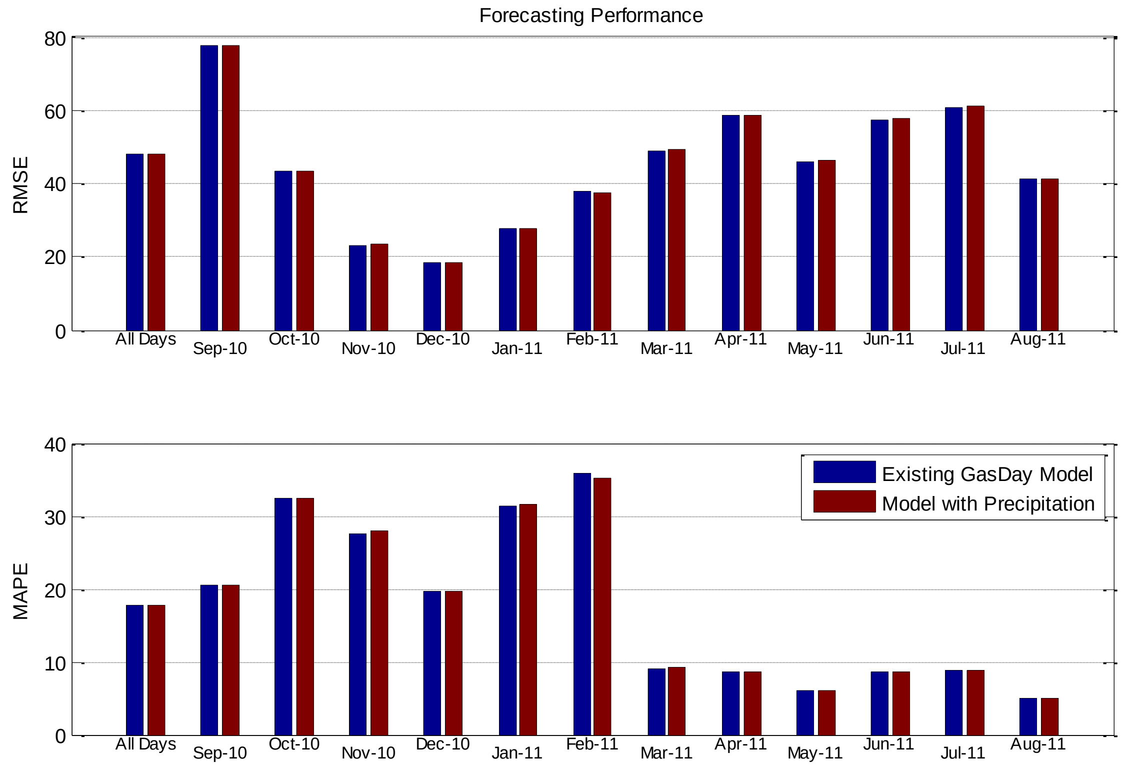

Figure 3.14 Adding Precipitation to the existing GasDay model

Precipitation and Dew Point are the inputs that we proposed to help improving forecast accuracy for this operating area. However, Figures 3.14 and 3.15 show us that their contribution is very limited. By adding Precipitation to our existing GasDay model, the forecast does not improve. Adding Dew Point can lower the error during November and December (Figure 3.15), but for the remainder of the testing set, it does not significantly change the accuracy of the existing GasDay model.

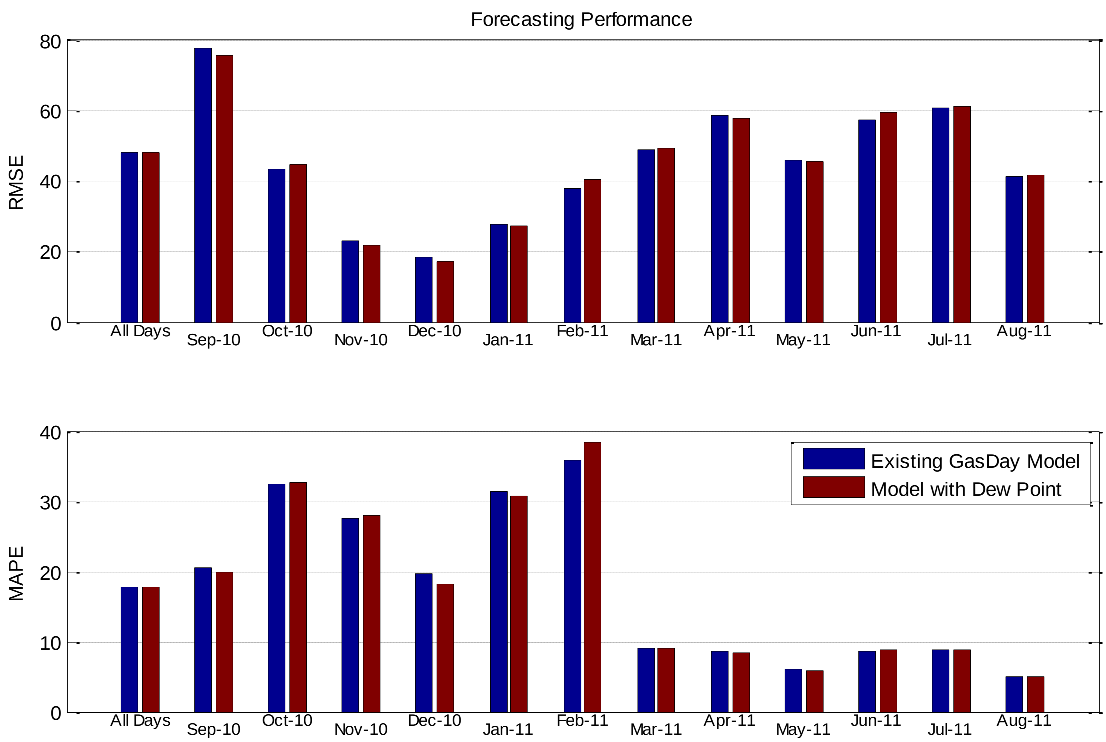

Figure 3.15 Adding Dew Point to the existing GasDay model

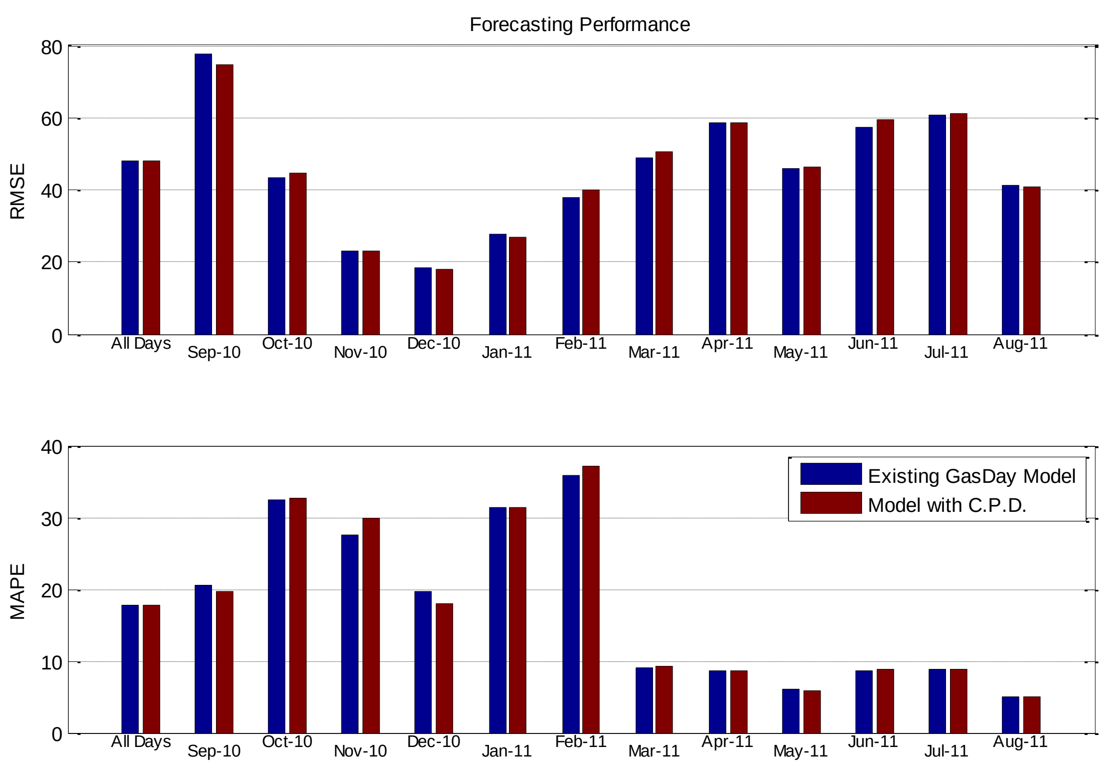

Figure 3.16 Adding Cloud Cover, Precipitation, and Dew Point (C. P. D.) to the existing GasDay model

The statistical testing suggests adding all the weather inputs to our existing model. However, unlike operating area M, this area is unlikely to benefit from adding new weather inputs. We give a summary of our results in Table 3.6. In this table, each row represents a model constructed from the existing GasDay model and the variable that is listed in the first column. The second and third columns represent the overall RMSE and MAPE value, which is the same value as "All Days" shown in Figures 3.13 to 3.16. The fourth and fifth columns show the improvement based on the current GasDay model (RMSE 48 DTh, MAPE 17.7%) during the same testing period. None of the added weather inputs improve our forecast accuracy for this operating area.

Table 3.6 Summary of forecasting performance

Current GasDay Model +	RMSE (DTh)	MAPE (%)	RMSE Improved (%)	MAPE Improved (%)
Cloud Cover	48	17.60	0	0.6
Precipitation	48.1	17.70	-0.2	0
Dew Point	48	17.70	0	0
C.P.D.	48.2	17.80	-0.4	-0.6

In the next section, we will examine the performance of the input candidates on the unusual days. Any new weather input that can significantly improve the forecasting on the unusual days should also be considered as a strong candidate to be added to the current GasDay model.

3.3.3 Unusual days evaluation for operating are D

According to the unusual day evaluation, Precipitation is not as helpful as the T-statistic suggested. It does not provide significant help for any of the unusual day types. Cloud Cover, as another input candidate, does not improve the result as well. Overall, these two variables have no significant contribution to the existing model in terms of either the usual days forecast or the unusual days forecast. Dew Point, as an indicator of humidity, helps lower error on the colder (warmer) today than yesterday and on the low humidity heating days, even though its contribution is very limited. Overall, the model with Dew Point has more poor performance than good performance according to the unusual days. We do not recommend adding this input in to this area. Since this is not a typical customer for GasDay, our evaluation for this operating area does not represent other LDCs and operating areas.

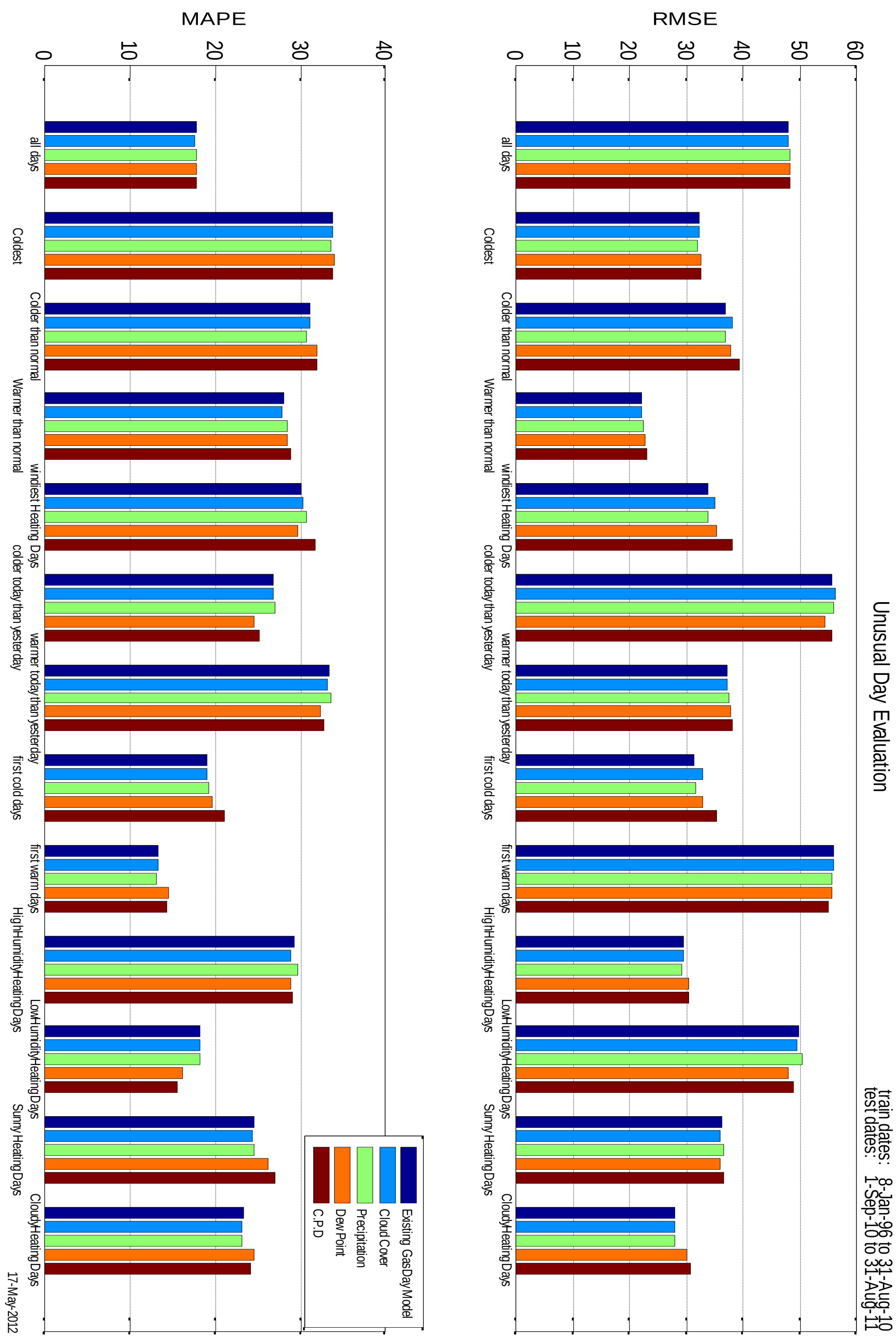

Figure 3.17 Unusual day evaluation for operating area D

Overall, no weather inputs provide significant help. Based on our evaluation, Operating area D’s gas consumption is not primarily temperature, Precipitation, or even weather related. However, for the research for this operating area of this paper, we end our discussion here. The possibilities for further research on this problem are stated in Chapter 5. In the next chapter, we are going to discuss the opportunities to improve our forecast in terms of the HDD65, HDD55, and CDD65.

CHAPTER 4

Model with Multiple Weather Stations

In Chapter 3, we discussed the possibilities of lowering our forecasting errors by including new weather instruments. We found that selected additional weather variables help to improve the accuracy of gas demand forecasting, at least in some areas. However, for most of our customers, natural gas use primarily depends on the temperature. Forecast the heating load is still the main goal of GasDay. In this chapter, we focus on HDD65, HDD55, and CDD65, looking for opportunities to improve the performance of our model in terms of these variables.

4.1 Current GasDay method and its limitations

As we stated in Chapter 1, some LDCs provide their services for a large geographic area where a single weather station's data are not enough to represent all weather characteristics of such an area. That is why GasDay supports weighted multiple weather stations in one operating area. The current GasDay solution is to divide a single geographic area into multiple small operating areas based on the gas flow data reported by the LDCs. However, occasionally, some LDCs only have single time series of daily gas flow data reported, which means their service areas are almost impossible to decompose. Figure 4.1 shows how GasDay handles operating areas which have multiple weather stations. In this chapter, we focus on one LDC which services a large geographic area we denote as operating area C.

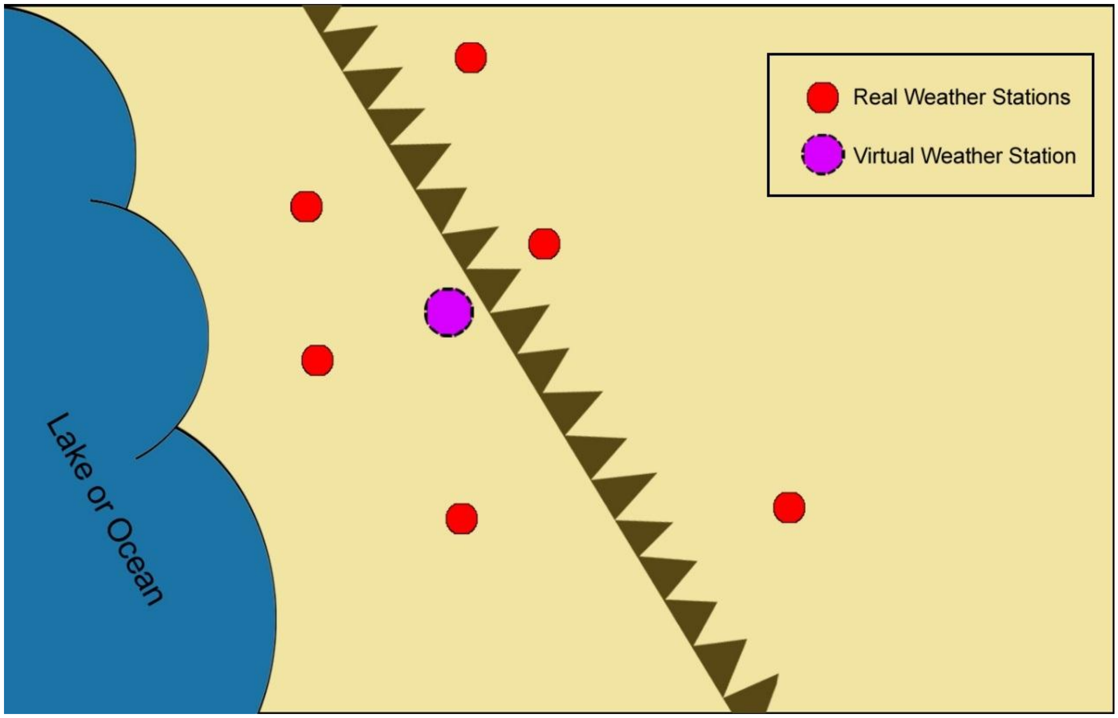

Figure 4.1 Distribution of weather stations in operating area C

For the operating area suggested in Figure 4.1, our current forecasting model uses six weather stations that are selected by the LDC. The current GasDay method uses a weighted combination of data from these six weather stations to obtain a combined temperature (shown in Equation 4.1). To accomplish this, we first calculate the weighted average of the actual temperature for the six weather stations as the temperature of a virtual weather station,

$$T_{(virtual)k} = \sum_{i=1}^{6}(T_{(WSi)k} * w_i)\,, \tag{4.1}$$

where $T(_{WSi)k}$ refers to the daily average temperature of the i^{th} weather station on the k^{th} day. w_i is the weighting factor from the current GasDay model with $w_i \geq 0$, and $\sum_{i=1}^{6} w_i = 1$.

Based on the temperature of the virtual weather station $T_{(virtual)}$ on the k^{th} day, we can calculate HDD65$_{(\text{virtual}),}$ HDD55$_{(\text{virtual})}$, and CDD65$_{(\text{virtual})}$. The current GasDay regression model is based on the variables shown in Equations 4.2 to 4.4.

$$HDD65_{(virtual)(k)} = \text{Max}(65 - T_{virtual(k)}, 0) \tag{4.2}$$

$$HDD55_{(virtual)(k)} = \text{Max}(55 - T_{virtual(k)}, 0) \tag{4.3}$$

$$CDD55_{(virtual)(k)} = \text{Max}(T_{virtual(k)} - 65, 0) \tag{4.4}$$

Reference model: the current GasDay base model: uses a virtual combination of weather stations:

$$\hat{S}_k = \beta_0 + \beta_1 * HDD65_{(virtual)k} + \beta_2 * HDD55_{(virtual)k} + \beta_2 * CDD65_{(virtual)k} + \beta_n * GDinput\,. \tag{4.5}$$

where *GDinput* represents the additional variable(s) that are used in the current GasDay regression model.

The current method of forming a virtual weather station causes a few problems. First, the virtual weather station is an optimal combination of the existing real weather stations, but it is not an actual station. The accuracy of forecasts depends on the computed weights. We may have problem when cold weather come across the region or weather stations comes from an unusual directions. Second, each real weather station (red dot in Figure 4.1) has its unique characteristics and temperature sensitivity, which depends on geographic locations and climates. As shown in Figure 4.1, some weather stations may be close to the ocean, while others are on the back side of mountains facing the mainland. If we simply combine them into one virtual weather station, the characteristics and sensitivity are hidden, and some weather volatility is lost. Finally, the method that calculates the HDDs may cause a problem when we combine the weather stations. For example, suppose we use HDD65 for an operating area with two weather stations. As long as one weather station has temperature above $65^{o}F$, and the other one is below $65^{o}F$, it is possible that the weighted combination of the temperature of the two weather stations equals $65^{o}F$. Then, according to Equation 4.3, HDD65 for the virtual weather station is 0 *HDD*. But is it really zero?

Consider a concrete example to illustrate this problem. Assume that we have two actual weather stations (A and B) for an operating area, and we provide a forecast based on the virtual weather station built on A and B using weights 50% and 50%. Suppose the daily average temperature of weather station A is $60^{o}F$, and the daily average temperature of weather station B is $70^{o}F$. The average temperature of A and B is $65^{o}F$. According to Equation 4.2, $\text{HDD65}_{(\text{virtual})}$ is 0 *HDD*. Comparing this result to Table 4.2, we can see the difference.

Table 4.1 Calculation of HDD65 for two weather stations

Weather station	Actual Temperature	Average Temperature	Average HDD65
A	$60\,^oF$	$65\,HDD$	$0\,HDD$
B	$70\,^oF$		

In contrast to the current method shown in Table 4.1, Table 4.2 tells us that if we calculate HDD65 for weather station A and B separately, the results are 5 *HDD* and 0 *HDD* respectively. If we average them, we get 2.5 *HDD* for HDD65 of the virtual weather station, contrasting with zero shown in Table 4.1. This 2.5 degree difference can lead to significant forecasting errors and additional costs for gas utilities and for their customers.

Table 4.2 Calculation of HDD65 for two weather stations

Weather stations	Actual Temperature	HDD65	Average HDD65
A	$60\,^oF$	$5\,HDD$	$2.5\,HDD$
B	$70\,^oF$	$0\,HDD$	

This example tells us that we need to be careful when combining multiple weather stations into a virtual weather station; the virtual weather station may simplify the problem, but it may lead to an error especially in difficult to forecast shoulder months. The forecast is inaccurate when we consider HDD65 to be zero, but actually it is not. The same issue occurs when we calculate HDD55 or CDD65.

4.2 MWS Model with multiple weather stations

In contrast to the current method, Equation 4.5 shows a new method for forecasting daily gas consumption on the k^{th} day. We extended the weather inputs by applying HDD65, HDD55, and CDD65 for each weather station as independent variables (Equation 4.5). We call this the Multiple Weather Stations (MWS) model.

$$\hat{S}_k = \beta_0 + \beta_1 * HDD65_{k(WSi)} + \beta_2 * HDD55_{k(WSi)} + \beta_3 * CDD65_{k(WSi)} + ... + \beta_n * GDinput \,, \quad (4.6)$$

where WS_i is the i^{th} weather station of this operating area; i = 1 to 6, which is the actual number of weather stations that the LDC used in this operating area. If we use the MWS model, weather stations are not combined, and the potential issues described in Section 4.2 may be avoided. Before we compare of the results, we discuss the data from the area we will use in this chapter.

4.3 Data description and conditions

The data are used in this chapter are daily time series from the GasDay Lab at Marquette University. The testing data sets are from 2009 to 2011. The training set starts from the beginning of the current GasDay data base, March of 1998, for this operating area. To capture seasonal effects, we give the error estimate by calendar months, which will allow us to track the performance of our model month by month.

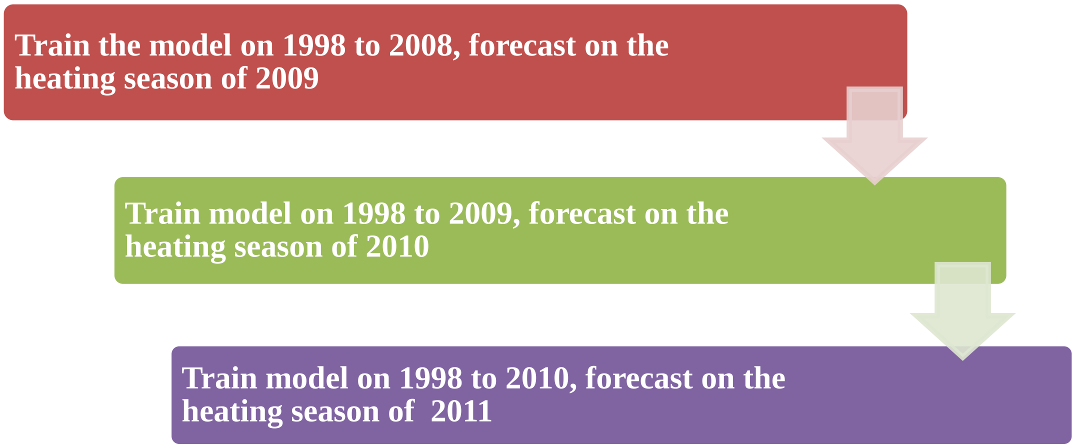

Figure 4.2 Training sets and forecasting sets

Equation 4.5 proposes the MWS model. Of course, this method has some constraints. First, the possible errors of the current method addressed by the MWS model occur when the actual temperature is close to the reference temperature, $65\,^{o}F$ for most LDCs. This means that we are expecting the largest improvement of our forecasts in the testing period with the temperature around the reference temperature. Hence, if $65\,^{o}F$ is our reference temperature, we are expecting the MWS model at least to improve the forecast in the shoulder months. In gas forecasting, the demand usually peaks during the winter (December to February) months of the year, and lowest over the summers (June to August). “Shoulder months” usually refer to the months in either the spring or the fall, which are the months that gas demand falls between the low level (summers) and the high level (winters). Over the shoulder months, gas consumption is not typically heating driven and often is difficult to forecast accurately. The demand frequently goes back and forth from only the base load to the sum of the base load and the heating load. On the other hand, the weather stations used in the MWS model are not assumed to be “identical.” In other words, the MWS model works better when those weather stations are relatively far away from each other. Thus, this method is designed primarily for large operating areas.

4.4 Steps of estimation

In the following section, we assess the performance of the existing method and the MWS model following the steps shown in Figure 4.3.

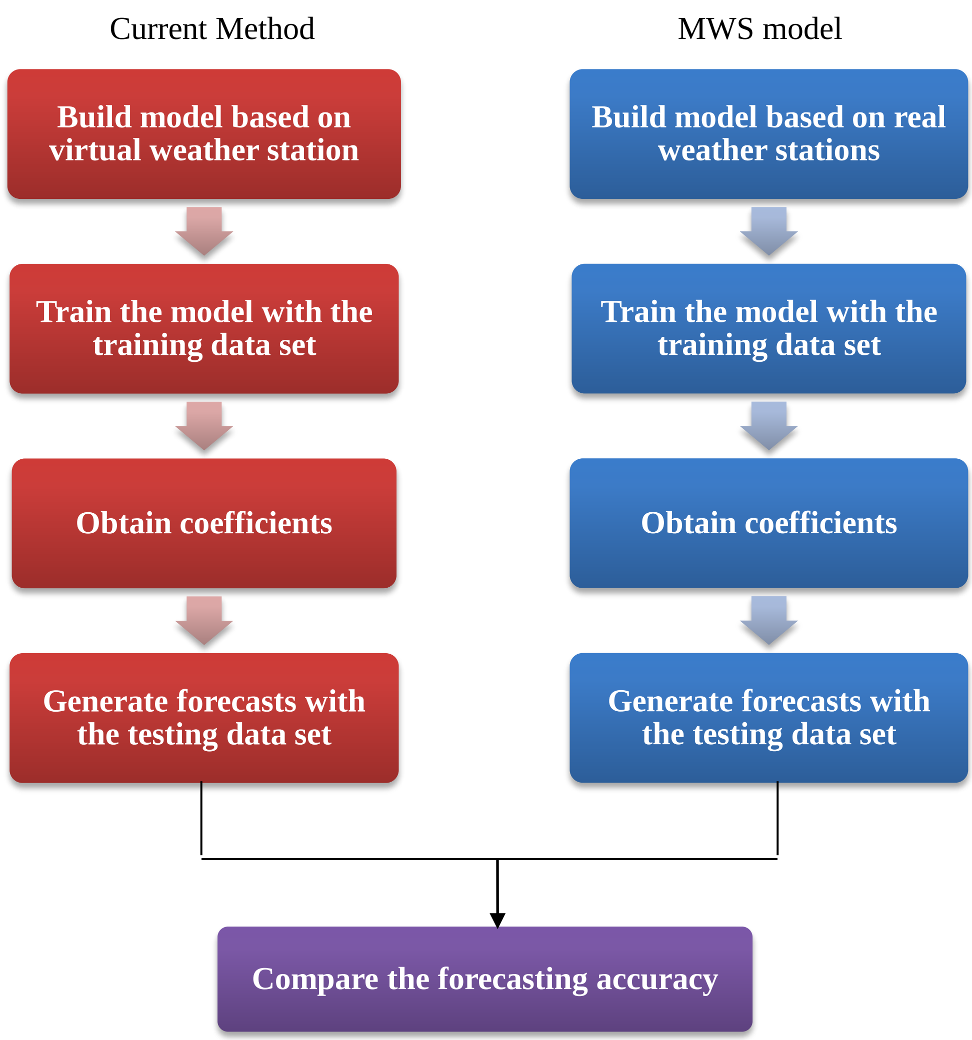

Figure 4.3 Steps of assessing the MWS model

As Figure 4.3 shows, we estimate the forecasting accuracy of the current method (Equation 4.4) and the MWS model (Equation 4.5) based on the testing strategy shown in Figure 4.2. In the following section, we are going to evaluate the forecasting results in terms of RMSE, MAPE, and weighted MAPE.

4.5 Error estimates: MWS model improves forecasting

The following figures compare the forecasting results of the current method and the MWS model. The upper subplots of Figures 4.4, 4.6, and 4.8 show the forecast values compared to the actual gas consumption in time series for the testing data sets from 2009 to 2011. The lower subplots are the forecasting errors for each forecasting period. Figures 4.5, 4.7, and 4.9 show the errors of the two methods for the same data sets.

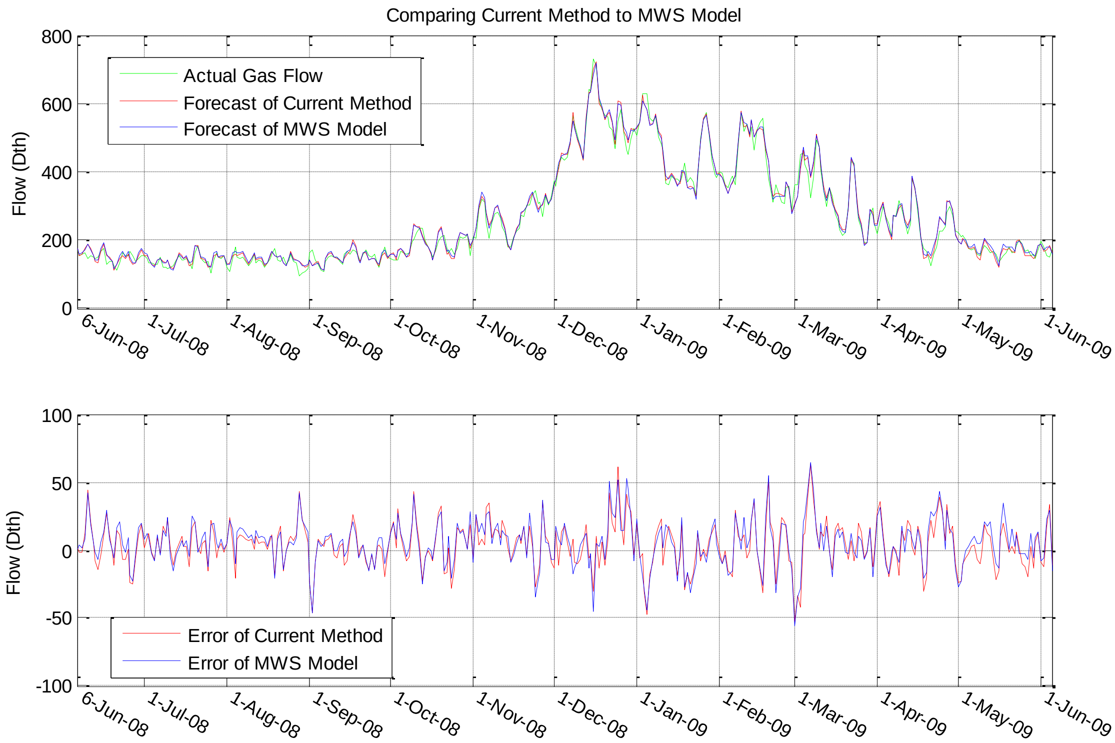

Figure 4.4 Gas demand forecasting for 2009

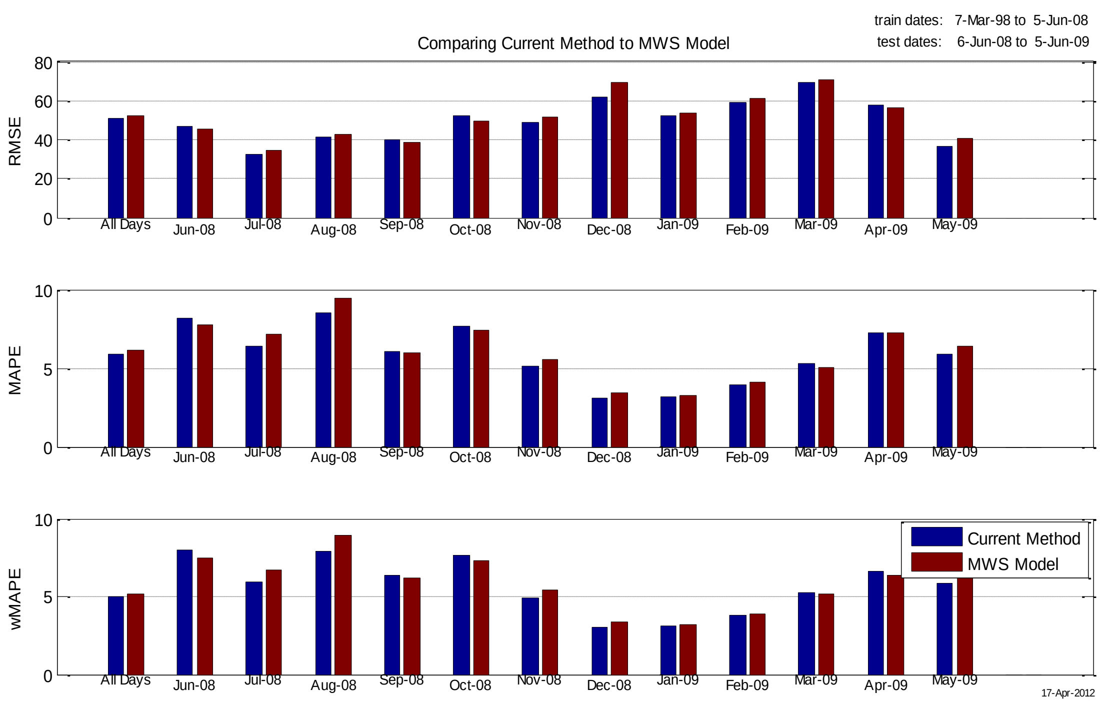

Figure 4.5 Error estimates for 2009

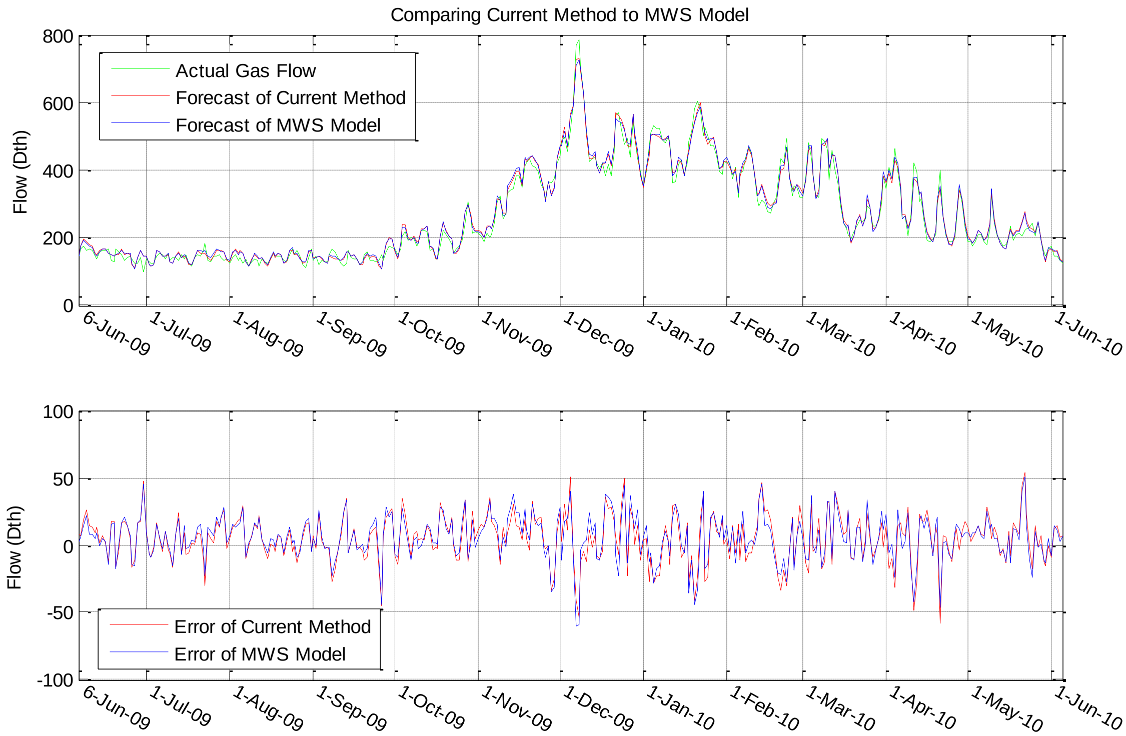

Figure 4.6 Gas demand forecasting for 2010

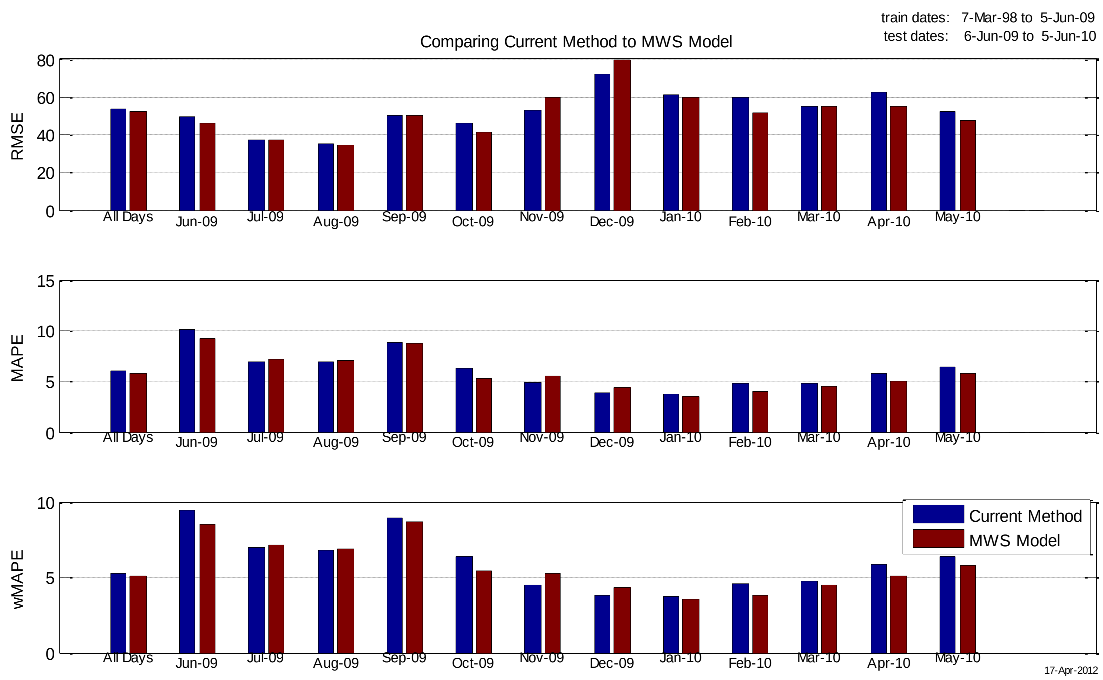

Figure 4.7 Error estimates for 2010

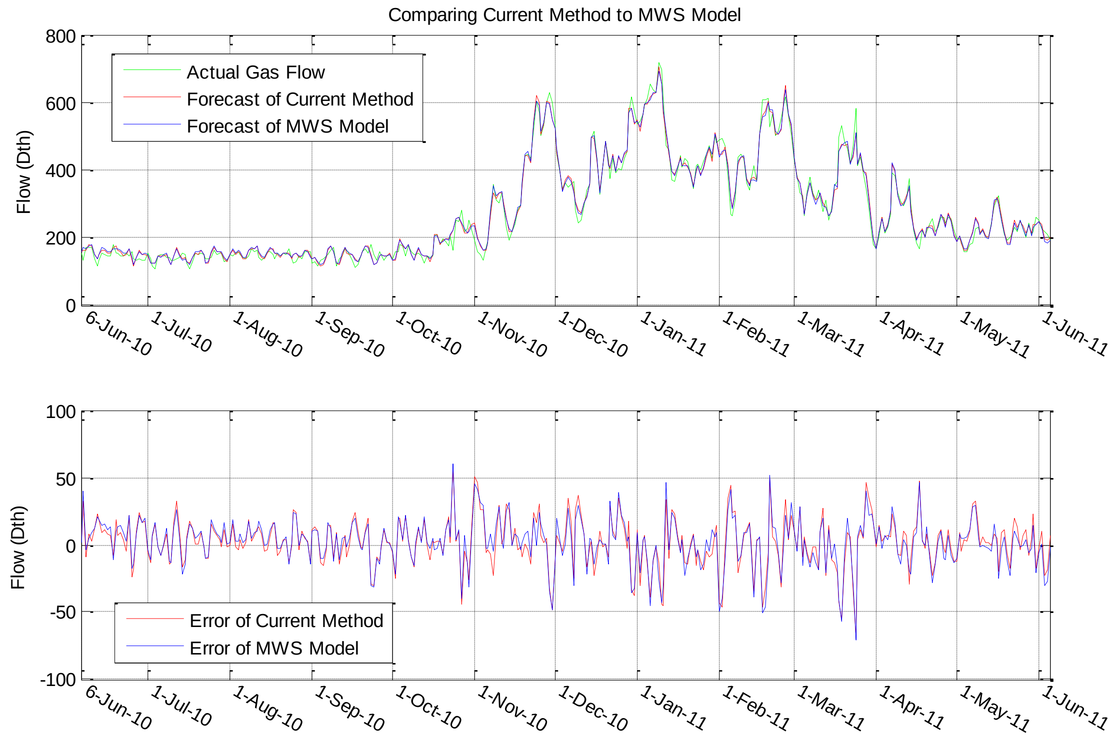

Figure 4.8 Gas demand forecasting for 2011

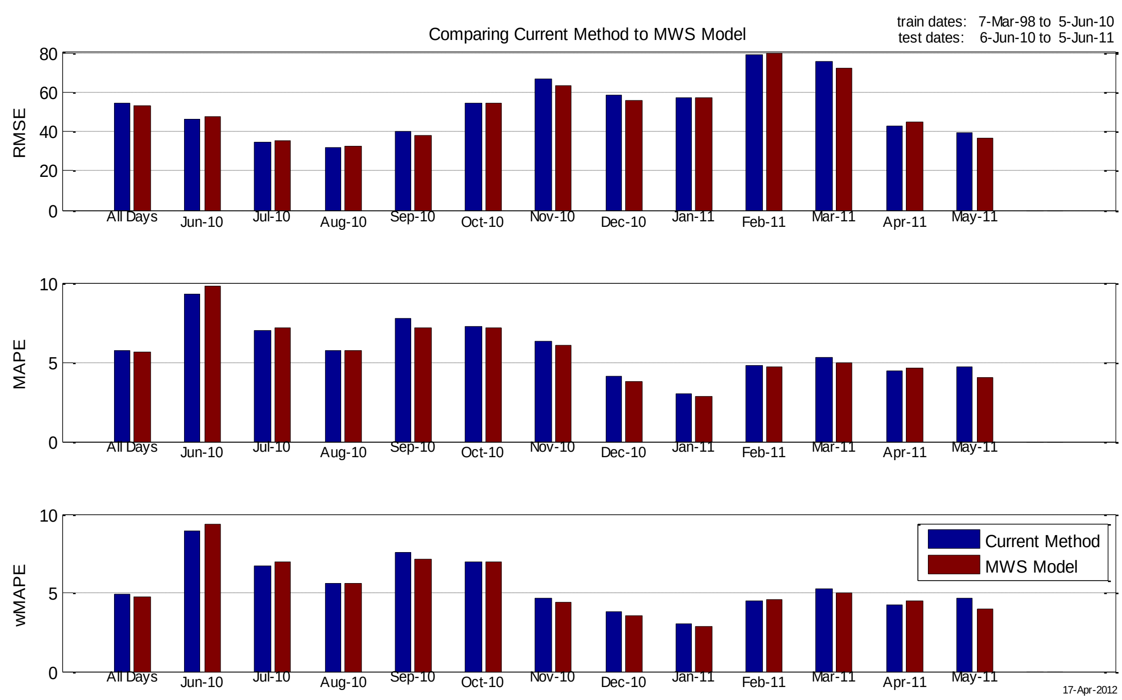

Figure 4.9 Error estimates for 2011

As we expected, the MWS model has improvement in terms of RMSE, MAPE, and wMAPE in the shoulder months for all the testing data sets. Overall, the new model improved the RMSE by about 5% and improved the MAPE by about 4%, comparing to the existing method. For all the testing winter months (November, December, January, and February), the forecasting accuracy are very similar. Both the current method and MWS model can predict demand with MAPE or weighted MAPE as low as 4%. However, in the shoulder months and summers, the MWS model is superior to the existing method on average by about 7% in terms of RMSE. The MWS model not only can improve the forecasting accuracy over the shoulder months; it can also help generate a better forecast during the summers where the CDD65 plays an important role. Further discussions of the results and the conclusions are given in Chapter 5.

CHAPTER 5

More Weather Inputs Improve Gas Demand Forecasts

In this thesis, our goal is to develop algorithms that extend the existing weather inputs of our current model. In Chapter 3, by applying statistical hypothesis testing, a forecasting performance test, and unusual day evaluation, we identified a series of new weather instruments that offer improvement to forecasts of natural gas consumption. In Chapter 4, by developing a new algorithm for an operating area with multiple weather stations, we significantly improved the forecast accuracy of all the testing period. In this chapter, we summarize the results from the previous two chapters. Some suggestions to the GasDay Lab and further possibilities to continue the research are presented as well.

5.1 New weather inputs help forecast

In Chapter 3, we discussed the significance of the new weather instruments that are conjectured to have impact on the gas consumption. Operating areas M and D are the two areas on which we tested the significance of each weather input. In the following sections, we provide a summary of the results for both operating areas.

5.1.1 Operating area M

For the operating area M, hypothesis testing identified significant variables from the new set of weather instruments. Dew Point and Cloud Cover are found to have a significant impact on the base load. Precipitation, on the other hand, is found to be statistically insignificant based on the T-test. Precipitation, Cloud Cover, and Dew Point all crossed with HDD are found to have significant contributions to the heating load. For the Wind Direction variable, all Wind Direction variables are found to have significant positive impact on the gas consumption statistically, which suggests that the cooling effect of wind is very important in natural gas demand forecast for the customers who use natural gas for heating. However, according to the forecasting performance test and the unusual days evaluation, Wind Direction is not helpful.

T-statistic gives an overview of the weather instrument candidates. Based on the forecasting performance, we found that Cloud Cover, Precipitation, and Dew Point have contributions to the existing model both individual and jointly. We have also found that these variables contribute more significantly during the shoulder months. On the other hand, we are reluctant to suggest adding Wind Direction to the current GasDay model. Even though it does slightly lower the RMSE and MAPE during the heating months, overall, the existing GasDay model has higher accuracy. Implementing unusual day evaluation allows us to evaluate the value of new input candidates from another prospective. We achieve improved accuracy by adding Cloud Cover, Dew Point, and Precipitation for most of the unusual day types. However, the forecasting accuracy is hurt by adding Wind Direction on most types of unusual days.

5.1.2 Operating area D

In operating area D, the gas use of the customers is different from that in operating area M. Therefore, different results are expected. For this area, most of the weather-related variables are found to be statistical insignificant for estimating gas consumption. However, forecasting performance has no significant change by adding new weather inputs. We still are not able to capture all the characteristics of this operating area. Our model has about 10% of wMAPE on the average for this operating area, compared with about 5% for operating area M. The most significant variable(s) are from the lagged gas flow, especially two days ago flow (S_{k-2}). The impact of new weather inputs are very limited in terms of forecasting at both time series and unusual days. It is reasonable to conclude that operating area D's gas consumption is not HDD, Precipitation, or even weather related. One potential reason is that the environment of this operating area is similar to a "desert oasis" [43]. Mild winters and abundance of sunny days are the main features of such an area. Within the data from NOAA (2005 to 2009), 7 % of the Precipitation data is greater than 0.1 inch/day, only 10% of the Precipitation data is greater than 0.05 inch/day; and about 58% of the Precipitation data is zero inch/day. Because of the insufficient Precipitation data, to model and forecast the gas consumption of operating area D based on Precipitation is difficult.

Beside the weather effects, other factor(s) might relate to the consumption of natural gas of this area. Operating area D is a relatively small territory with population less than 1000 [2]. At such a low density of population, the growing schedules of individual farmers and contracts to purchase their farm produce may affect their gas consumption significantly. To further improve the performance of our models, we should include variable(s) that can represent characteristics such as farmers' irrigation schedules and their harvest patterns. For the research purposes of this thesis, we do not investigate further on the problem of forecasting for operating area D.

5.2 Multiple weather stations help forecast

In this section, we provide brief conclusions from Chapter 4. Overall, the MWS model better represented actual weather conditions for operating area C than the current method. We improved our forecasts over the heating days without loss of forecasting accuracy on non-heating days, especially for the shoulder months when the temperature varies more frequently. The performance of MWS model occasionally is not as good as the current method, but overall it is competitive. Hence, we can improve our forecast accuracy by applying the MWS model during heating season and the shoulder months, with slightly deterioration of accuracy of summers. Hence, we suggest using MWS model forecasting the shoulder months and summer, use the existing GasDay model forecasting the winter. Overall, our goal is to improve forecast accuracy by extend our input space. The current GasDay model accomplishes this goal by supporting four temperature readings in one day, we call the "temperature quarters." In this paper, we give the evidence to show forecasting improvement by extending our inputs by the numbers of weather stations.

5.3 Result summary and suggestions for GasDay

With a series of exogenous weather inputs, the accuracy of forecasts experiences a significant improvement. The new weather instruments we discussed in Chapter 3 could help improve the forecast by capturing more characteristics for the model for the customer with both heating and non-heating purposes. The multiple weather stations, on the other hand, better reproduced the actual temperature conditions on the customer side. Based on this study, we offer a few suggestions for GasDay.

1. For temperature sensitive areas such as operating area M, the additional weather inputs help to improve forecasting accuracy. Based on the T-statistics, Cloud Cover, Precipitation, and Dew Point are found to impact on gas demand significantly. By including the Precipitation, the forecast can be improved by 2.5% overall. Cloud Cover, as an indicator of radiation effect, helps to improve the RMSE by 1.2%. Dew Point, as a humidity indicator, offers 2.4% improvement of RMSE. The three variables jointly provide 5% improvement of RMSE comparing to the existing GasDay model. We recommend adding Cloud Cover, Precipitation, and Dew Point to the existing GasDay model.

2. For Wind Direction, all four Wind Direction factors are found to be statistically significant, which reinforces the significant cooling effect by the wind, especially in the heating season. Among these variables, the T-statistics are not significantly different from each other, which indicate that the wind effect does impact gas consumption with about equally significance for all four Wind Directions. However, our method of modeling the wind effect is not as good as the existing model. We

suggest keeping the wind-adjusted HDDs as a critical input in the GasDay model.

3. For non-temperature-sensitive operating areas such as area D, none of the new weather instruments are found to be highly correlated with gas consumption. Cloud Cover, Precipitation, and Dew Point have no apparent improvement of accuracy according to the forecasting on both usual days and the unusual days. More information and data are needed if we intend to improve our forecast for such an operating area.

4. In Chapter 4, we discussed of including multiple weather stations into the model. For GasDay, it is recommended to use real weather stations' data as long as it is available instead of combining them into a single virtual weather station. However, we note that the MWS model may lead to an extra operating cost. The development of MWS model should be under a pre-determined development cost (time consumed) to ensure that current work will not be affected. Overall, from a long-term view of the development of GasDay, including multiple weather stations in the model is recommended.

5.4 Further thinking

Although this paper has investigated a feasible method to test the impact of new weather inputs on natural gas consumption, there are still many improvements which can be made to our methods. We list several possible improvements and extensions to improve the work.

1. We identified several weather instruments that are statistically significant. However, for some variables, the T-statistics shown that the coefficients are different from zero does not imply a strong impact on gas consumption. Further statistical tests might be applied to the data sets.

2. For those variables that have been tested in this thesis, we recommend to experiment using the same techniques with multiple LDCs or operating areas. The results may be sensitive to the operating area that we examined.

3. For operating area D, it is possible to improve the accuracy of the forecast by including more inputs related to the individual farmers' behavior, which might provide much more information to the model.

4. For the new weather inputs discussed in the prior chapters, one can apply other techniques to re-evaluate the significance of those inputs both individually and jointly.

5. In this paper, we have tested Cloud Cover, Dew Point, Precipitation, and Wind Direction as new weather inputs. It may be possible to test the significance of other external inputs based on our existing GasDay model.

6. Techniques developed in this paper might also be used to forecast demand of other energies.

BIBLIOGRAPHY

[1] C. Aksu and S. I. Gunter, "An empirical analysis of the accuracy of SA, OLS, ERLS, and NRLS combination forecasts," *International Journal of Forecasting*, vol. 8, pp. 27-43, 1992.

[2] American Fact Finder, 2010, http://factfinder.census.gov/home/saff/main.html?_lang=en, (Accessed 1 Feb. 2011).

[3] American Gas Association (AGA), 2010 http://www.aga.org/Pages/Default.aspx, (Accessed 1 Sep. 2010).

[4] J. S. Armstrong (ed.), *Principles of Forecasting: A Handbook for Researchers and Practitioners*, Norwell, MA: Kluwer Academic Publishers, 2001.

[5] J. G. Asbury, C. Maslowski, and R. O. Mueller, "Solar availability for winter space heating: An analysis of SOLMET data," *Science*, vol. 206(4419), pp. 679-681, 1979.

[6] D. G. Baker, "Effect of observation time on mean temperature estimation," *Journal of Applied Meteorology*, vol. 14(4), pp. 471-476, 1975.

[7] L. Biacino and G. Gerla, "Fuzzy logic, continuity and effectiveness," *Archive for Mathematical Logic*, pp. 643–667, 2002.

[8] P. Bretschneider, T. Rauschenbach, and J. Wernstedt, "Hybrid forecast strategy using an adaptive fuzzy classification algorithm," *EUFIT '98; 6th European Congress on Intelligent Techniques and Soft Computing*, vol. 3, pp. 1916-1919, 1998.

[9] M. Brian, "Forecasting daily processing tomato harvest tonnage in california," Marquette University, Department of Electrical and Computer Engineering, Milwaukee, WI, M.S. thesis, 2004.

[10] R. H. Brown, "OK, who is playing with the thermostat?" *Gas Forecaster's Forum.* Albuquerque, NM, 2003.

[11] R. H. Brown, P. Kharouf, I. Matin, and L. P. Piessens, "Development of Artificial Neural Network Models to Predict Daily Gas Consumption," American Gas Association Forecasting Review, vol. 5, pp. 1-22, March 1996.

[12] R. D. Cess, M.H. Zhang, P. Minnis, L. Corsetti, et al., “Absorption of solar radiation by clouds: Observation versus models,” *Science*, vol. 267, pp. 496-499, 1995.

[13] C. K. Chan, B. G. Kingsman, and H. Wong, "A Comparison of unconstrained and constrained OLS for the combination of demand forecasts: A case study of the ordering and stocking of bank printed firms," *Annals of Operations Research*, vol. 87, pp. 129-140, 1999.

[14] H. Chen, Y. Du, and J.N. Jiang, “Weather sensitive short-term load forecasting using knowledge-based ARX models,” *IEEE Transactions on Neural Networks*, Austin, TX, vol. 1, pp. 190-196, 2005.

[15] D. Cramer and H. Dennis, *The Sage Dictionary of Statistics*, SAGE publications Ltd, 2004.

[16] P. M. Dare, “A study of the severity of the midwestern winters of 1977 and 1978 using heating degree days determined from both measured and wind chill temperatures,” *Bulletin of the American Meteorological Society*, vol. 62(7), pp. 974-982, 1981.

[17] Energy Dictionary, 2010 http://www.energyvortex.com/energydictionary/EnergyVortex.htm, (Accessed 3 Mar. 2011).

[18] M. K. Evans, *Practical Business Forecasting*: Malden, Mass, Blackwell Publishers, 2003.

[19] R. A. Fisher, “Statistical methods for research workers,” Edinburgh: Oliver and Boyd, pp.43, 1925.

[20] R. A. Fisher, *Statistical Methods and Scientific Inference*, New York, Hafner, 1956.

[21] D. N. Gujarati, *Basic Econometrics*, McGraw Hill, Boston, MA, 2003.

[22] S. Halim, “Selection of inputs for generating combinatorial daily natural gas demand forecasts,” Marquette University, Department of Electrical and Computer Engineering, Milwaukee, WI, M.S. thesis, 2004.

[23] J. Hanke, A. Reitsch, and D. Wichern, *Business Forecasting* (7th Ed.), Prentice Hall, 2001.

[24] J. Hawkins, “Economic forecasting: history and procedures,” *Economic Roundup, Autumn*, The Treasury, Canberra, 2005.

[25] L. Hilbelink, "Consolidating weather information in gas load estimation," Marquette University, Department of Electrical and Computer Engineering, Milwaukee, WI, M.S. thesis, 1999.

[26] O. Hyde and P. F. Hodnett, "Modelling the effect of weather in short-term electricity load forecasting," *Mathematical Engineering in Industry*, vol. 6, 1997.

[27] A. Khotanzad and H. Elragal, "Natural gas load forecasting with combination of adaptive neural networks," *Proceedings of the International Joint Conference on Neural Networks*, vol. 6, pp. 4069-4072, 1999.

[28] A. Khotanzad, H. Elragal, and T.-L. Lu, "Combination of artificial neural-network forecasters for prediction of natural gas consumption," *IEEE Transactions on Neural Network*, Dallas, TX, vol. 11, pp. 464-473, 2002.

[29] D. Kondepudi, *Introduction to Modern Thermodynamics*, Wiley, Chichester, pp. 108-114, 2008.

[30] M. G. Lawrence, "The relationship between relative humidity and the Dew Point temperature in moist air: A simple conversion and applications," Bulletin of the American Meteorological Society, pp. 225–233, 2005.

[31] E. Lehmann and J. Romano, *Testing Statistical Hypotheses* (3rd Ed). New York: Springer, 2005.

[32] H. L. Lim and R.H. Brown, "Hourly gas load forecasting model input factor identification using a genetic algorithm." *44th IEEE Midwest Symposium on Circuits and Systems*, Dayton, OH, pp. 670–673, 2001.

[33] National Geographic (2011) http://photography.nationalgeographic.com/photography/enlarge/irrigated-fields-kendric_pod_image.html, (Accessed 2 Dec. 2011).

[34] National Oceanic and Atmospheric Administration (NOAA), 2010 http://www.noaa.gov/, (Accessed 5 Feb. 2010).

[35] Natural Gas.org, 2009 http://www.naturalgas.org/naturalgas/distribution.asp, (Accessed 2 Dec. 2009).

[36] W. Newey and K. West, "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, vol. 55(3), pp. 703-708, 1987.

[37] M. Pidwirny and S. Jones, *Fundamentals of Physical Geography*, 2010, http://www.physicalgeography.net/, (Accessed 11 Nov. 2010).

[38] J. V. Sutcliffe and Y. P. Parks, *The hydrology of the Nile*. IAHS Special Publication, pp. 161, http://iahs.info/bluebooks/SP005/BB_005_0161.pdf, 1999.

[39] L. Talavera, "An evaluation of filter and wrapper methods for feature selection in categorical clustering", *Advances in Intelligent Data Analysis VI*, pp. 440-451, 2005.

[40] K.E. Trenberth, D.P. Stepaniak, and J.M. Caron, "The global monsoon as seen through the divergent atmospheric circulation," *Journal of Climate*, Vol.13, pp. 3969-3993, 2000.

[41] U.S. Department of Energy (DOE), 2010 http://www.energy.gov/, (Accessed 7 Apr. 2010).

[42] U.S. Energy Information Administration (EIA), 2010 http://www.eia.doe.gov/, (Accessed 17 May. 2010).

[43] U.S. Geological Survey, 2010 http://www.usgs.gov/, (Accessed 30 Mar. 2010).

[44] U.S. Department of Agriculture (USDA), 2010 http://www.usda.gov/ , (Accessed 13 Dec. 2010).

[45] S.R. Vitullo, R.H. Brown, G.F. Corliss, and B. Marx, "Mathematical model for natural gas forecasting," vol. 17(4), *Canadian Applied Mathematics Quarterly*, 2009.

[46] S.R. Vitullo, "Disaggregating time series data for energy consumption by aggregate and individual customer," Marquette University, Department of Electrical and Computer Engineering, Milwaukee, WI, Ph.D. Dissertation, 2011.

[47] R.R. Yager and D. P. Filev, "Essentials of fuzzy modeling and control," *SIGART Bulletin*, New York, NY: John Wiley, vol. 6(4), pp.22-23, 1994.

[48] H. White, "A Heteroskedasticity Consistent Covariance Matrix and a Direct Test for Heteroskedasticity," *Econometrica*, vol.48, pp. 817-838, 1980.

[49] Wikipedia, Dew Point, 2009. http://en.wikipedia.org/wiki/Dew_point, (Accessed 28 Mar. 2011).

[50] Wikipedia, Student's t-test, 2012. http://en.wikipedia.org/wiki/Student%27s_ttest#Equal_sample_sizes.2C_equal_variance, (Accessed 11 Jul. 2012).

[51] Wikipedia, F-test 2012. http://en.wikipedia.org/wiki/F-test#Regression_problems, (Accessed 11 Jul. 2012).

[52] P. Zuidema and C. Fairall, “On air–sea interaction at the mouth of the Gulf of California", *Journal of Climate*, published by the American Meteorological Society, vol. 20(9), May 2007.