Design of a Weather-Normalization Forecasting Model

Progress Report

Abram Gross
Yafeng Peng
Jedidiah Shirey

3/4/2014

TABLE OF CONTENTS

1.0 INTRODUCTION

Northern Virginia Electric Cooperative (NOVEC) is an electricity distributor headquartered in Manassas, Virginia. In order to meet energy demand, NOVEC purchases power in two ways: long-term bulk purchases and spot purchases. Bulk purchases occur up to five years in advance and are meant to satisfy estimated demand over this time period. In the event bulk purchases are insufficient to meet any demand over this timeframe, spot purchases provide the energy to cover the difference. Temperature fluctuations, mainly during the summer months, are a significant contributor to increased power demand in excess of the bulk purchase amount.

In order to minimize the amount of ad hoc purchases without overcompensating for their avoidance with excessive bulk purchases, NOVEC has developed a forecasting model that estimates future energy purchases over a 30-year horizon. Rather than predicting the temperature patterns for the future, the model uses the long-run average value as determined from historical temperature data. Historical power consumption since 1983, weather data since 1963, and economic forecasts at the state and county level are all inputs to the model. The output of the model is a monthly power demand forecast over a 30-year horizon.

2.0 PROBLEM STATEMENT

In order to predict future power demand, the model performs weather normalization for 50 years of hourly weather data and evaluates economic data provided by Moody's economic forecast. In accounting for these variables, NOVEC believes that the current model may no longer be the best available and that a new weather-normalization method may better reflect recent changes in weather trends. Improving the accuracy of the forecast would limit the amount of power that NOVEC has to buy beyond the bulk amount, thus decreasing costs. NOVEC requests analytical support to develop a new weather-normalization forecasting model or to determine that the existing model is the best available.

3.0 SCOPE

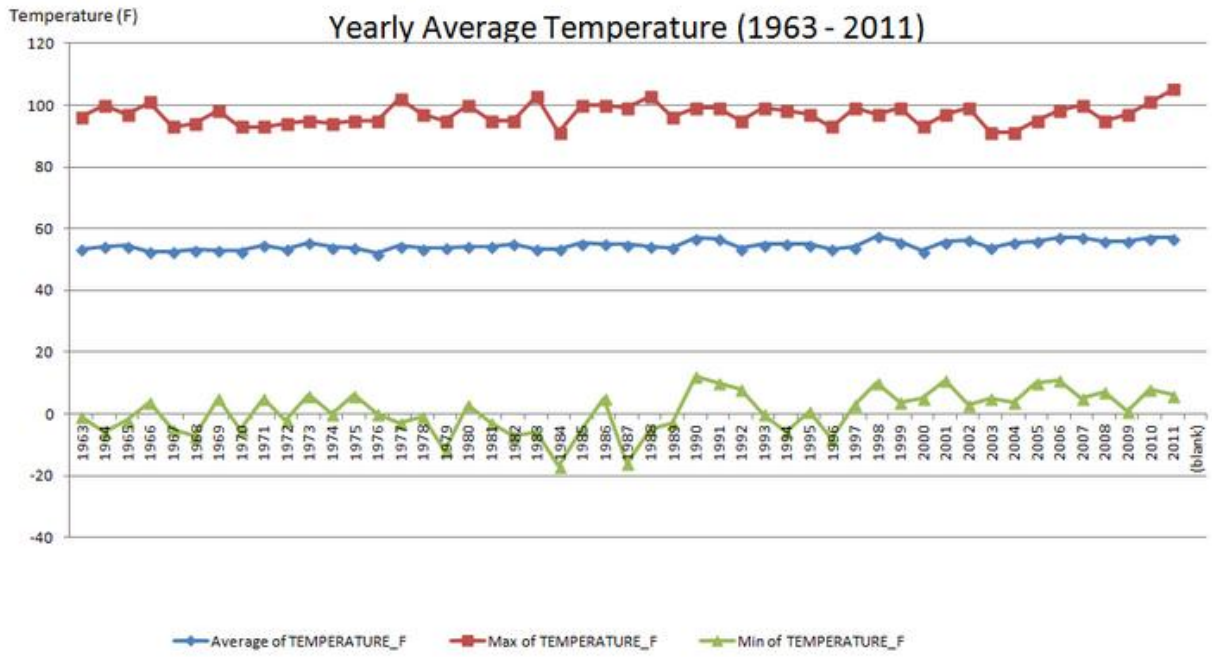
The purpose of this project is to develop a new weather normalization methodology to improve NOVEC's forecasting model by more accurately modeling future power demand. The model will take into account historical data as inputs: customer and energy purchase totals by

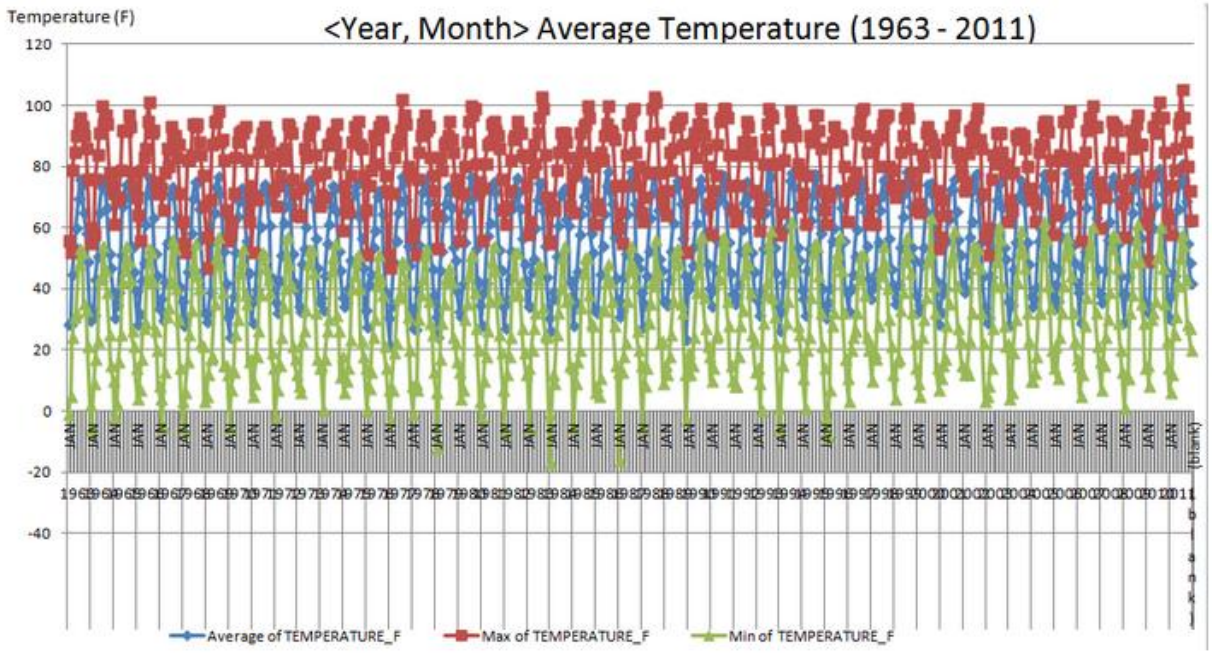
month starting from 1983 and hourly weather data starting from 1963. Additionally, this analysis will leverage Moody's state, county, and Washington, D.C. metro economic data starting from the 1970s. In particular, per sponsor guidance, data relating to employment, housing stocks, and GDP will be used to predict the growth or decline of NOVEC's customer base, though other metrics are available for analysis. Moody's economic data includes projections of econometrics across varied scenarios, only one of which is currently used to inform NOVEC's forecasts. Testing the model under additional scenarios offers a means to conduct sensitivity analysis and inform the sponsor's decisions with some measure of risk related to modeling assumptions. This project will not focus on understanding or modifying the existing model as the sponsor is interested solely in a completely new methodology.

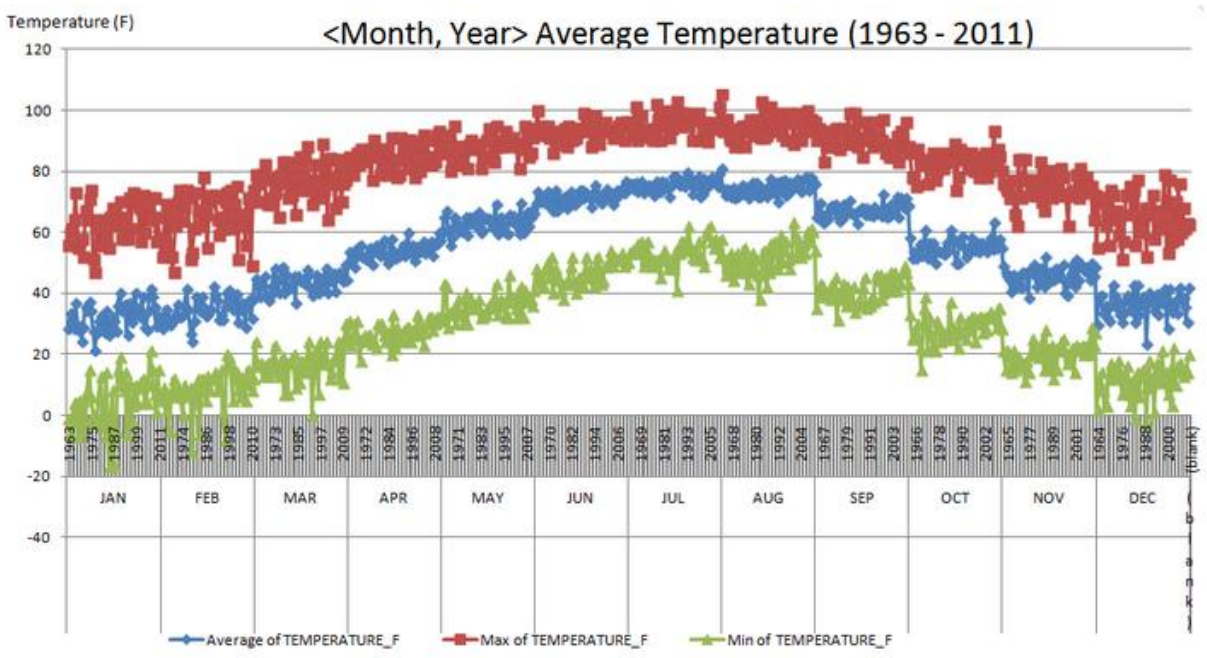
4.0 PROGRESS

The first step in developing the model is data exploration. After completing the data exploration, we will move on to developing the weather-normalization model in two stages. The first stage will be the specific weather-normalization procedure and the second stage will be the economic factors. These two stages will form a complete model that outputs a 30-year forecast at a monthly level.

Currently, we are spending most of our time on the data exploration and the weather-normalization stage. Using Excel, we have been able to manipulate the data into more useful structures that help with data visualization. As recommended at our meeting last week, we have created three charts (average yearly temperature against year from 1963 to 2011; average monthly temperature against month from 1963 to 2011; and min, max, and average monthly temperature against month) that help to visualize the temperature trends over time.


Temperature (F)
Yearly Average Temperature (1963 - 2011)
120
100
80
60
40
20
0
-20
-40
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
(blank)
Average of TEMPERATURE_F
Max of TEMPERATURE_F
Min of TEMPERATURE_F


Temperature (F)
<Year, Month> Average Temperature (1963 - 2011)
120
100
80
60
40
20
0
-20
-40
JAN
(blank)
Average of TEMPERATURE_F
Max of TEMPERATURE_F
Min of TEMPERATURE_F



The weather-normalization component of the model will output two variables: the number of heating degree days (HDD) and cooling degree days (CDD) for each month of the year. The equations for HDD and CDD are as follows:

$$HDD = \sum_{i=1}^{N} (T_b - \bar{T}_i)^+$$

$$CDD = \sum_{i=1}^{N} (\bar{T}_i - T_b)^+$$

N is the number of days per month, T_b is the base temperature coinciding with a non-heating/cooling temperature, and $\bar{T}_i$ is the average temperature per day. Since our temperature data as available at an hourly resolution rather than a daily, our expectation is that we will be able to determine a more accurate HDD and CDD amount. We have a working model that calculates the heating/cooling degree *hours* in a day and sums the number of hours for each month and divides by 24 to determine the number of HDD and CDD in a month. The other component of the weather-normalization component is the change in temperatures during the period of historical data. Our goal is to reflect the changing temperature over time in our forecast of HDD and CDD in the future. To do this, we are considering utilizing a regression line that will adjust the HDD and CDD days in the forecast.

As of now, we have not yet begun working in earnest on the economic variables component of the model. After consulting with our sponsor, we have determined the variables from the economic data that NOVEC currently uses in their weather-normalization model. We plan on starting with these variables and, as time allows, consider methods of integrating other economic factors into the model.

5.0 Problems and Risks

Currently, we have met with no significant problems and foresee limited risks in the future. We have met with our sponsor multiple times and are pleased with his availability. His input and guidance has helped us to narrow the problem and make corrections to our work. Also, we received all of the data that we need for our model from him at our first meeting and have been able to spend significant amounts of time working on it.

6.0 Way forward

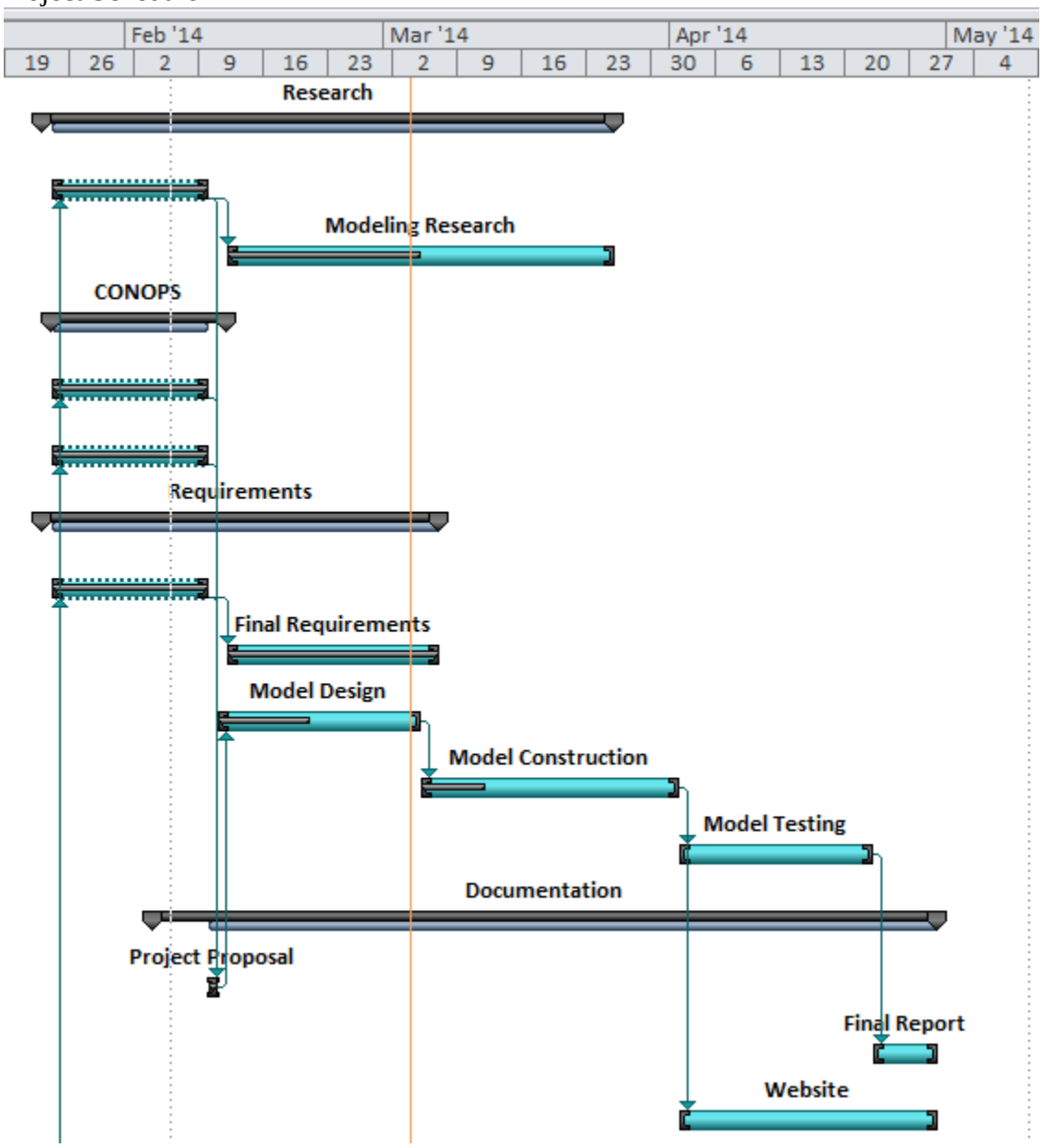
Our next steps are as follows:

1) Finalize HDD and CDD calculations method and convert to monthly value.
2) Finalize methodology to account for changing temperatures over time.
3) Integrate Steps 1 and 2 so that the HDD and CDD per month change throughout the duration of the forecast.
4) Determine economic factors to utilize and implement them in the model.
5) Research website development and begin implementation (concurrent with previous steps).

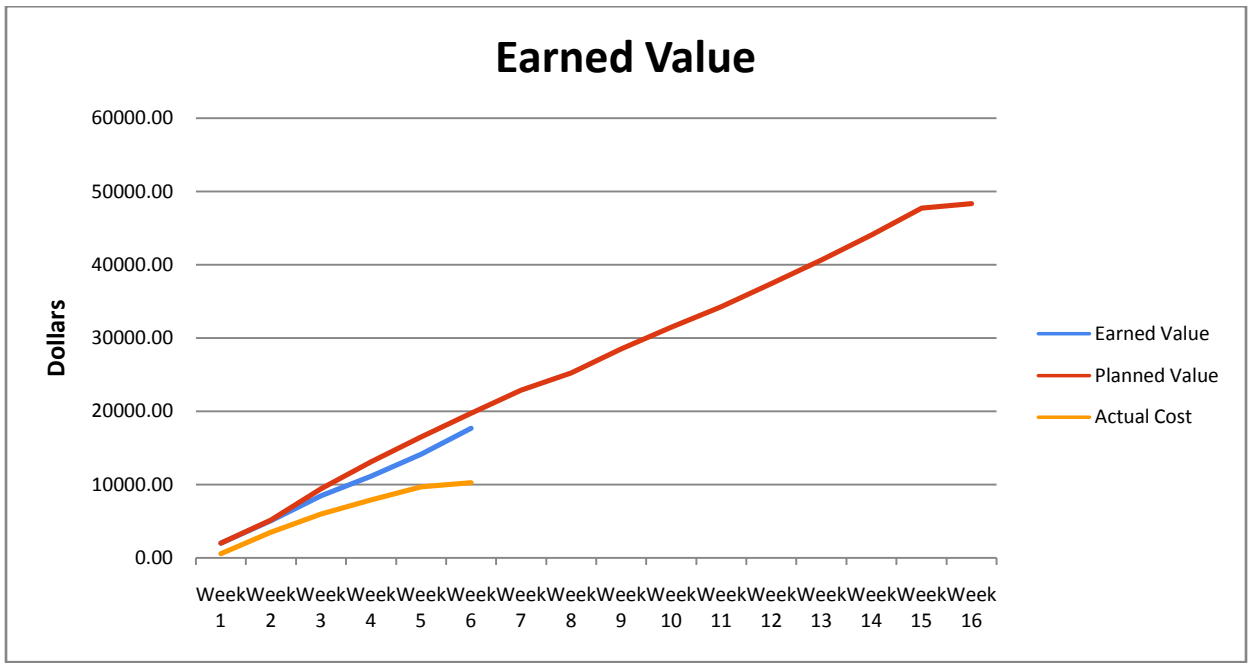
Steps 1, 2, and 3 we plan on completing in the next 2 weeks. Step 4 should be completed in the 1-2 weeks beyond that.

APPENDIX

Project Schedule



Earned Value



CPI and SPI