Time Series Analysis

Energy Consumption Data

Filtering, Forecasting and Visualization

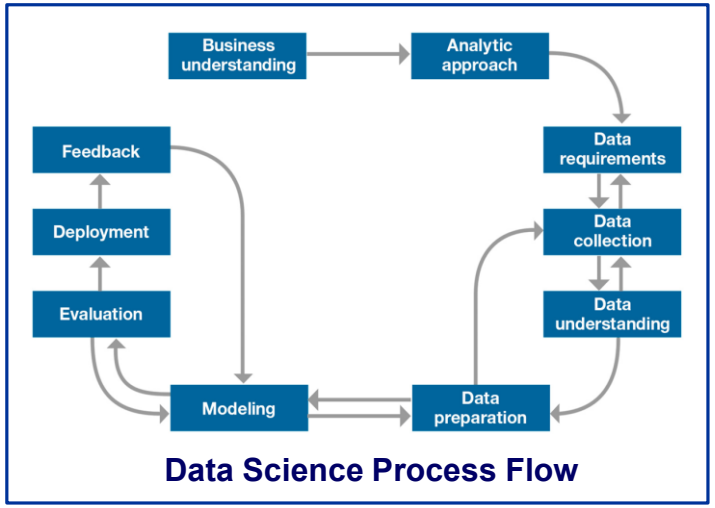
Greg Hepguler

BUSINESS OBJECTIVES

- Gather information and data on the amount of electric power consumption in different areas of a smart home, where electrical system is monitored by multiple submeters
- Investigate electric power consumption from the different submeters, sorted out by 3 specific areas in the house
- Model patterns of energy usage in different time frames such as by time of day or day of the year in a typical smart home
- Discover the specific rooms or appliances that are using the most electric power
- Observe changes, trends, patterns in power consumption

BUSINESS AND IoT ANALYTICS OBJECTIVES

- Apply IoT Analytics to evaluate electric submeter data acquired in a smart home
- Follow the data science process flow
- Create visualizations of power consumption data acquired every minute by the 3 electric submeters in this household



Data Science Process Flow

- Create these visualizations by filtering the electric submeter (ESM) data into to smaller datasets
- Investigate larger time frames than one minute
- Demonstrate how we can forecast future consumption
- Observe and present the components of ESM data
- Investigate different forecasting options

Review of ESM data

- Retrieved 2,049,279 measurements gathered between December 2006 and November 2010 (47 months) in one minute intervals
- Documented and described the contents of the data
- Sorted out specific types of data in each measurement acquired by the minute
-
- Checked if data contains any error
- Checked if there is any missing value
- Checked data statistics to see upper and lower end of data to see if there are outliers

Description of the Electric Submeter Data

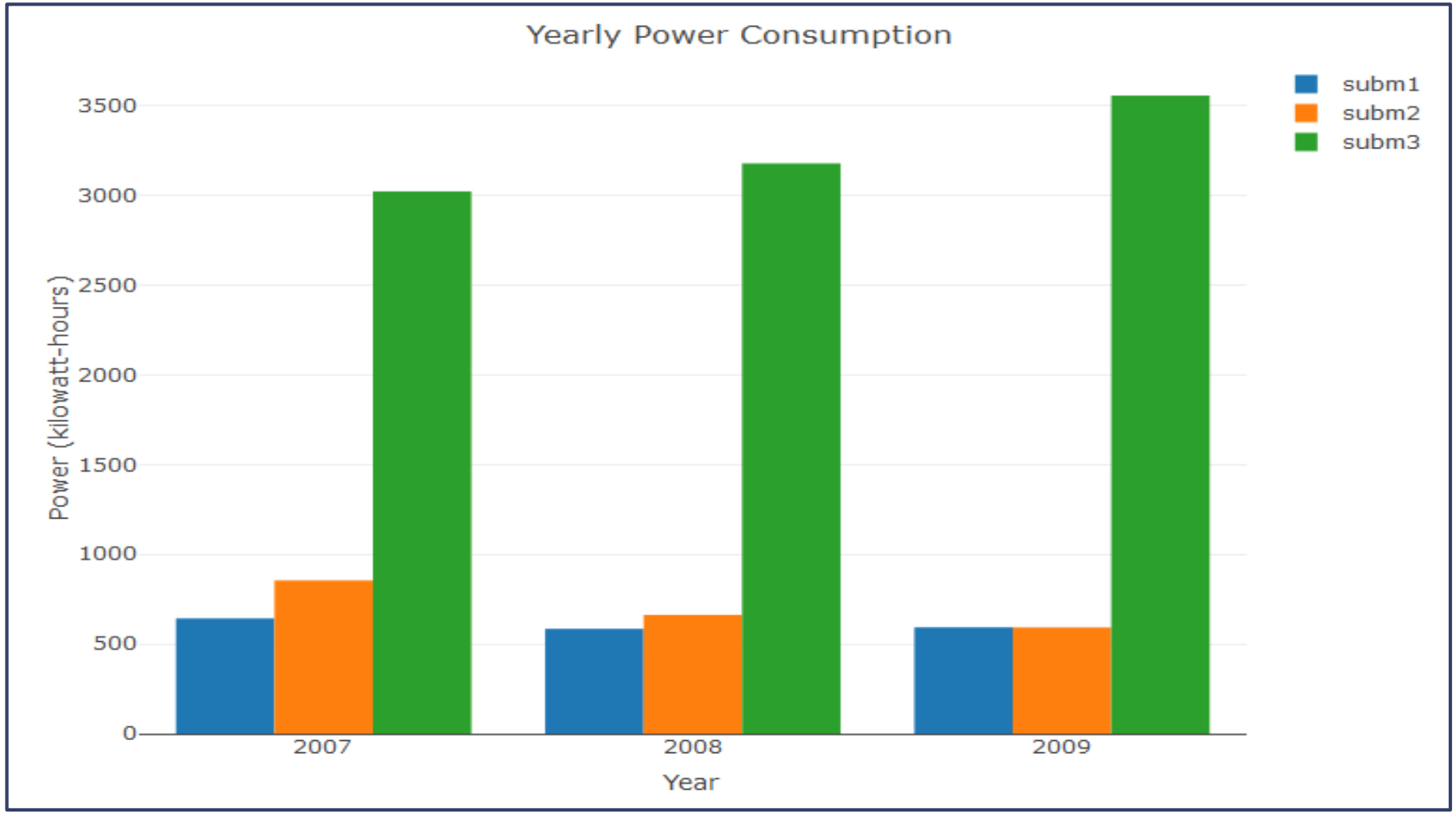
We decided to focus on 5 specific features from the electric submeter dataset, namely

- Date provided as dd/mm/yyyy
- Time provided as hh:mm:ss
- Power consumption from energy submeter No. 1, which corresponds to the kitchen, containing mainly a dishwasher, an oven and a microwave
- Power consumption from energy submeter No. 2, which corresponds to the laundry room, containing a washing-machine, a tumble-drier, a refrigerator and a light
- Power consumption from energy submeter No. 3, which corresponds to an electric water-heater and an air-conditioner

Visualizations in 4 time frames (Filtering)

- Created visualizations by filtering the electric submeter (ESM) data into to smaller datasets to investigate larger time frames than one minute records:
  1. Total annual consumption (kWh) for each submeter over the Jan-07 thru Dec-09 period.
  2. The average daily consumption (kWh) for each ESM by weekday for the winter seasons (December through February) over Jan 2007 thru Oct 2010 period. The results would present 7 values, one for each weekday and reflect the typical usage per weekday during the winter season.
  3. The average hourly kWh used for each hour of the day during January 2010. The result should only have 24 values and reflects the typical usage per hour of day during Jan-10.
  4. Power consumption at 12:00pm (noon) for the 1st day of each month in 2009.

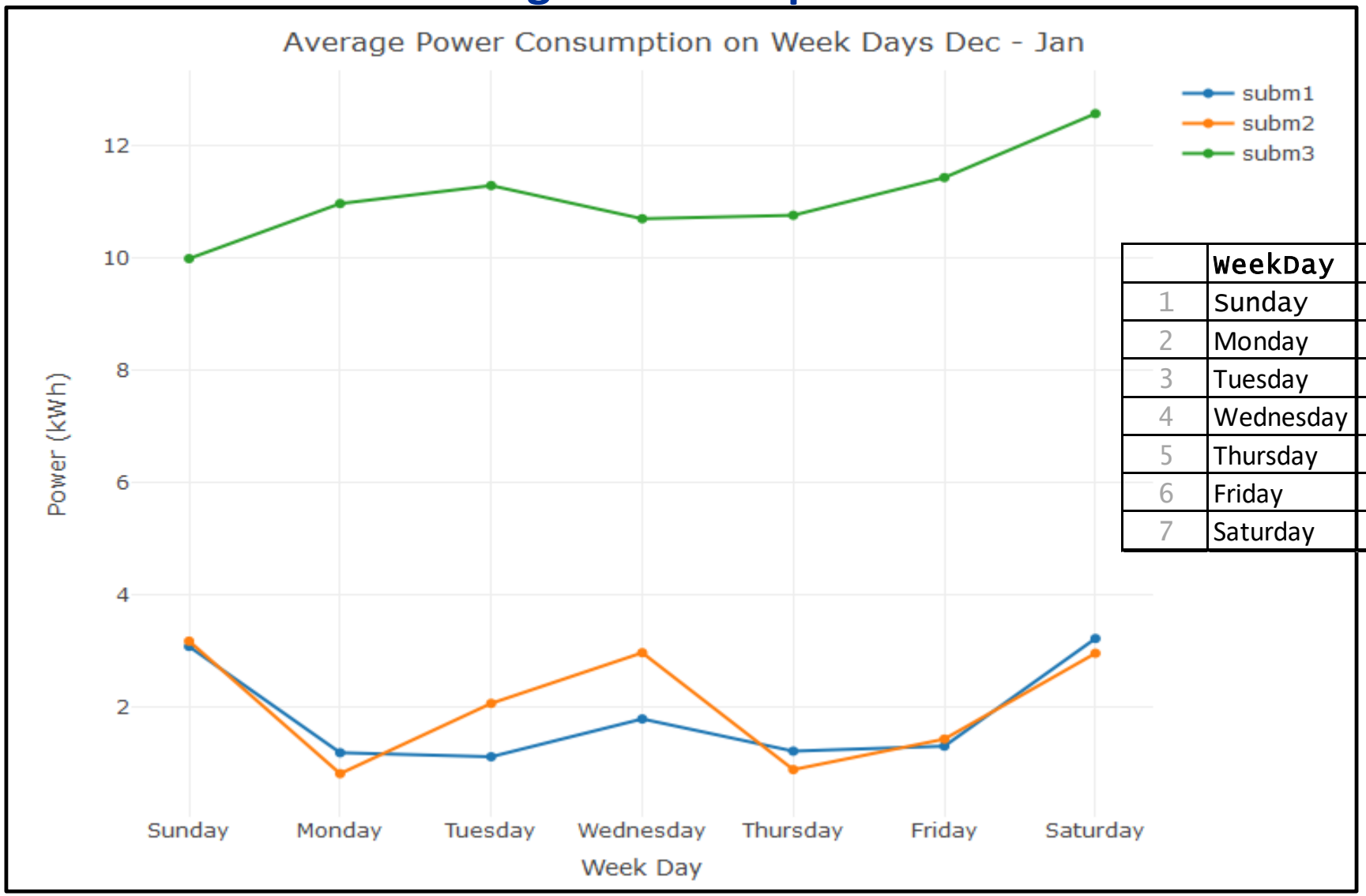
1. Total annual consumption (kWh) for each submeter over the Jan-07 thru Dec-09 period



	Year	SM1	SM2	SM3	DateTime
1	2007	643	854	3023	2007-01-01 00:00:00
2	2008	584	662	3179	2008-01-01 00:00:00
3	2009	593	592	3557	2009-01-01 00:00:00

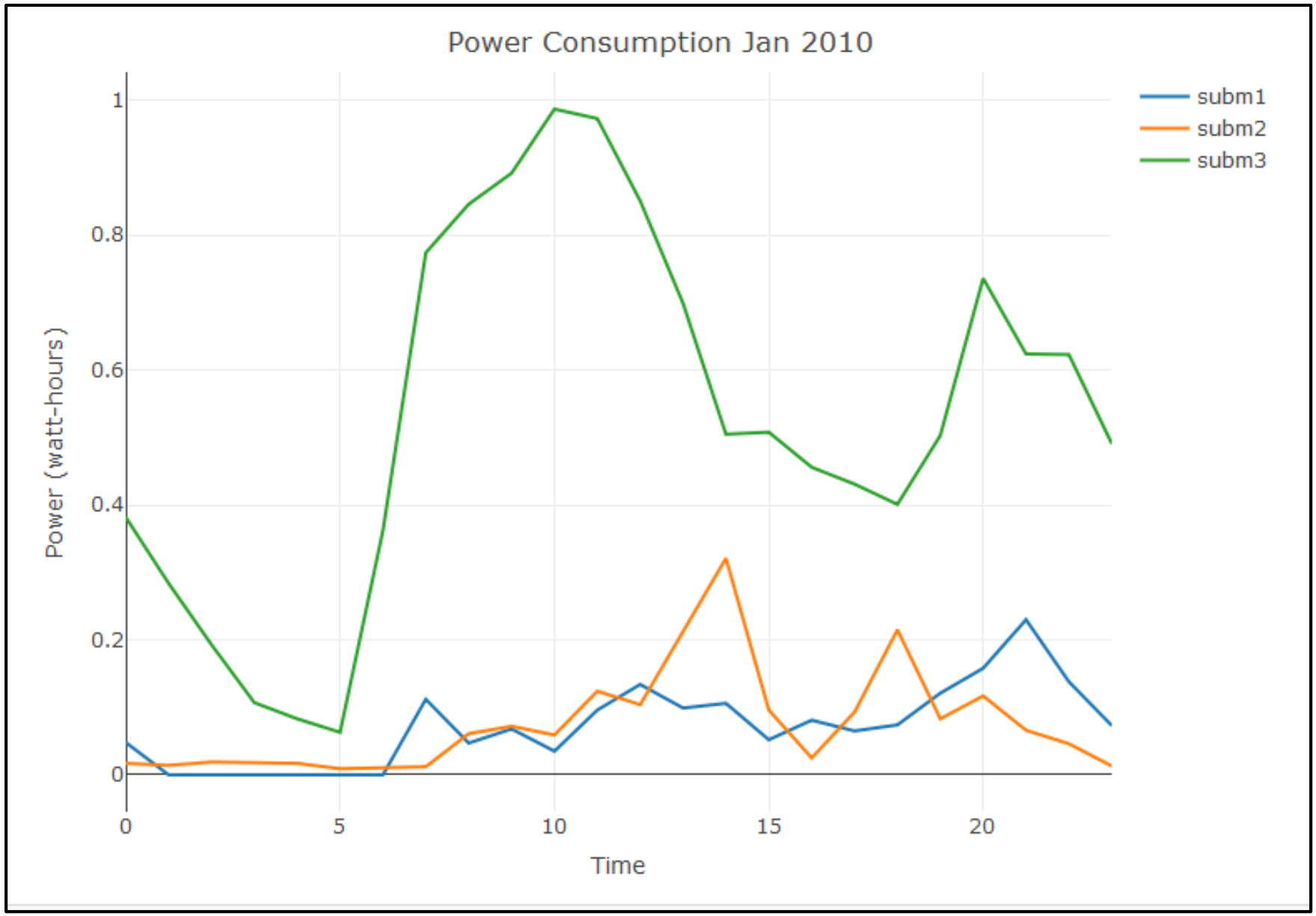
2. Average daily consumption for each ESM by weekday for the winter seasons (Dec-Feb)

Jan 2007 through Oct 2010 period



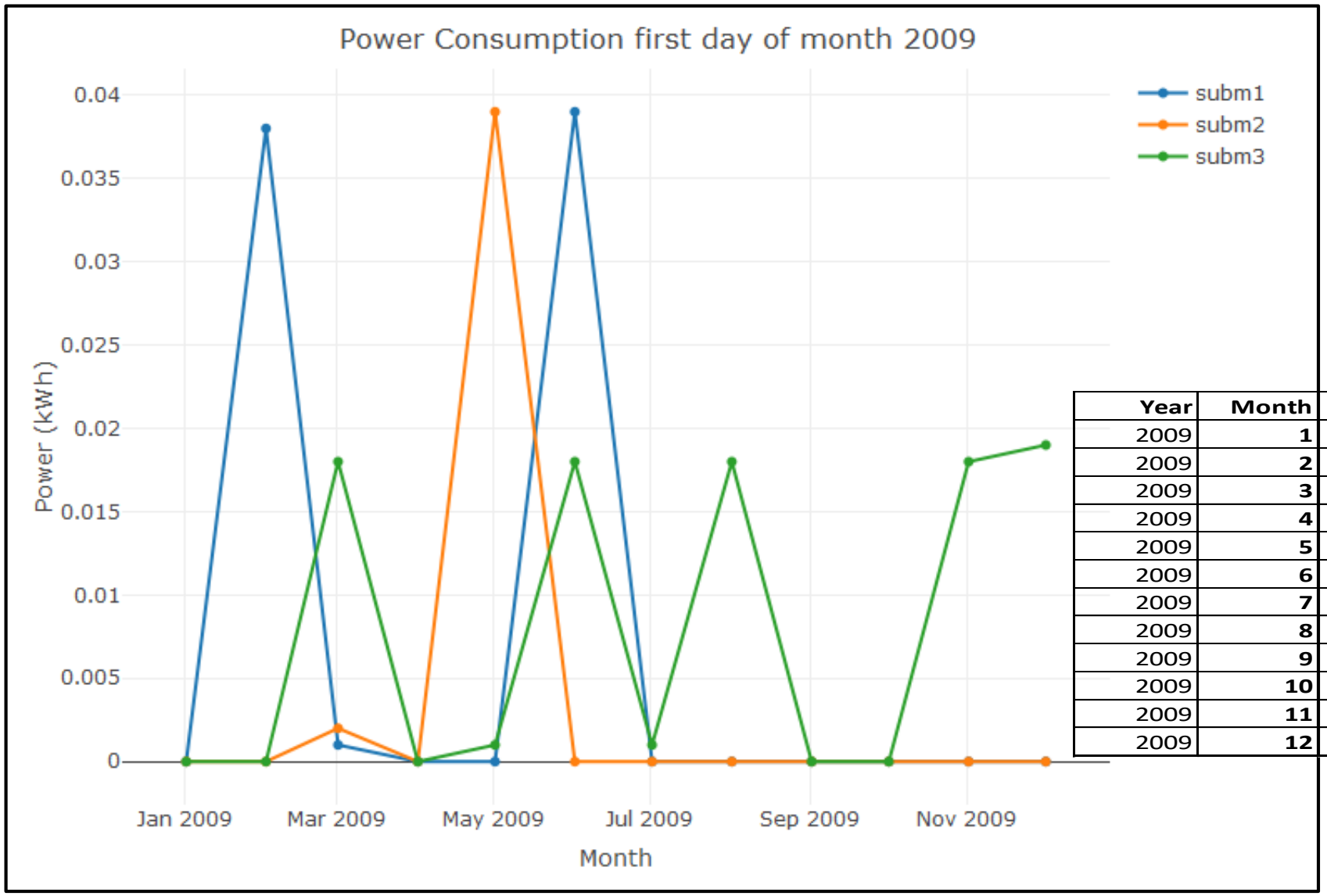
	WeekDay	SM1	SM2	SM3	DateTime
1	Sunday	3.08	3.18	10	01/07/2007 00:00:00
2	Monday	1.19	0.817	11	01/01/2007 00:00:00
3	Tuesday	1.12	2.07	11.3	01/02/2007 00:00:00
4	Wednesday	1.79	2.97	10.7	01/03/2007 00:00:00
5	Thursday	1.22	0.889	10.8	01/04/2007 00:00:00
6	Friday	1.31	1.43	11.4	01/05/2007 00:00:00
7	Saturday	3.22	2.96	12.6	01/06/2007 00:00:00

3. The average hourly kWh consumption for each hour of the day during January 2010



	Hour	SM1	SM2	SM3
1	0	0.048	0.017	0.382
2	1	0	0.014	0.284
3	2	0	0.019	0.193
4	3	0	0.018	0.107
5	4	0	0.017	0.083
6	5	0	0.009	0.063
7	6	0	0.011	0.362
8	7	0.112	0.012	0.774
9	8	0.047	0.061	0.846
10	9	0.068	0.072	0.892
11	10	0.035	0.059	0.987
12	11	0.096	0.124	0.973
13	12	0.134	0.104	0.851
14	13	0.099	0.212	0.699
15	14	0.106	0.321	0.505
16	15	0.052	0.096	0.508
17	16	0.081	0.025	0.456
18	17	0.065	0.093	0.431
19	18	0.074	0.215	0.401
20	19	0.121	0.083	0.503
21	20	0.158	0.117	0.736
22	21	0.23	0.066	0.624
23	22	0.138	0.046	0.623
24	23	0.073	0.013	0.491

4. Power consumption at 12:00pm (noon) on the 1st day of each month in 2009



Year	Month	SM1	SM2	SM3	DateTime
2009	1	0	0	0	01/01/2009 12:00:00
2009	2	0.038	0	0	02/01/2009 12:00:00
2009	3	0.001	0.002	0.018	03/01/2009 12:00:00
2009	4	0	0	0	04/01/2009 12:00:00
2009	5	0	0.039	0.001	05/01/2009 12:00:00
2009	6	0.039	0	0.018	06/01/2009 12:00:00
2009	7	0	0	0.001	07/01/2009 12:00:00
2009	8	0	0	0.018	08/01/2009 12:00:00
2009	9	0	0	0	09/01/2009 12:00:00
2009	10	0	0	0	10/01/2009 12:00:00
2009	11	0	0	0.018	11/01/2009 12:00:00
2009	12	0	0	0.019	12/01/2009 12:00:00

Power Consumption Forecast and Visualizations (1)

- In our forecast, we focus on ESM #3 to predict future power consumption.
- We complete 3 separate types of analysis. These analysis are called Time Series Analysis because each power consumption record that submeters collect every minute has a time stamp within this continuous time period. ESM data are all time dependent, covering a particular contiguous time period.

1. We use a set of equations denoted as the Time Series Linear Model and Forecast power consumption
   a) Filter out yearly data for ESM3 for the years 2007 through 2009. Forecast ESM3 power consumption for the years 2010 and 2011.
   b) Filter out total kWh per month for ESM3 for the months Jan 2007 through October 2010. Forecast for Nov. 2010 through Dec 2011.

Power Consumption Forecast and Visualizations (2)

2. We investigate the components of the power consumption. This is called decomposition.

   Decomposing the time series means separating the time series into three components: a trend component, a seasonal component and an irregular component.

   a) Filter out monthly data for ESM3 to get kWh by month over the Jan 2007 thru Oct 2010 time period, decompose this data into seasonal, trend and random components.

   b) Filter out ESM3 records that shows kWh by the hour over each day during Feb 2010 and decompose this into seasonal, trend and random components.

Power Consumption Forecast and Visualizations (3)

3. Holt-Winters (HW) method can be used to forecast data points in a time series, when the data in this series is “seasonal”, i.e. repetitive over some period.

   We use ESM3 for the 4 seasons:

   - Winter 09/10 from Dec 2009 thru Feb 2010,
   - Spring 2010 from Mar 2010 thru May 2010,
   - Summer 2010 from Jun 2010 thru Aug 2010,
   - Fall 2010 from Sep 2010 thru end of the data on Nov. 2010.

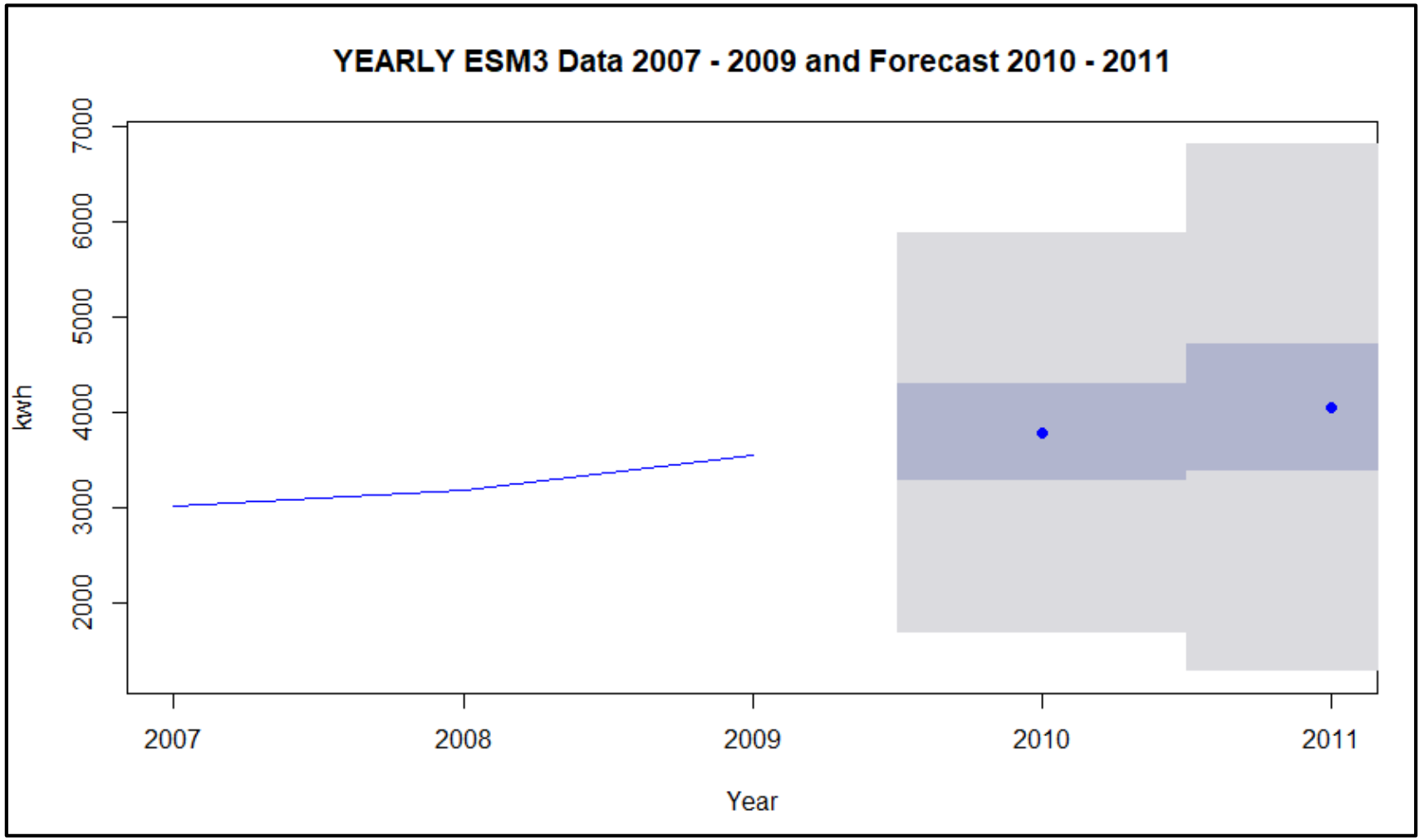
   We forecast ESM3 power consumption for the next 30 days for each of the 4 seasons.

   We first use (a) LM method for forecasting and decomposition as described in previous slide,

   Remove seasonal component and Apply (b) HW filtering method to fit a smooth line,

   (c) Forecast power consumption for the next 30 days for each season.

1. a) Yearly ESM3 data for 2007 through 2009 and Forecast of the power consumption for 2010 - 2011



	Year	SM3	DateTime
1	2007	3023	01/01/2007 09:19:00
2	2008	3179	01/01/2008 00:00:00
3	2009	3557	01/01/2009 06:10:00

	PointForecast	Lo 80	Hi 80	Lo 95	Hi 95
2010	3786.881	3278.68	4295.082	1688.773	5884.989
2011	4053.869	3381.582	4726.156	1278.333	6829.405

1. b) Monthly ESM3 data, Jan 2007 through Oct 2010 and Forecast for Nov 2010 through Dec 2011

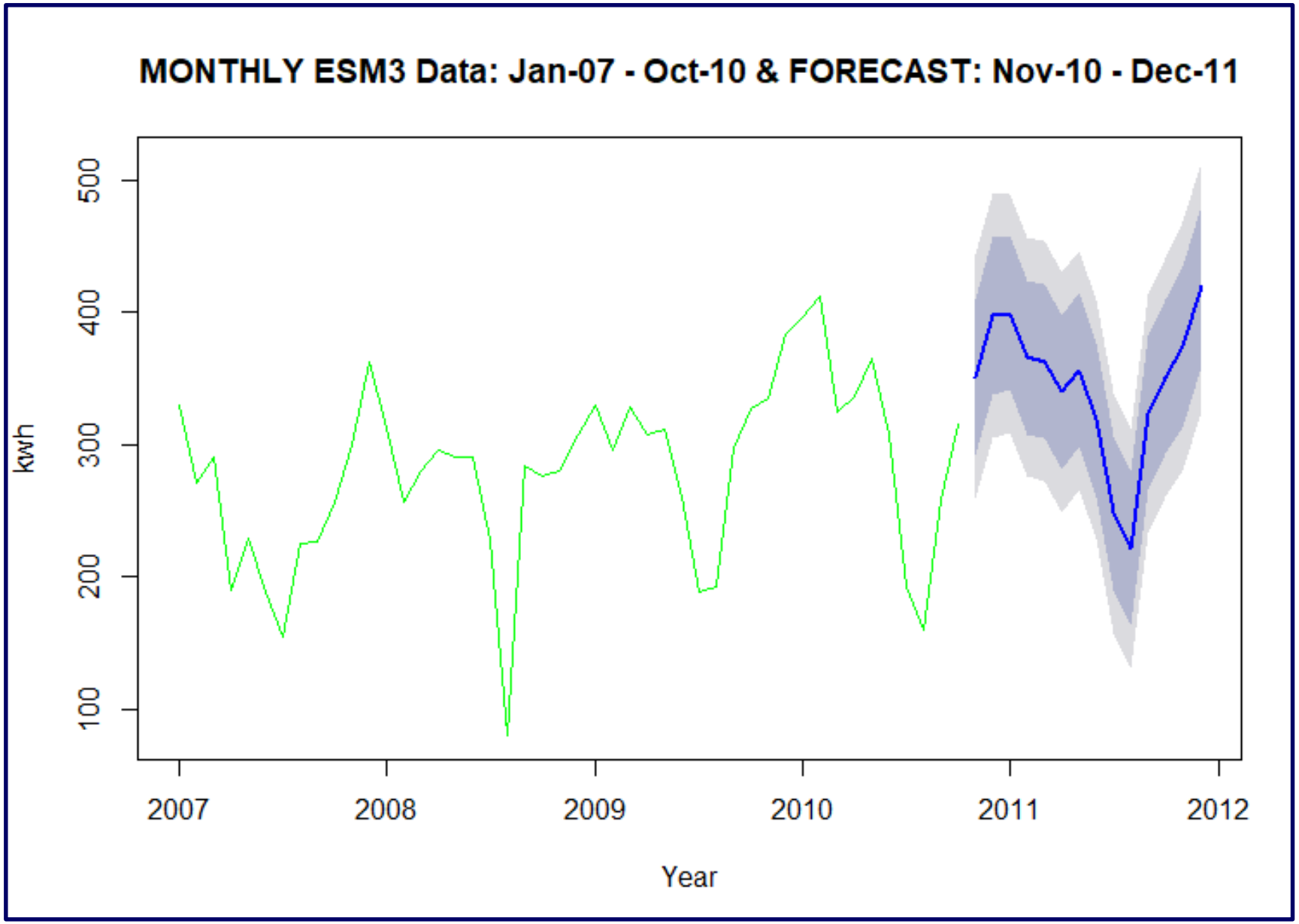
Monthly ESM3 data, Jan 2007 through Oct 2010

Year	Month	SM3	DateTime
2007	1	330	01/01/2007 00:00:00
2007	2	270	02/01/2007 00:00:00
2007	3	290	03/01/2007 00:00:00
2007	4	190	04/01/2007 00:00:00
2007	5	229	05/01/2007 00:00:00
2007	6	189	06/01/2007 00:00:00
2007	7	155	07/01/2007 00:00:00
2007	8	225	08/01/2007 00:00:00
2007	9	226	09/01/2007 00:00:00
2007	10	256	10/01/2007 00:00:00
2010	2	412	02/01/2010 00:00:00
2010	3	324	03/01/2010 00:00:00
2010	4	336	04/01/2010 00:00:00
2010	5	365	05/01/2010 00:00:00
2010	6	307	06/01/2010 00:00:00
2010	7	193	07/01/2010 00:00:00
2010	8	160	08/01/2010 00:00:00
2010	9	258	09/01/2010 00:00:00
2010	10	316	10/01/2010 00:00:00

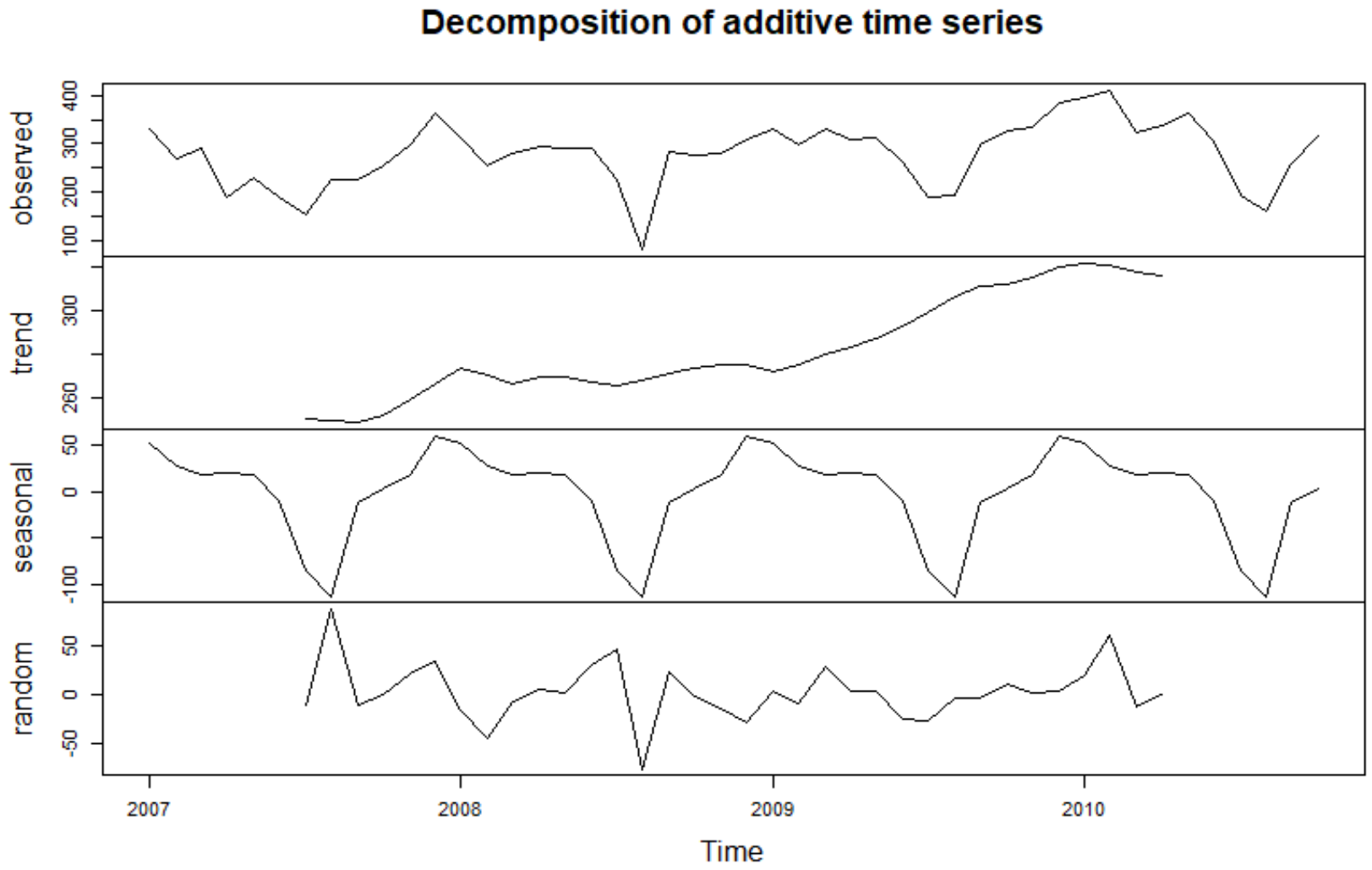
Forecast for Nov 2010 through Dec 2011

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Nov-10	351.1559	292.2140	410.0978	259.4569	442.8550
Dec-10	397.1039	338.1620	456.0458	305.4049	488.8030
Jan-11	399.1654	341.1026	457.2283	308.8340	489.4969
Feb-11	365.8654	307.8026	423.9283	275.5340	456.1969
Mar-11	363.0754	305.0126	421.1383	272.7440	453.4069
Apr-11	339.6254	281.5626	397.6883	249.2940	429.9569
May-11	356.2827	298.2198	414.3455	265.9513	446.6141
Jun-11	318.8117	260.7488	376.8745	228.4803	409.1431
Jul-11	248.0702	190.0073	306.1330	157.7388	338.4016
Aug-11	221.6874	163.6246	279.7503	131.3560	312.0189
Sep-11	323.6137	265.5508	381.6765	233.2823	413.9451
Oct-11	351.1327	293.0698	409.1955	260.8013	441.4641
Nov-11	374.0949	313.2450	434.9448	279.4275	468.7623
Dec-11	420.0429	359.1930	480.8928	325.3755	514.7103

1. b) Monthly ESM3 data, Jan 2007 through Oct 2010 and Forecast for Nov 2010 through Dec 2011

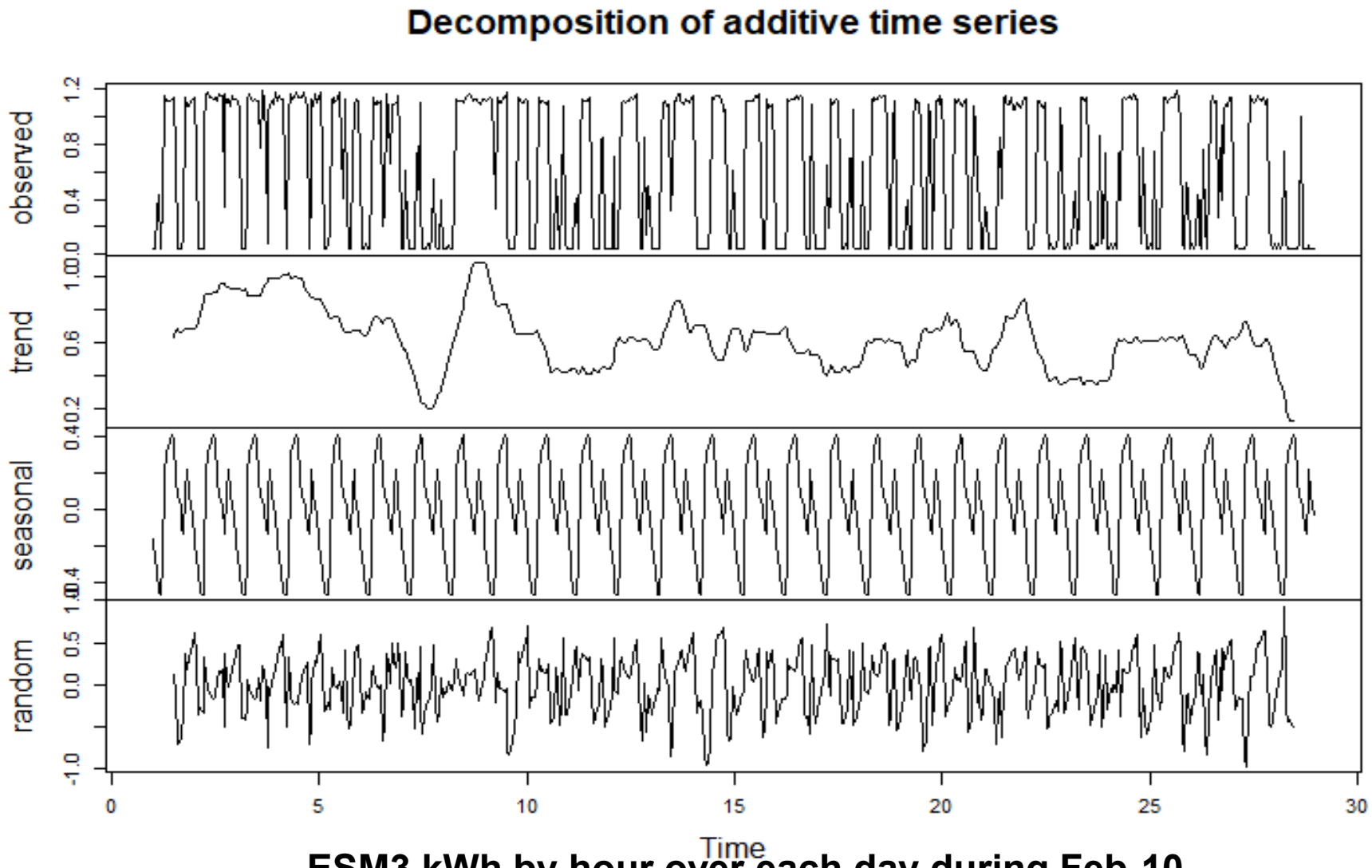


2. a) Decomposition of Monthly ESM3 Data Jan-07 thru Oct-10



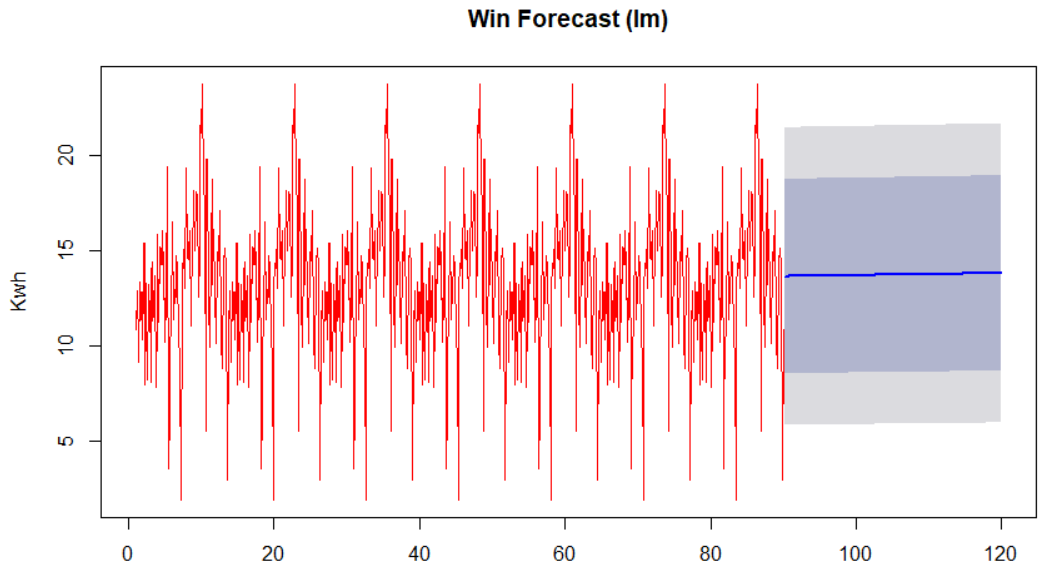
Decomposition of Monthly data for ESM3 over the Jan 2007 thru Oct 2010 time period into seasonal, trend, and random components.

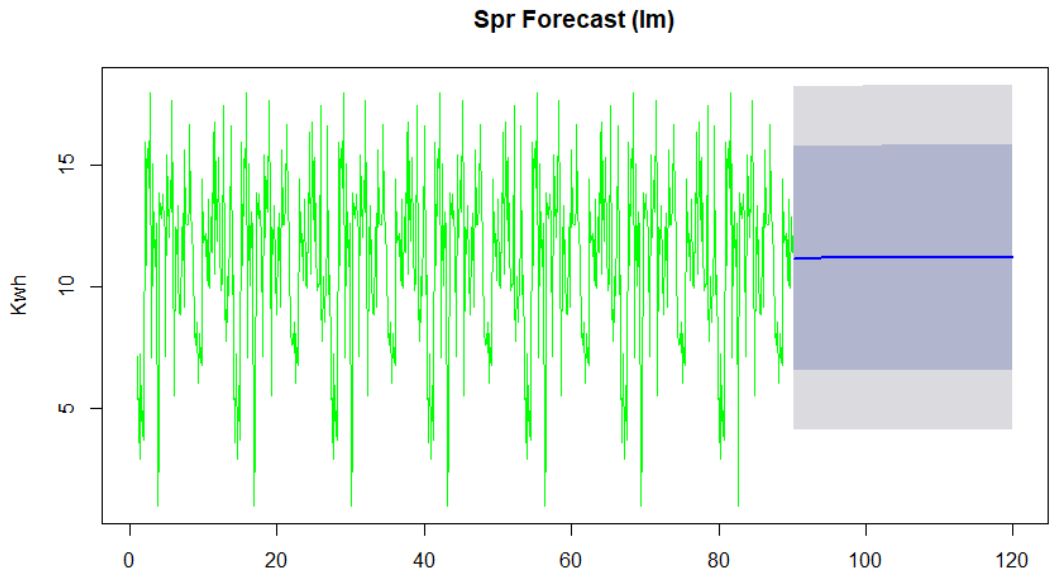
2. b) ESM3 kWh by HOUR over each day during Feb-10, Decomposed into seasonal, trend and random components

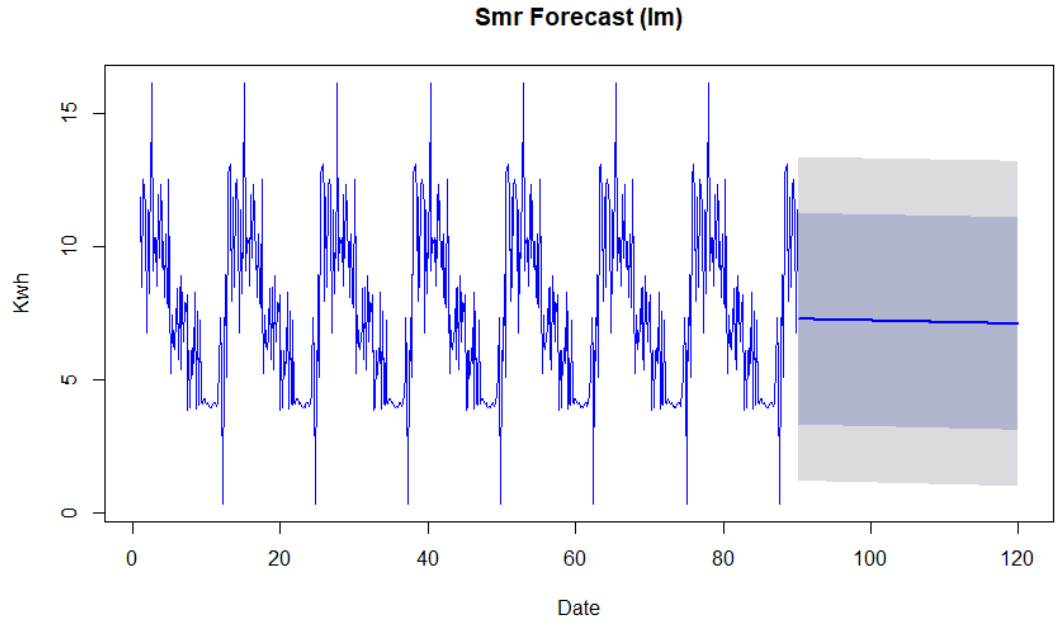


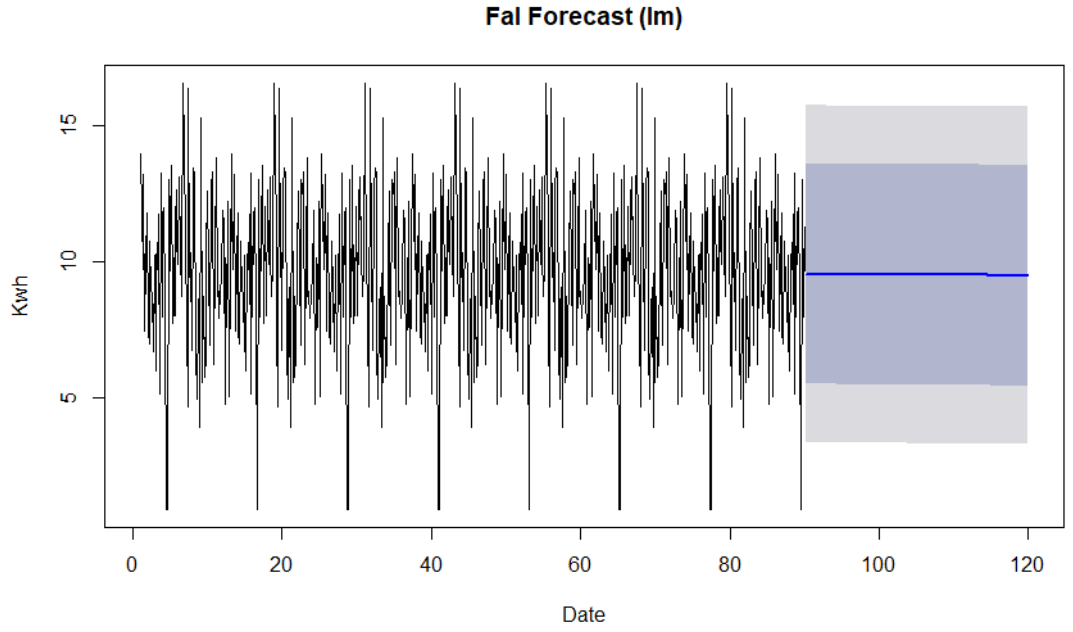
ESM3 kWh by hour over each day during Feb-10, Decomposed into seasonal, trend, and random components

3. a) LM Forecast for 4 seasons



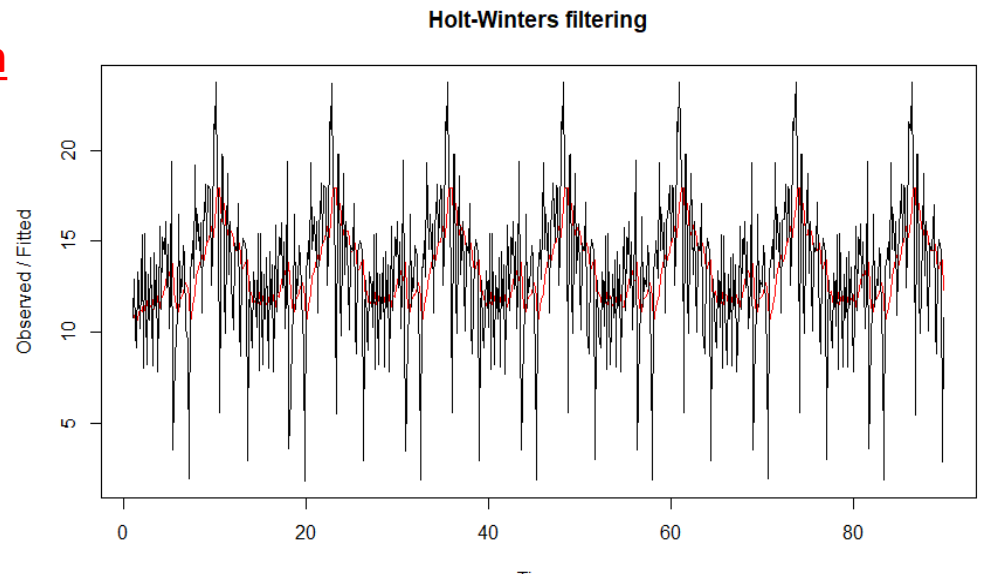




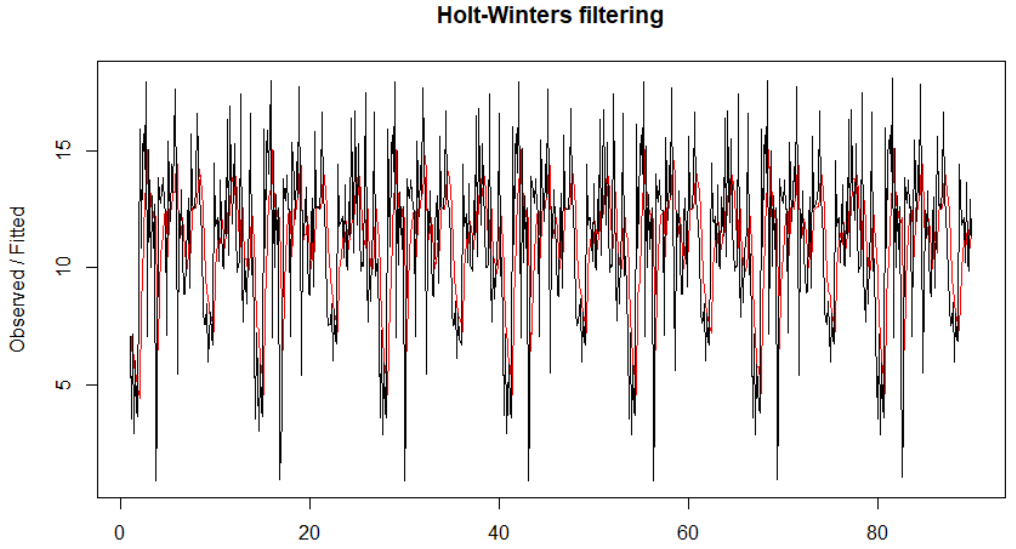


3. b) HW filtering method to fit a smooth line (HW exponential smoothing)

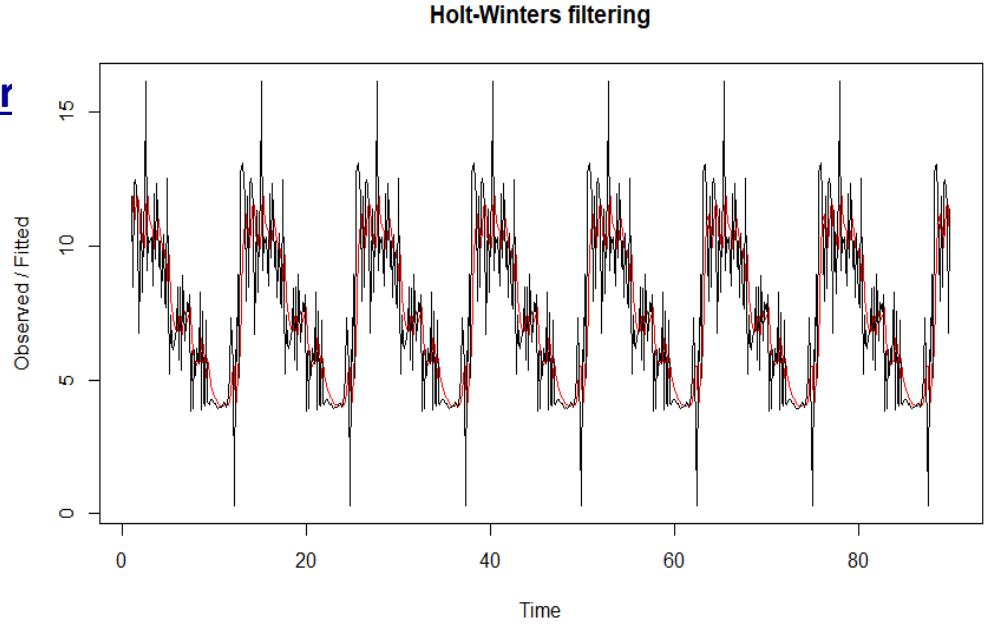
Win



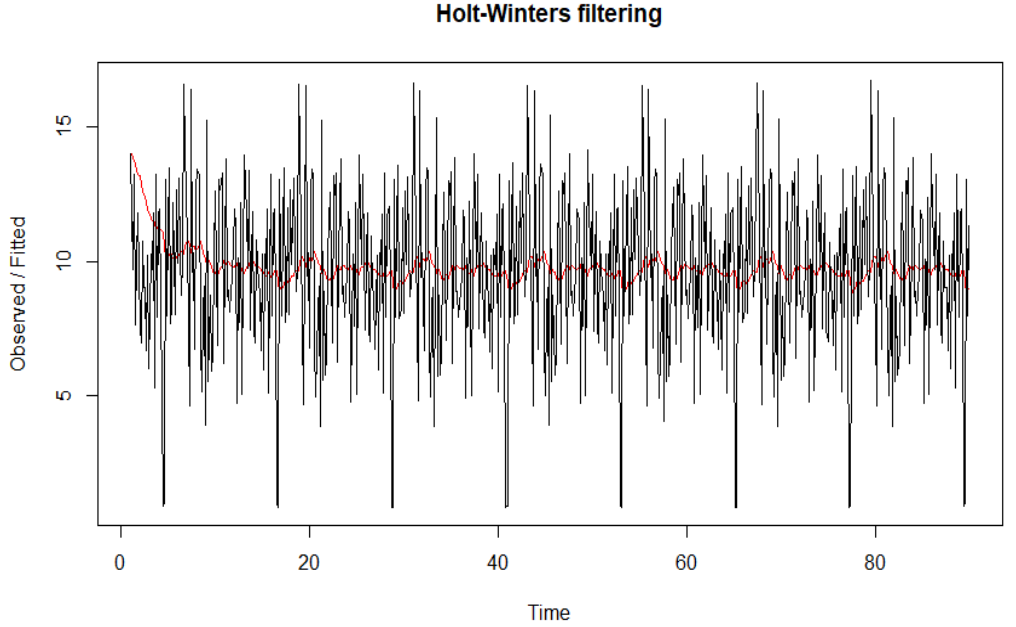
Spr



Smr

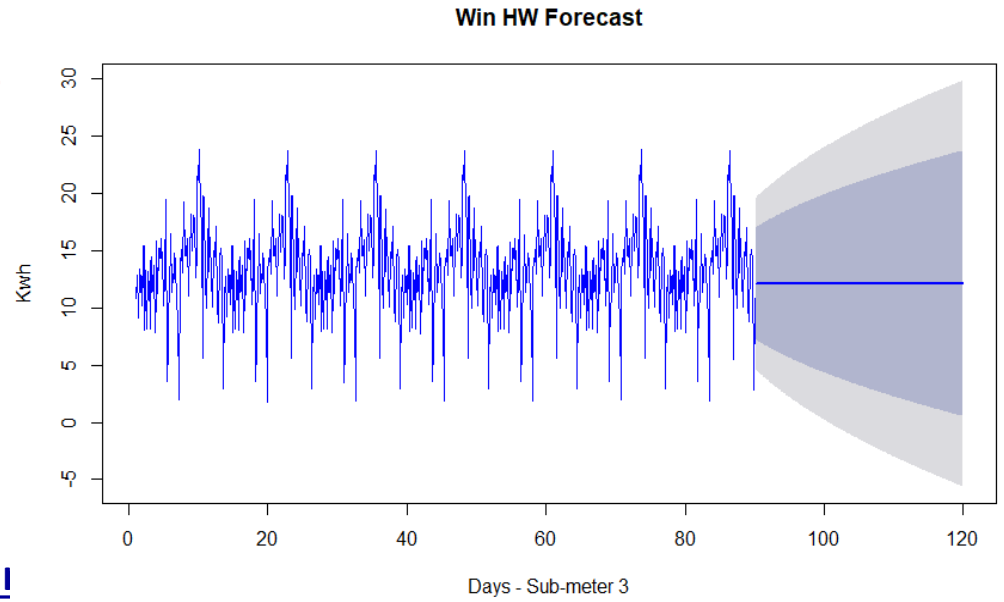


Fal

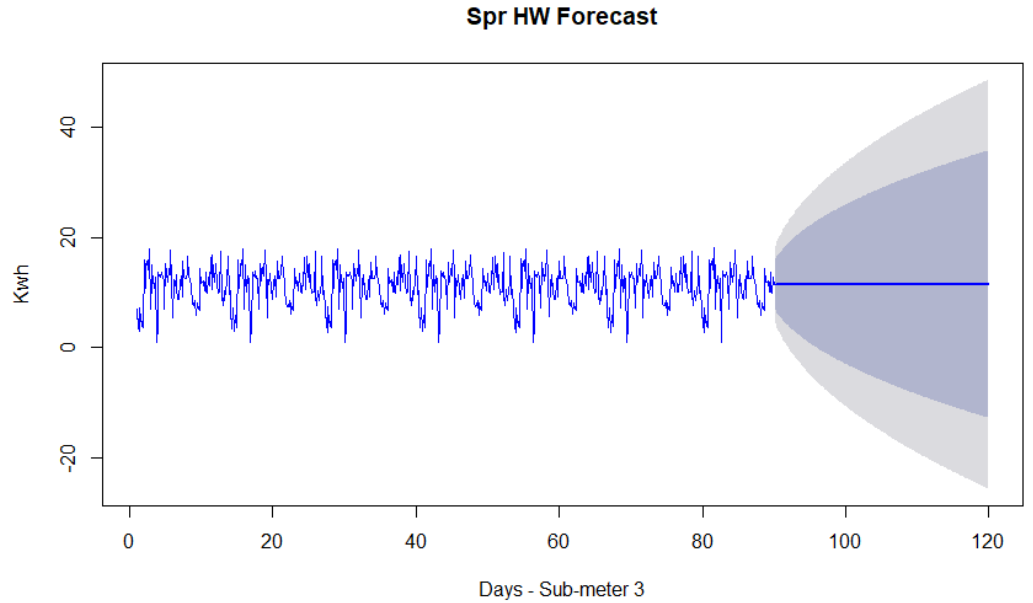


3. c) HW Forecast for 4 seasons

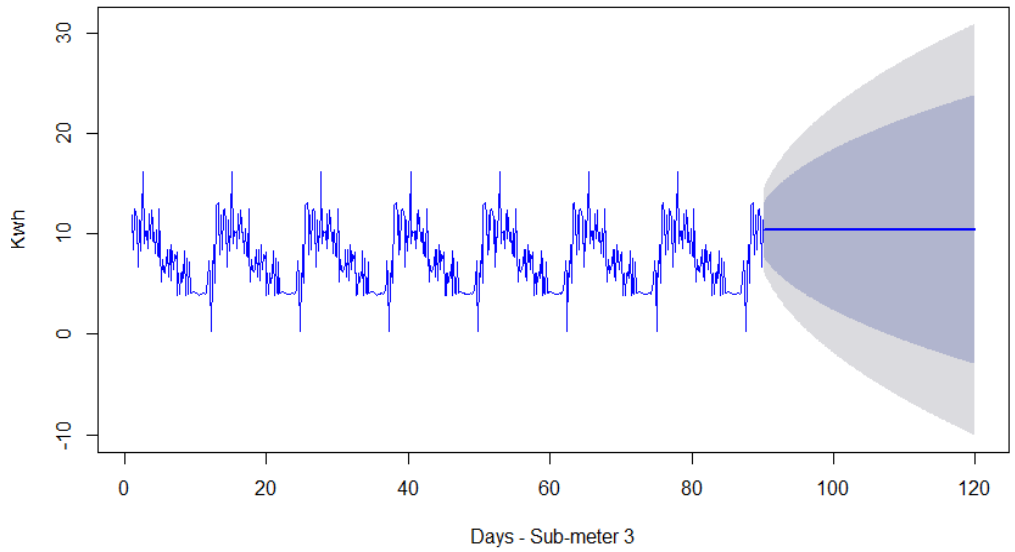
Win



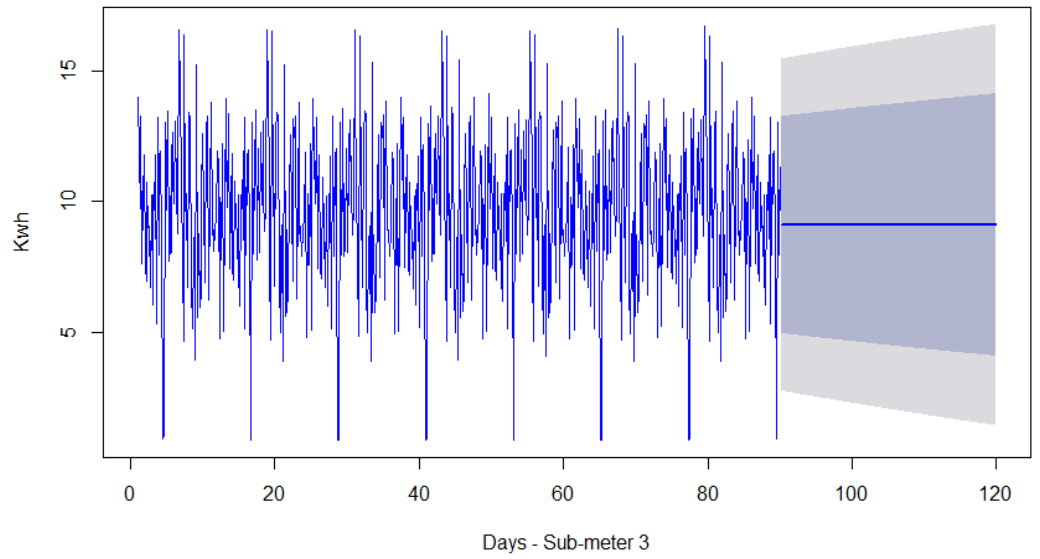
Spr



Smr



Fal



BUSINESS RECOMMENDATIONS

1. IoT Analytics presented here should be a live process where the results and visualizations are updated periodically.
2. ESM3 indicates that highest power consumption is due to air conditioner and water heater. Options to reduce this consumption, .e.g., tankless water heater would be worth the investment.
3. Forecast indicates increasing power consumption at ESM3, but declining consumption at other areas of the house. Air conditioner can be programmed to reduce this consumption between 7 AM and 2 PM, 9PM and Midnight.
4. Seasonal component of power consumption in the decomposition of monthly data as well as the HW forecasting show that winter is the season with highest power usage. Switching to natural gas for some equipment such as the water heater would potentially reduce power consumption and the energy cost.
5. IoT analytics should be coupled with Hourly, Daily, Monthly power cost to help home owner see their energy expenses.

LESSONS LEARNED

- Creating daily, weekly, monthly, yearly ESM data is necessary to see power consumption patterns. Granularity describes the frequency of observations within a dataset such as the ESM data acquired very minute. Adjusting granularity to larger time frames maximizes the information to be gained from ESM records.
- Transforming ESM records acquired every minute to groups of wider time periods and obtaining average consumption in those groups require careful filtering and averaging.
- Transforming data to meaningful time frames allow us to focus on the periods of time that highlight patterns of power usage in Visualization, i.e. plots and charts.
- Visualization is critical to seeing and understanding what the data are trying to tell us.