APPENDIX J

Non Household Demand Forecast

ABSTRACT

Appendix for the forecast of Non household demand. Two approaches are presented, a Top Down and a Bottom Up forecast. Whilst the Top down approach uses a regression on total demand, the Bottom Up forecast uses a regression models on of the SIC07 sub components of demand which sum to a total demand forecast.

Contents

Overview of Approach

Two approaches are presented, a Top Down and a Bottom Up forecast. Both methods draw from the same data sets using regional econometric factors and historic weather conditions. Whilst the Top down approach regresses on total measured non-household demand, the Bottom Up forecast uses many regression models on the SIC07 sub components of demand which are aggregated up to total demand.

Data

Econometric Data

Raw Data

Econometric data was sourced from the Office for National Statistics (ONS) 'Table 6: Regional gross value added (income approach) reference tables' (15 Dec 2015). The data set provides a breakdown of GVA for the period 1997 to 2015. The table provides estimates of GVA by SIC07 for the following groups:

Code	SIC07 Group
A	Agriculture, forestry and fishing
BDE	Production
C	Manufacturing
F	Construction
GHI	Distribution; transport; accommodation and food
J	Information and communication
K	Financial and insurance activities
L	Real estate activities

MN	Business service activities
OPQ	Public administration; education; health
RST	Other services and household activities
All	All industries

TABLE 1: BREAKDOWN OF SIC07 GROUPS

The dataset is broken down into ONS Nomenclature of Territorial Units for Statistics (NUTS) areas. Portsmouth Water sits across three NUTS areas; Portsmouth, South Hampshire, and West Sussex (South West). The data from all three areas is initially included but screened or combined depending on the approach.

The three NUTS areas are broadly comparable in terms of the makeup of GVA between the SIC categories with some individual differences.

- The Portsmouth area is characterised by having a larger proportion of GVA made up by 'Public administration; education; health' and 'Information and Communication Services' with little 'Agriculture, forestry and fishing'. Portsmouth also has a higher proportion of GVA made up of Manufacturing than the UK average.
- West Sussex has a high proportion of 'Real Estate activities' and 'Agriculture, forestry and fishing'. West Sussex also has higher proportion of 'Financial and insurance activities' but less than the UK average.
- In South Hampshire, 'Real Estate activities' is proportionally higher than the UK average but lower than West Sussex. South Hampshire also has a very notably large %GVA of Manufacturing, over twice that of the UK average.

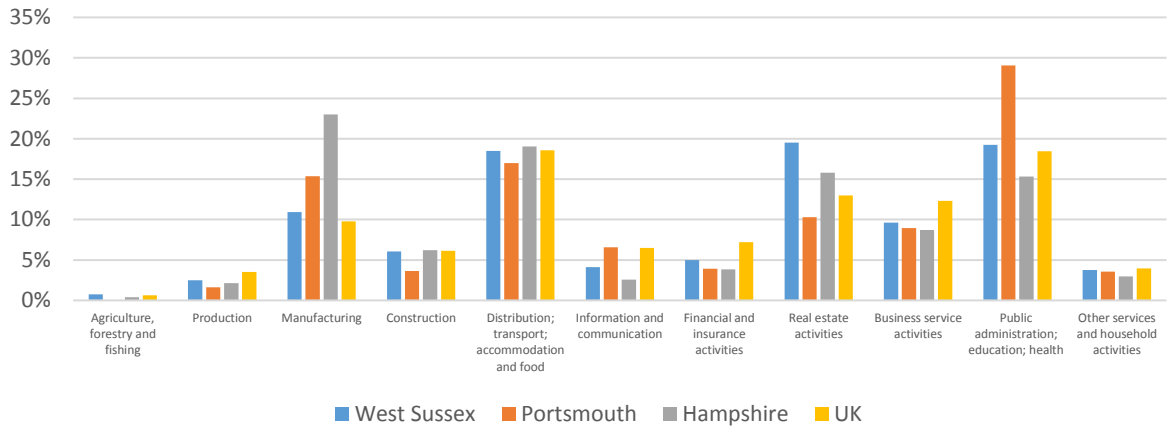

FIGURE 1: 2015/16 % SPLIT OF GVA BY AREA

Forecasted Data

The raw data does not include projections of the regional GVA. A standard an Autoregressive Integrated Moving Average (ARIMA) approach is used to forecast the GVAs for each of the economic variables.

Figure 2 presents the aggregated cumulative growth for the three NUTS areas. Pre 2015 data points are outturn values whilst post 2015 values are forecasted. The forecast shows continued growth at a reasonable

level where a strong trend is present. Where strong consistent trend is not present the GVA tends to remain flat.

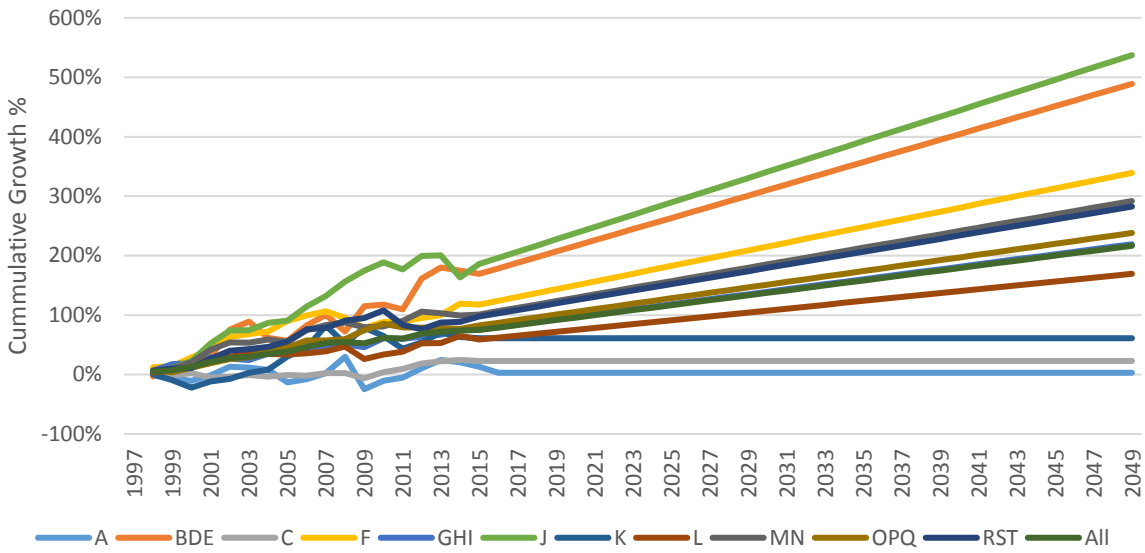

FIGURE 2: CUMULATIVE GROWTH IN SIC07 GVA (ALL AREAS)

Weather Data

Weather data consists of rainfall (mm) and temperature (°C) data which. Rainfall data is collected by Portsmouth Water whilst temperature data is sourced from publically available Met Office HadCET data.

The raw data is in a daily format, from the raw data the following variables are created:

- Total Rainfall (Annual)
- Total Summer Rainfall (May-Sept)
- Average Max Temperature (Annual)
- Average Summer Max Temperature (May-Sept)
- Total Dry Days (Annual)
- Total Dry Days of Summer (May-Sept)

Weather factors are presumed to remain flat over time with 2015/16 notionally representing the normal year and 2013/14 notionally representing the dry year. Generally, most measured demand is insensitive to weather variables.

Data Selection

For both the Top Down and Bottom Up models a Multiple Linear Regression (MLR) approach is applied. One of the challenges of modelling measured non household demand in this instance is the dataset is very collinear where all the variables are strongly correlated to each other, Figure 1. The economic factors tend to be highly positively correlated which each other and highly negatively correlated with measured consumption. Furthermore, some weather factors have moderate relationships with other weather factors and some economic factors. Collinearity is challenging with MLR type models as there is no means of identifying which variable is producing a decrease in volumes over time and therefore must be accounted for. Collinearity is handled differently in the Top Down and Bottom Up models;

- **Bottom Up**: Only a subset of the relevant variables is made available for each SIC07 group a final selection is made based upon the variables which give the best Residual Means Squared Error (RMSE).
- **Top Down**: Principle Component Analysis (PCA) is used to combine the economic and weather predictor variables into different uncorrelated variables.

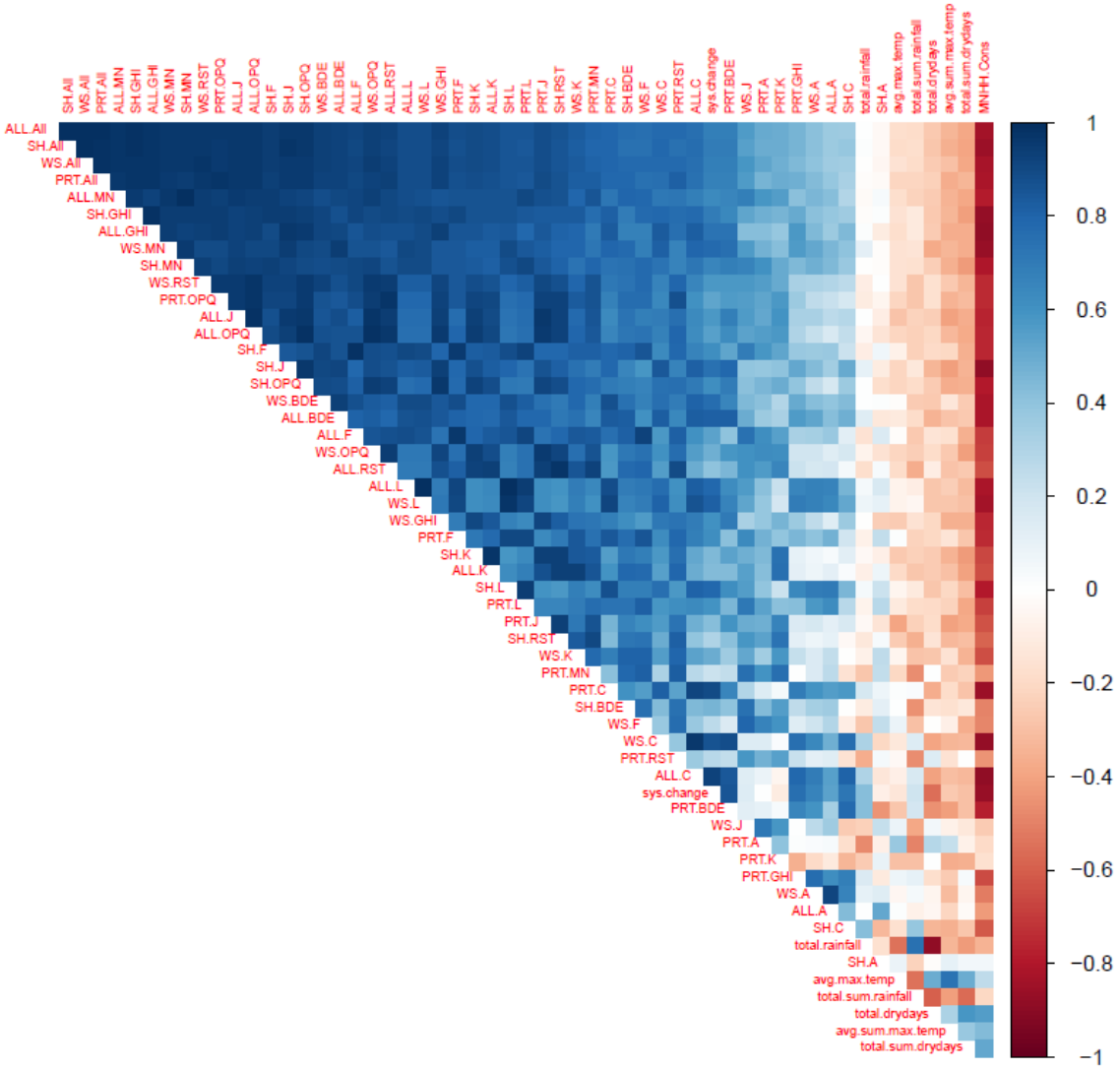

FIGURE 3: CORRELATION MATRIX

Principal Component Analysis

Principal Component Analysis (PCA) is a method of extracting important variables from a large set of variables available in a data set. It extracts a smaller number of variables from a data set with many variables with the motive to capture as much information as possible. In PCA the first component always contains the most amount of information which diminishes with each new component.

In this analysis two components, or features, are extracted from the data. The components use weightings from the original variables which can either be positive or negative, Figure 4.

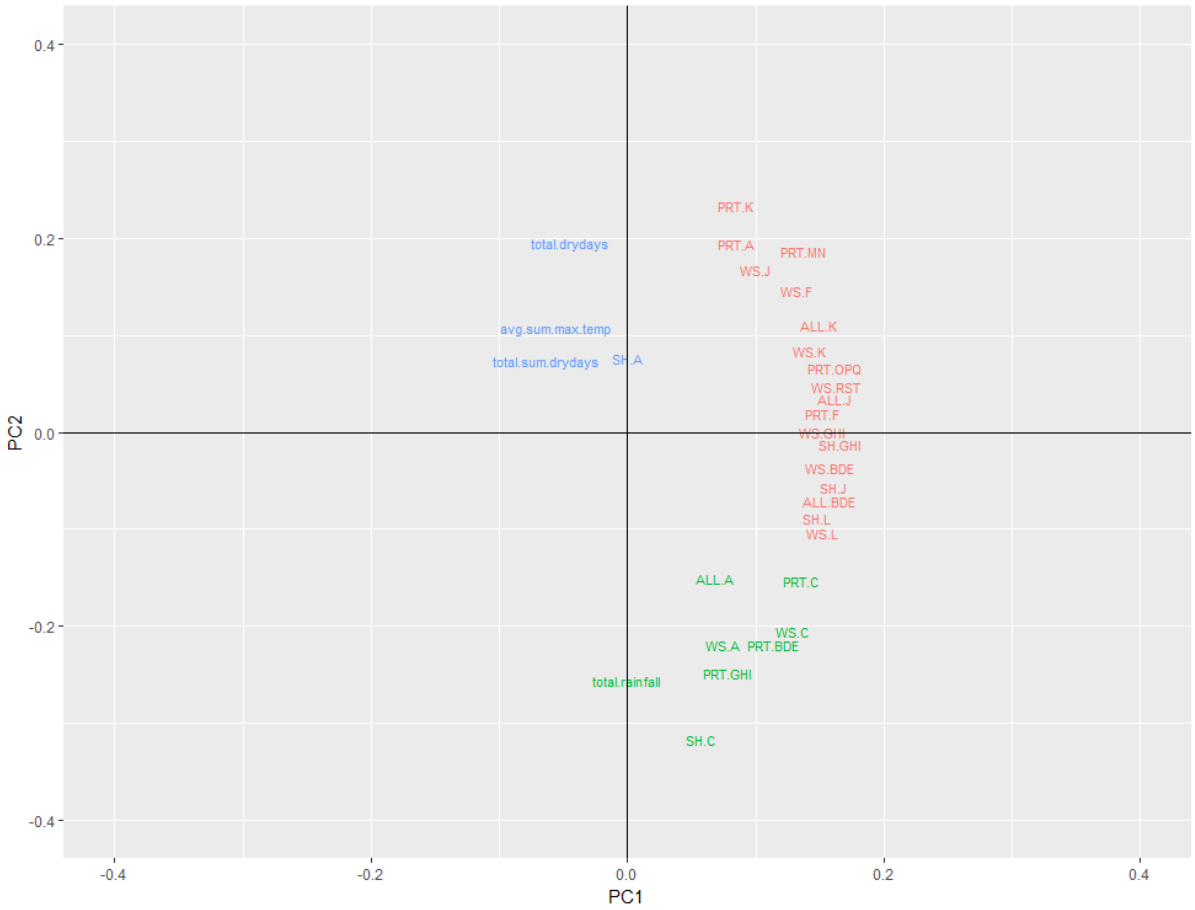

FIGURE 4: PRINCIPLE COMPONENTS WEIGHTING

Component	Description
PC1	PC 1 largely captures broadly equal weighting of the economic GVAs with a slightly weaker positive weighting for Agriculture and Manufacturing activities.
PC2	PC2 reflects most of the variance in the data as a result of weather. It has a positive weighting for 'dry' weather factors and a negative weighting for 'wet' weather factors. It also includes a negative weighting for Agriculture and Manufacturing activities.

TABLE 2: DESCRIPTION OF PRINCIPAL COMPONENTS

Model Results

Top Down

The Top-Down model produces a very good result with an R2 score of 0.96.

Mean Average Percentage Error (%)	R^2
2.01	0.96

TABLE 3: RESULTS OF TOP DOWN MODEL

Bottom Up

The Bottom Up forecasts also produce a good result in most instances.

SIC07 Group	Mean Average Percentage Error (%)	R^2
A	4.51	0.97
B	5.80	0.85
C	3.96	0.81
D	13.78	0.87
E	7.21	0.98
F	4.78	0.94
G	8.15	0.86
H	6.36	0.95
I	2.64	0.96
J	29.95	0.97
K	2.9	0.93
L	9.32	0.93
M	6.27	0.85
N	6.21	0.90
O	4.65	0.98
P	3.71	0.99
Q	3.02	0.98
R	1.51	0.99
S	2.84	0.99
Defence	4.68	0.99

TABLE 4: BOTTOM UP MODEL RESULTS

Plots

The below plots show the sub forecasts for the bottom up models.

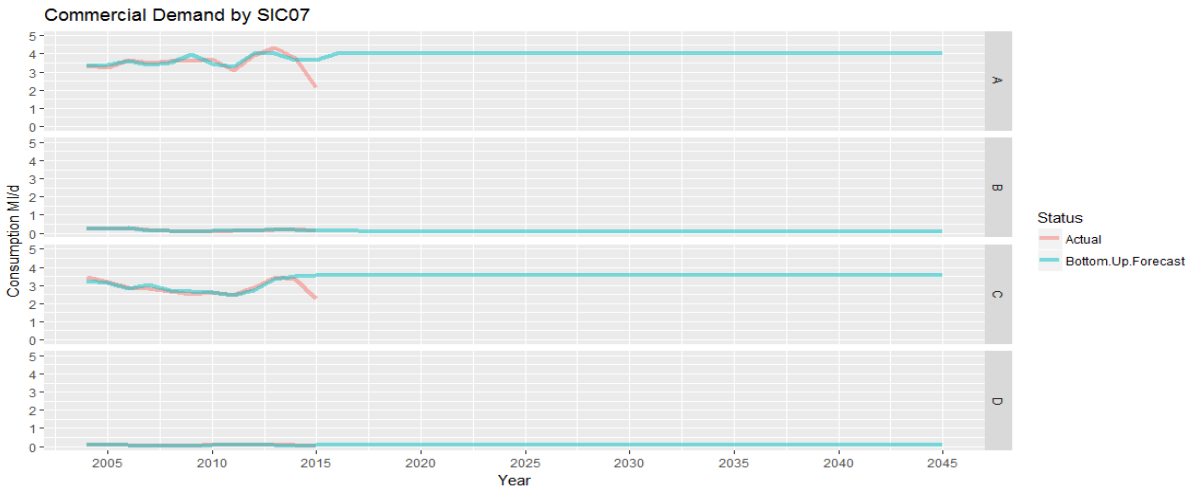

FIGURE 5 BOTTOM-UP OUTTURN PREDICTION & FORECAST SIC A-D

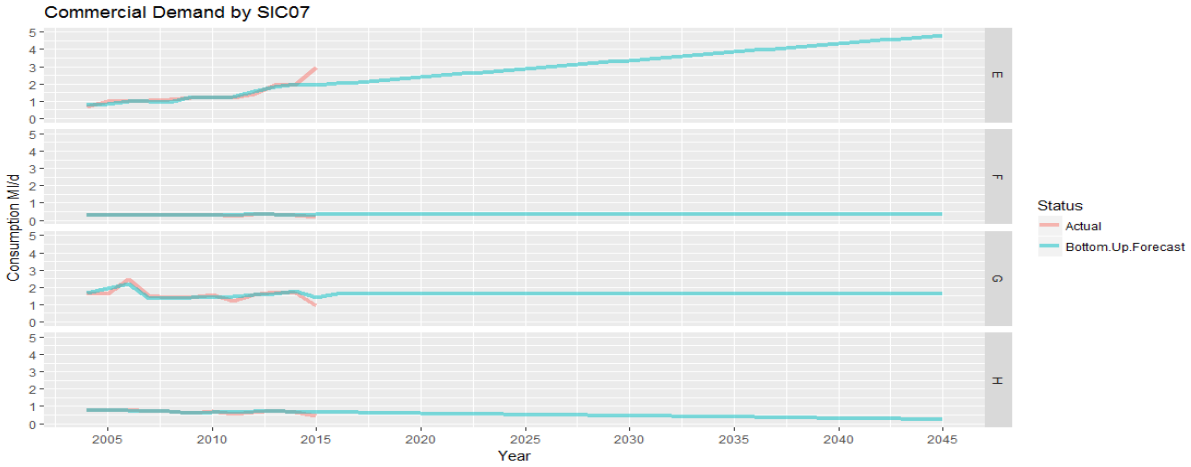

FIGURE 6 BOTTOM-UP OUTTURN PREDICTION & FORECAST SIC E-H

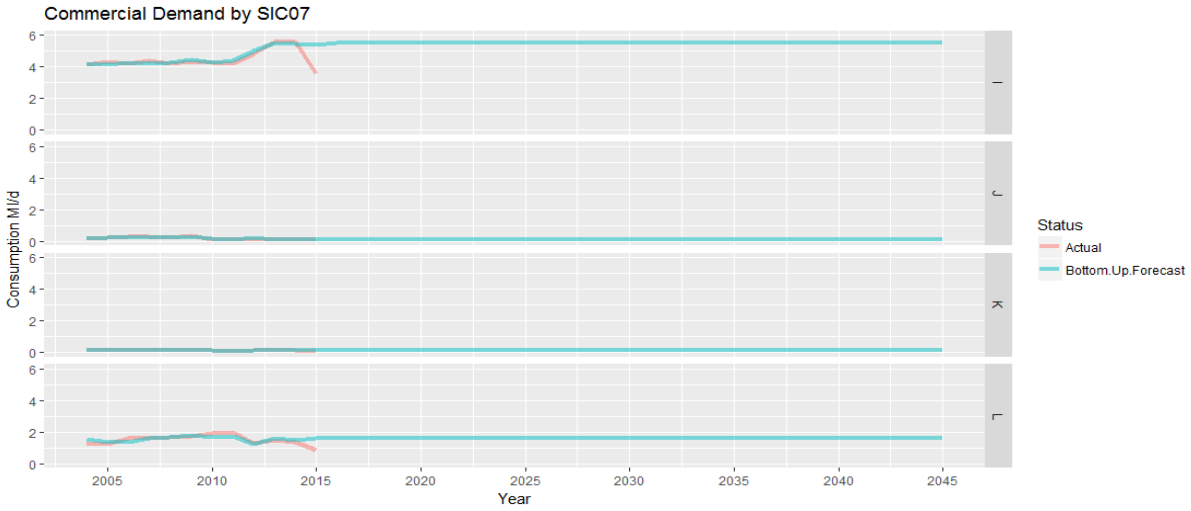

FIGURE 7 BOTTOM-UP OUTTURN PREDICTION & FORECAST SIC I-L

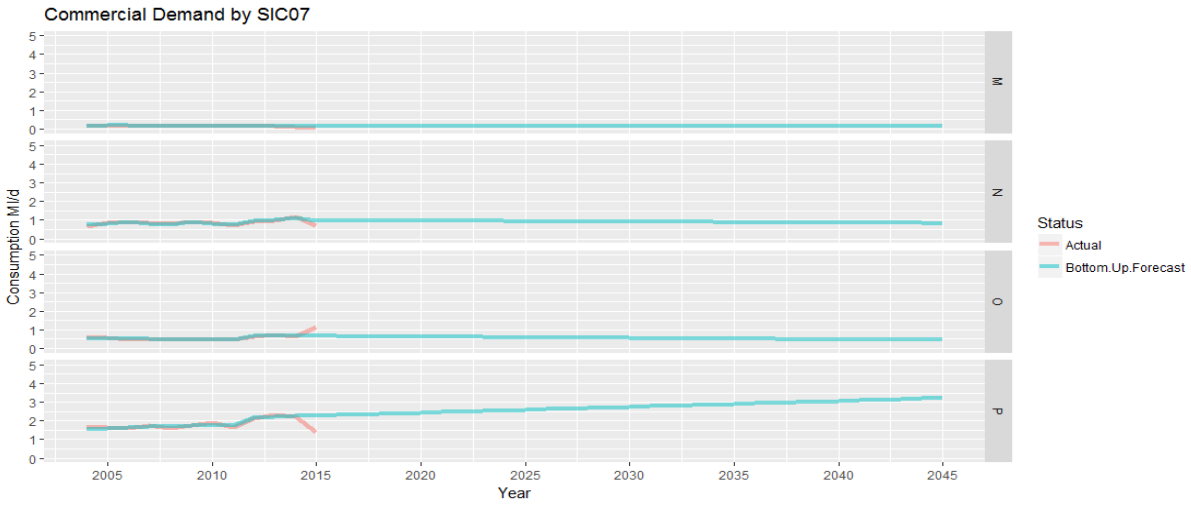

FIGURE 8 BOTTOM-UP OUTTURN PREDICTION & FORECAST SIC M-P

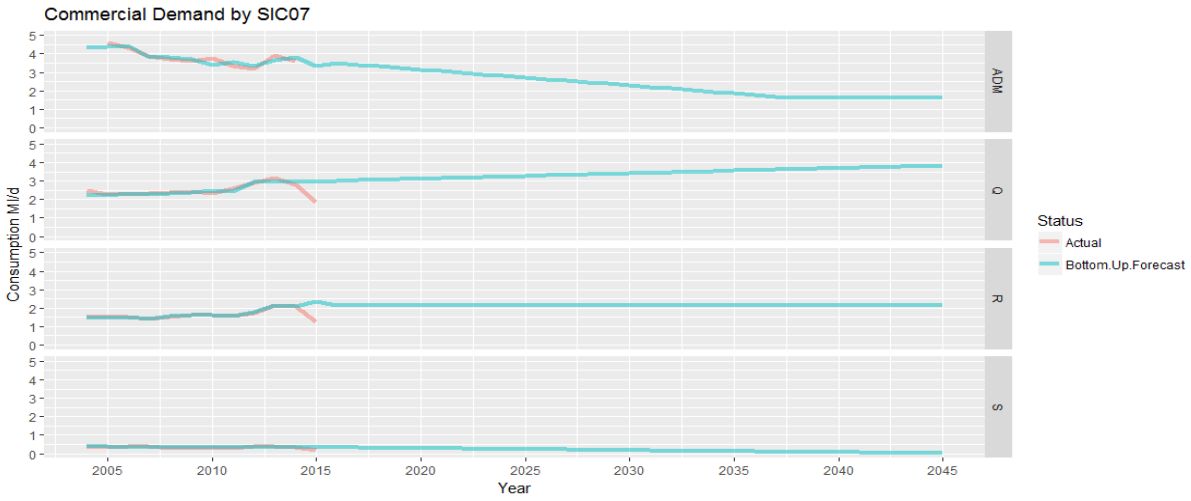

FIGURE 9 BOTTOM-UP OUTTURN PREDICTION & FORECAST SIC ADM-S

**ADM: Admiralty and Airforce*

Figure 10 shows the Top Down and Bottom Up forecasts.

The Bottom Up forecast shows demand to be relatively flat with a slight increase towards the end of the period as a result of increasing Production (E), Education (P) and Health (Q) activities.

The Top Down forecast shows demand to falling over the period eventually below 25 Ml/d by 2044/45.

Observing the outturn trend since 2005, measured non-household demand has continued to fall until 2012 when demand has since remained relatively flat. Both the 'Top Down' and 'Bottom Up' models produce a reasonable forecast given the uncertainties over the impact of the opening of the Non-household retail market and Brexit.

Given that both the 'Top Down' and 'Bottom Up models produce a reasonable forecast a 'Hybrid' approach is used which is simply an average of the two models. The resulting forecast shows a smooth decline in measured non-household demand eventually falling to 30 Ml/d by 2044/45.

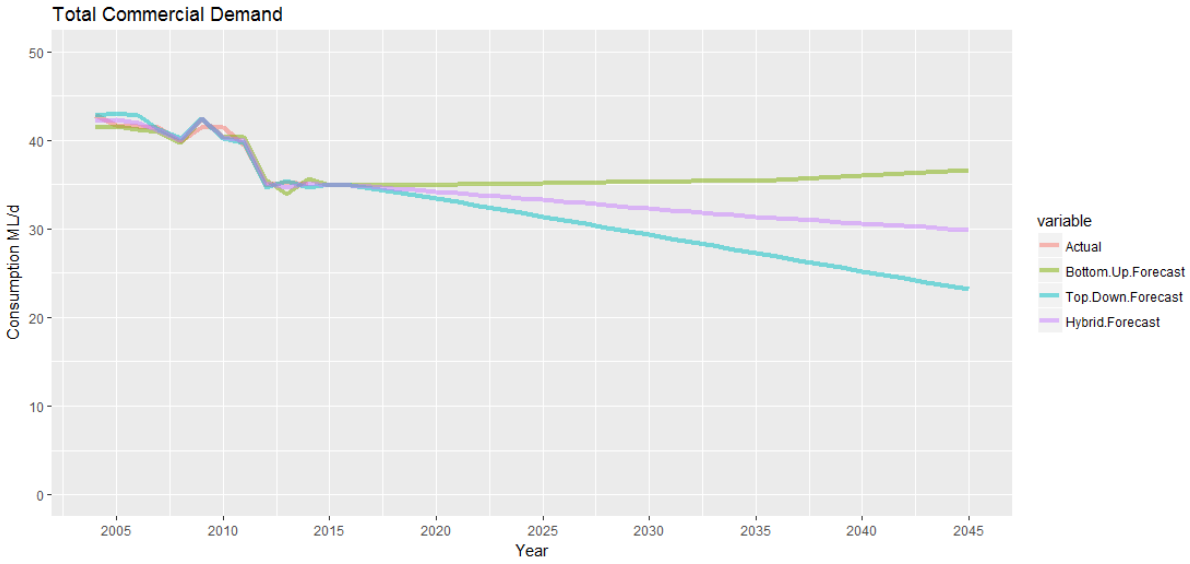

FIGURE 10 ALL FORECASTS