Southern Water Non-Household Forecast Update

Project no.:	2631	**Date:**	18/07/2023
Author(s):	Aidan Gibbons, Basil Langham	**Email:**	aidan.gibbons@artesia-consulting.co.uk; basil.langham@artesia-consulting.co.uk
Reviewed by:	Joe Cahill, Sarah Rogerson	**Reference:**	AR1528

1 Introduction

This technical note covers changes made to the Southern Water non-household consumption forecasting, as an update from the forecasts provided in 2020. This involves updates to demographics and historical consumption datasets, which are then applied to existing MLR models.

One of the primary drivers of this piece of work is due to several large-scale events (Covid-19 pandemic, Brexit, and the war in Ukraine) that have impacted the socio-economic situation in the United Kingdom, resulting in significant changes in economic and employment forecasts, as well as non-household water consumption through changing consumer behaviours.

The original process consisted of the training of an MLR model to predict NHH consumption trends using a combination of population, employment, economic and consumption metrics, as well as a MOSL flag for consumption data source. Central, upper and lower scenarios were also modelled to provide an understanding of the uncertainty bands around the prediction levels depending on future events.

Properties are classified using the government SIC codes, which are mapping up to one of five industry groups: Agriculture, non-service, service (economy-driven), service (population-driven) and unclassified. Each of these industry groups has an MLR model, which is then calibrated to base-year consumption from the latest annual review (AR) data.

This piece of work takes the existing models and applies them to updated data from several data sources, to provide an understanding of how the changing socio-economic climate in the UK has impacted NHH consumption forecasts. Updated factors were initially intended to be GVA (economic), employment, consumption (AR) and population datasets, although GVA and population weren't used in the final forecast update. The forecasted consumption is then adjusted to the upper and lower scenarios, by the same factors as in the previous forecasts, rather than re-modelling scenarios.

2 Data Updates

2.1 Oxford Economics

The models used Gross Value Added (GVA) and employment forecasts from Oxford Economics (OE) through to 2040. A data update was purchased from Oxford Economics, with the intention of using this to update non-household property forecasts with the latest available data by SIC group, this is particularly important to understand how the economic landscape of the UK is forecasted to change post-Brexit, post-covid and with additional factors such as the war in Ukraine.

Upon our initial QA checks upon receipt of this data, it was clear that there were significant changes to the economic forecast, not only looking ahead, but also historically. This causes significant issues as the models are trained on specific data, and this data needs to be consistent for an updated forecast to be calculated. Figure 1 shows the differences in GVA forecast after the data updates. There are historical changes in all SIC groups, but most notably in groups C and E.

Figure 1 - Updated vs Original Oxford Economics forecast by SIC group

After further analysis and discussion with OE, the following conclusions were reached as to what could have changed historical data:

1. The original dataset was set to 2016 £s, whereas the new data uses 2019 £s, meaning inflation will have adjusted the numbers.

2. A larger effect will be coming from the fact that the ONS made methodological changes to their GVA calculations in 2021 (back casted for all previous years), which OE then incorporated into their models.

Due to point 2, and after performing some sensitivity analysis on the effects of these changes, it was decided that it wouldn't be possible to use updated economic data for this forecast update.

There were no issues noted with the OE employment data.

2.2 Consumption Data

Additional AR consumption data was provided for AR21 and AR22 at WRZ-level, this is used to create calibration factors for the model consumption.

2.3 Population data

WRZ-level population data is used as an explanatory variable in the models, but due to lack of data availability, existing population data was used as before.

3 Results

The updated datasets and forecasts show the sharp downwards spike in consumption due to the Covid-19 pandemic and resulting step-change in the base year consumption. Under almost all aggregations, there is considerable downwards movement in the position of the forecasts.

The outputs are provided in three different aggregations: company, industry group and WRZ. Tabular and visual outputs for each of these are shown below. Within these aggregations, the most notable (relative) changes within a group are as follows:

- WRZ-level
  - Kent Thanet has a 19% drop in forecast consumption.
  - Hampshire Rural has an 16% drop in forecast consumption.
  - Hampshire Winchester has a 16% drop in forecast consumption.
- Industry-level
  - Service-population has an 14% drop in forecast consumption.
  - Non-service has a 13% drop in forecast consumption.
  - Service-economy has an 13% drop in forecast consumption.
- SWS company has an 13% drop in forecast consumption.

No groupings show an increase in forecast consumption.

4 Outputs

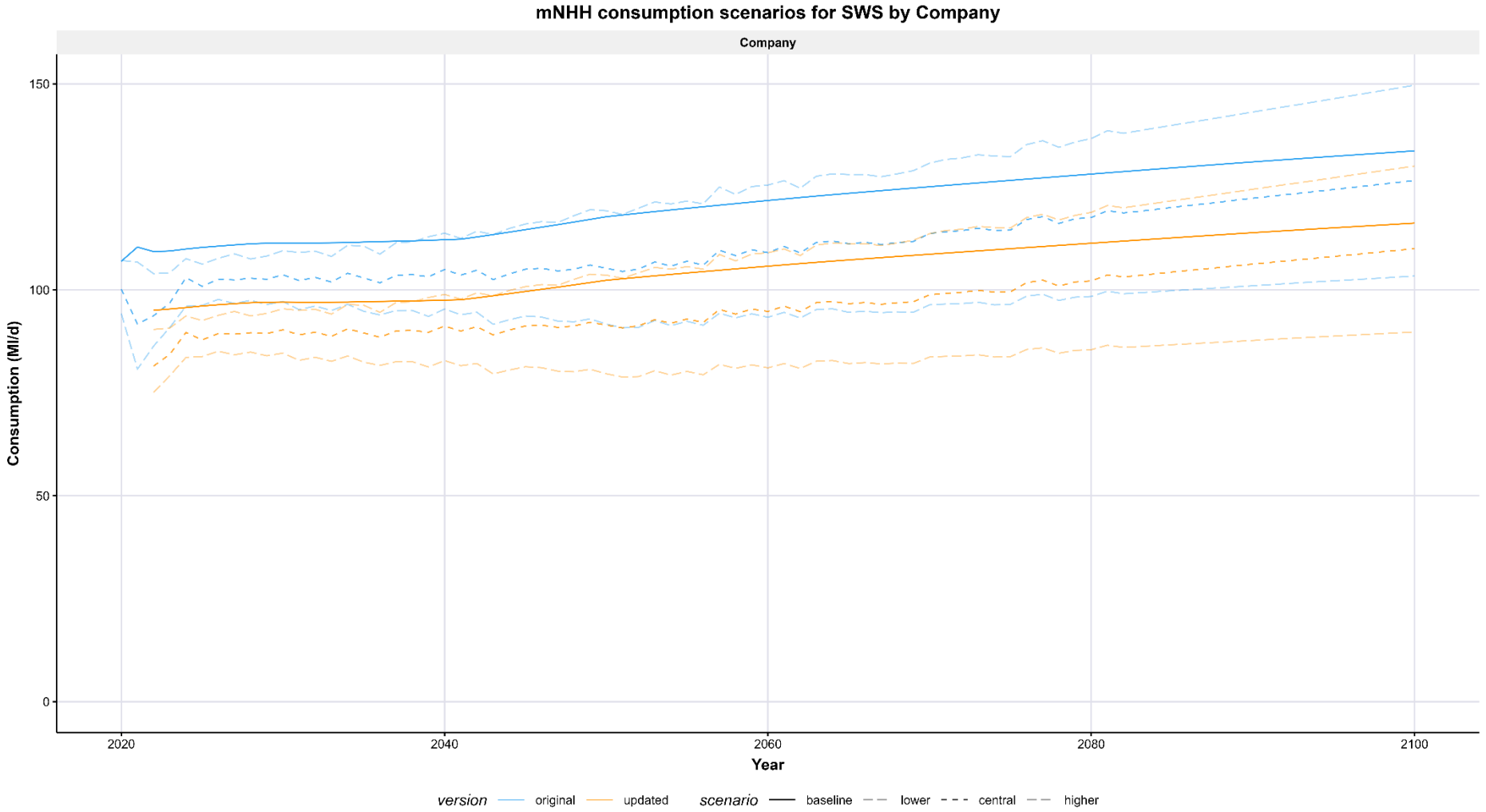
Table 1, ordered by largest reduction in consumption, shows how much each group within each aggregation has changed with the updated forecasts.

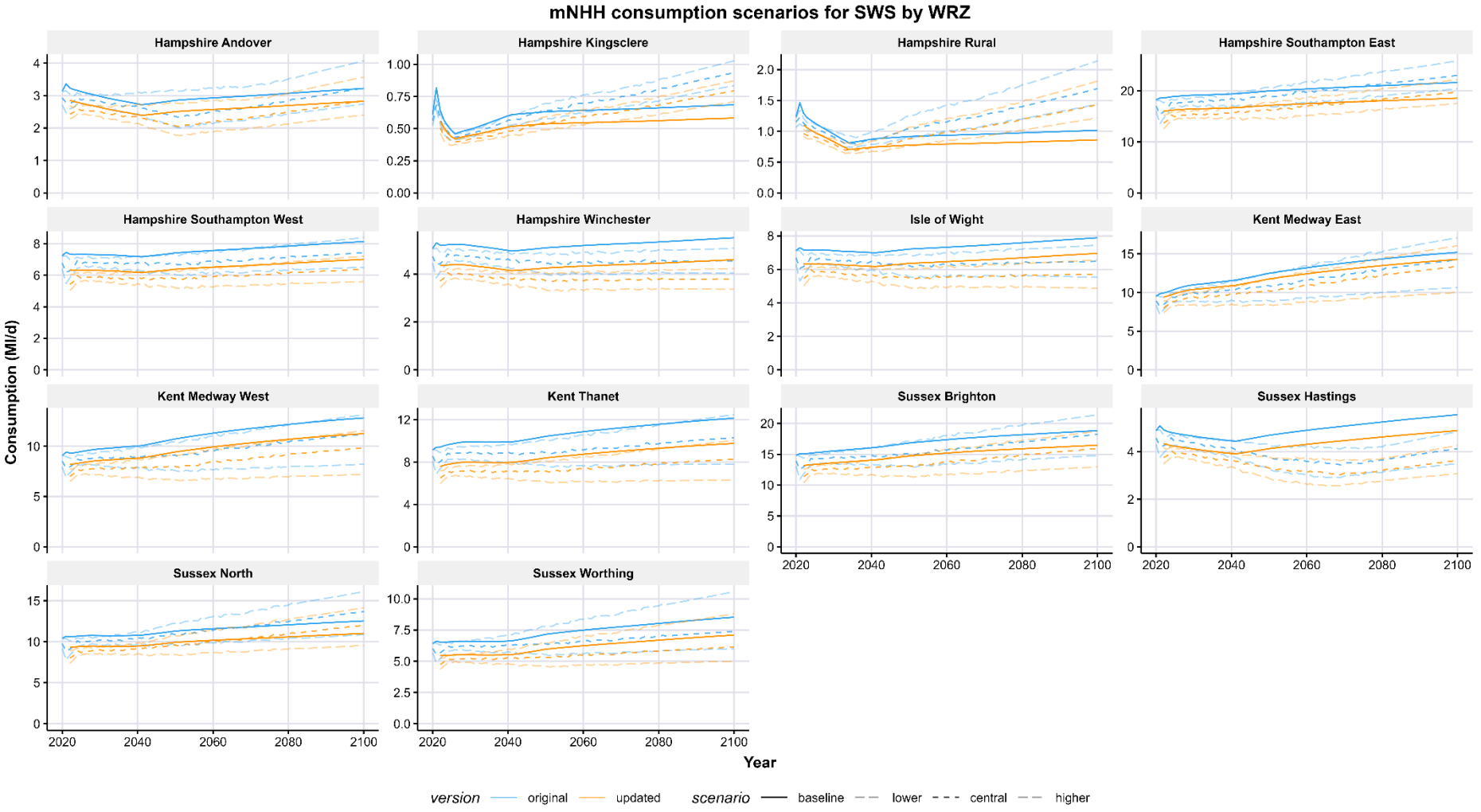
Table 1 - Change in average (across timeseries) consumption forecast by groupings

Meter Status	Area	Sector	Grouping Type	Original (ml/d)	Updated (ml/d)	Raw Change (ml/d)	Change (%)
measured	Kent Thanet	Measured	WRZ	10.8	8.74	-2.08	-19%
measured	Hampshire Rural	Measured	WRZ	0.969	0.817	-0.152	-16%
measured	Hampshire Winchester	Measured	WRZ	5.23	4.37	-0.862	-16%
measured	Sussex Worthing	Measured	WRZ	7.43	6.21	-1.22	-16%
measured	Hampshire Kingsclere	Measured	WRZ	0.625	0.534	-0.0911	-15%
measured	Company	Service industries (population driven)	Industry Grouping	43	37.2	-5.84	-14%
measured	Hampshire Southampton East	Measured	WRZ	20.2	17.4	-2.81	-14%

measured	Hampshire Southampton West	Measured	WRZ	7.6	6.55	-1.05	-14%
measured	Company	Company	Company	121	105	-15.5	-13%
measured	Company	Non-service industries (excl. Agriculture)	Industry Grouping	14.7	12.8	-1.9	-13%
measured	Company	Service industries (economy driven)	Industry Grouping	43.6	37.7	-5.83	-13%
measured	Hampshire Andover	Measured	WRZ	2.99	2.63	-0.367	-12%
measured	Isle of Wight	Measured	WRZ	7.37	6.51	-0.863	-12%
measured	Kent Medway West	Measured	WRZ	11.2	9.85	-1.31	-12%
measured	Sussex Brighton	Measured	WRZ	17.1	15.1	-2.1	-12%
measured	Sussex Hastings	Measured	WRZ	4.96	4.37	-0.593	-12%
measured	Sussex North	Measured	WRZ	11.5	10.1	-1.38	-12%
measured	Company	Agriculture	Industry Grouping	6.17	5.5	-0.676	-11%

measured	Company	Unclassified	Industry Grouping	13.6	12.3	-1.31	-10%
measured	Kent Medway East	Measured	WRZ	12.9	12.3	-0.68	-5%





5 Conclusions and Recommendations

Overall, there is seen to be a significant decline in NHH consumption forecasts. This is primarily driven by the lower base-year consumption caused by the drop during the covid-19 pandemic, alongside changing employment forecasts in the Oxford Economics datasets.

A positive result of this update is the fact that almost all the updated baseline forecasts are contained within the scenario bounds from the original forecast. This shows that the original scenarios sufficiently accounted for potential variations in the forecasts due to unknown changes.

Our primary recommendations would be to carry out a re-modelling process using the latest available GVA data, as the data update wasn't able to be used due to changes in the underlying method of deriving this metric. This means a true understanding of how the economic forecasts have changed since 2020 are limited to just the employment forecasts. Updated population data should also be used.

We also recommend re-modelling and redefining the scenarios that are used, to ensure that they are in-line with the latest socio-economic changes and providing uncertainty bounding around the baseline models with additional context around the economic uncertainty from the data sources.

6 Appendix

6.1 SIC group to industry group mapping

sic_code_group	industry_grouping
A	agriculture
B	nonservice
C	nonservice
D	nonservice
E	nonservice
F	nonservice
G	serviceeconomy
H	serviceeconomy
I	serviceeconomy
J	serviceeconomy
K	serviceeconomy
L	serviceeconomy
M	serviceeconomy
N	serviceeconomy
O	servicepopulation
P	servicepopulation
Q	servicepopulation
R	servicepopulation

S	servicepopulation
T	servicepopulation
U	unclassified