UNC Workgroup 0754R

07/07/2021

Provided by:

Contents

- Workgroup Meeting 2 Recap (12th May 2021)
    - Background / Rationale
    - Key Discussion Points
- Update on Progress:
    - Approach to Analysis
    - Data Requirements, Success Criteria, Measures and System Set Up
- Timeline
- Next Steps

Useful Links

- Uniform Network Code Section H
- Demand Estimation Methodology
- Demand Modelling Approach (2021 version)
- UIG Task Force Findings
- NDM Algorithm Consultation Material
- UNC Request for 0754R Workgroup

Glossary

- For those not familiar with all the industry abbreviations please find full name of those used in this presentation below:
  - ALP: Annual Load Profile
  - AUGE: Allocation of Unidentified Gas Expert
  - CDSP: Central Data Services Provider
  - CWV: Composite Weather Variable
  - DAF: Daily Adjustment Factor
  - DESC: Demand Estimation Sub Committee
  - DM: Daily Metered
  - DOW: Day of Week
  - EUC: End User Category
  - ILF: Indicative Load Factor
  - LDZ: Local Distribution Zone
  - MAPE: Mean Absolute Percentage Error
  - MPE: Mean Percentage Error
  - NDM: Non-Daily Metered
  - PLF: Peak Load Factor
  - SNCWV: Seasonal Normal Composite Weather Variable
  - UIG: Unidentified Gas
  - UNC: Uniform Network Code
  - WCF: Weather Correction Factor

Meeting 2 Re-cap
(12th May 2021)

Background

- UIG Task Force produced a number of recommendations to help reduce temporary UIG levels/volatility. This included findings associated with the modelling error within the NDM Algorithm
- DESC is responsible for the NDM Algorithm (UNC Section H) and has an obligation to review it every 3 years (UNC H 2.2.2)
- Prior to moving forward with the above a consultation was performed during Q4 of 2020 to assess the levels of support for making improvements to the NDM Algorithm
- A more detailed view of the background to this Workgroup and current state overview is provided in the March meeting papers here

Rationale for Workgroup 0754R:

- Supports DESC's UNC obligation to review the NDM Algorithm
- UIG Task Force findings will be explored and progressed
- Clear industry support for investigating advanced analytical approaches
- A Workgroup maintains focus and increases visibility across the industry
- Improved NDM Allocation will result in a reduction in UIG volatility and subsequent Meter Point reconciliation/UIG volumes (temporary)

Key Discussion Points

The main headlines from Meeting 2 of 754R were……

1.Areas to Investigate :
Workgroup confirmed it was happy with the 3 areas of focus (highlighted green in Fig 1), these are:

- Area 1: Trial alternative approaches to deriving SNDt
- Area 2: Improve Validation Processes
- Area 3: Review End User Category definitions

For each of the 3 areas there was general agreement with the high level view of proposed approach, data and systems

2.Resource Requirements:
Systems / Set-Up
- SAS Platform/Products to be used
- Significant work required 'setting up'

Costs
- Existing service provision to be used, may be additional costs if/when need to perform at scale

Resource/Expertise
- In House Demand Estimation Team resource to support the WG analysis, more availability post peak BAU activities

3.Data Availability:
Workgroup reviewed the proposed data item categories (Fig.2) to be used during the analysis, these are:

- NDM Gas Consumption
- Weather
- Supply Point Attributes
- Reconciliation Data
- Calendar

4. Trial LDZs and EUCs:
Workgroup agreed with approach to focus on:

- 2 LDZs, with preference for North/South split which is representative of others (TBC)
- 2 EUCs with preference for Domestic "01BND" plus an I&C EUC (TBC)

Fig. 1

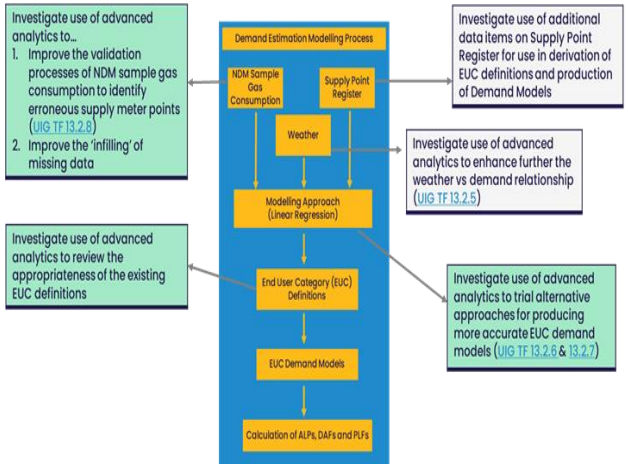
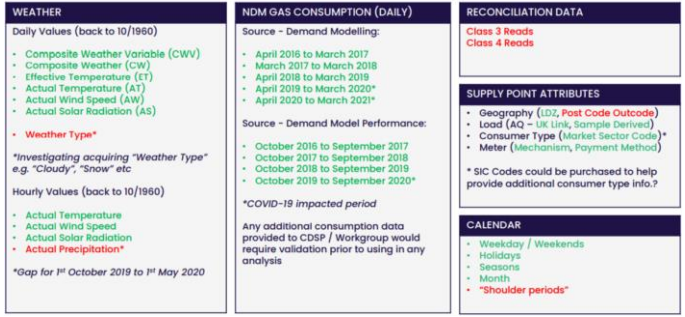


Fig. 2



Update on Progress:
Approach to Analysis

Area 1: Trial alternative approaches to deriving SND_t

Linear Regression between NDM Sample Demand and Daily Composite Weather Variable

Current Approach

EUC Demand Models (39 per LDZ)

Model Smoothing, the process of averaging [3] individual models, is applied to create the final form of the EUC demand model.

$SND_t = P_t * (C1 + C2 * SNCWV_t)$,
$WSENS_t = P_t * C2$

P_t is a multiplicative factor which is set to 1 for a non-holiday Monday to Thursday and potentially a different value for weekend or holiday day.

C_1 is the constant from the (non-holiday) Monday to Thursday regression of daily demand against daily CWV.

C_2 is the slope of this regression line which represents the weather sensitivity of demand for (non-holiday) Mondays to Thursdays.

Annual Load Profiles

The ALP_t shall be determined as:

$$\mathrm{ALP}_t = \frac{SNDE_t}{\frac{[\sum_{t=1}^{N} SNDE_t\ ]}{N}}$$

where

t denotes the value for a particular day

E denotes the EUC

N is the number of days in the Gas Year;

Daily Adjustment Factor

The DAF_t shall be determined as:

$$\mathrm{DAF}_t = \frac{WVCE_t}{SNDE_t}$$

where

E denotes the EUC

t denotes the value for a particular day

SND – Seasonal Normal Demand
WVC - Weather Variable Coefficient;

Peak Load Factor

The PLF for the EUC sample is calculated as:

$$\frac{\text{Aggregate AQ from the EUC model}}{\text{1 in 20 peak demand from the EUC model} * 365}$$

An aggregate AQ is derived from the smoothed EUC demand model by setting the composite weather variable to its seasonal normal level in the model and summing the resulting demand values over the 365 days of the forecast year (excluding any 29th February).

For NDM EUCs a 1 in 20 peak day demand estimate is derived from each gas demand EUC model by simulation using the smoothed EUC demand model in conjunction with the database of historic daily composite weather variable values for the appropriate LDZ.

Area 1: Trial alternative approaches to deriving SND_t

Objective:

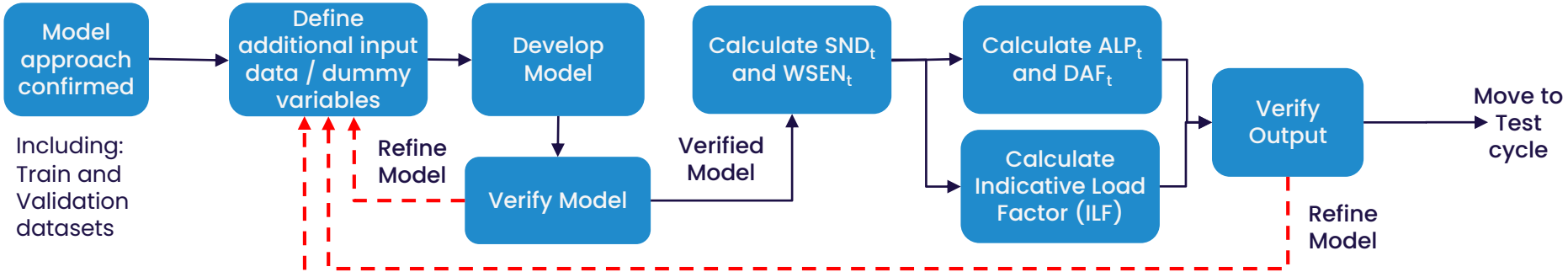
Explore alternative modelling approaches (outside of linear regression) to identify whether a more accurate view of SND_t and subsequent ALPs, DAFs and PLFs exist

Identify any weaknesses, improvements and make recommendations which link to evidence of a reduction in NDM modelling error

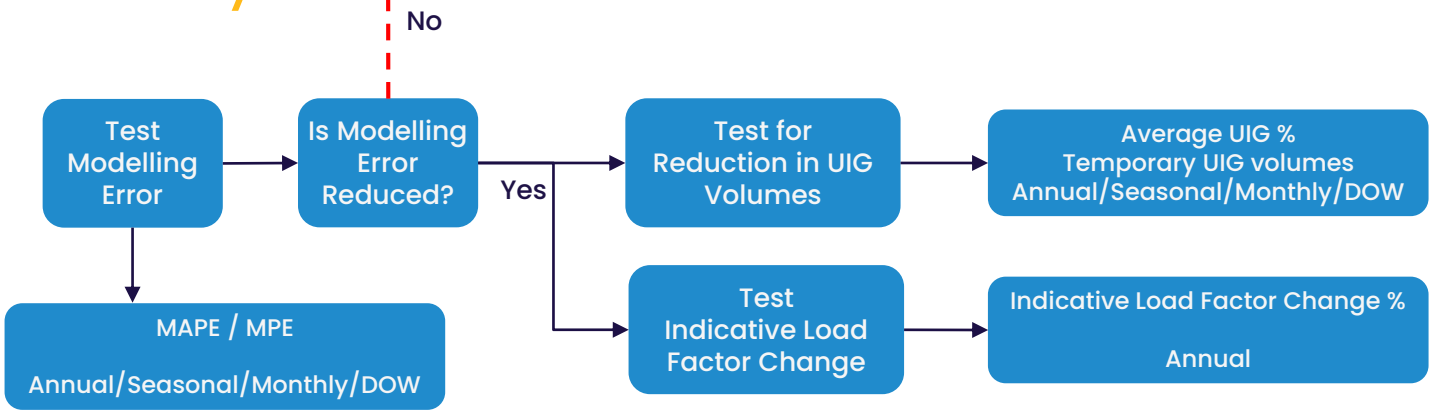
Approach, Data and Systems:

Potential Approach	Data	Systems
Time Forecasting (e.g. ARIMA) Neural Network UIG TF 13.2.6 and 13.2.7	Daily Gas Consumption Daily Weather Supply Point Attributes (AQ, MSC) Dummy Variables (e.g. month, season)	SAS Enterprise Miner
Amendments to existing approach e.g. dummy variables for month, Individual day of the week		SAS Demand Estimation Modelling

Development Cycle



Test Cycle



Evidence

For each model collate:

Approach details,
Modelling Error,
UIG outcomes,
ILF changes
Process observations

Area 1: Trial EUC Selection

It is proposed the following **EUC Models** will be focussed on:

- Domestic model **"01BND"** represents nearly 90% of NDM supply points
- I&C model **"02BNI"** represents the second largest I&C consumer group within the Small NDM AQ range (0 to 2,196 MWh pa)
- I&C model **"05B"** represents the largest I&C consumer group within the Large NDM AQ range (2,196 to 58,600 MWh pa)

Small NDM

EUC Band	AQ Range	No. of Models
01	0 to 73.2 MWh pa	4
02	73.2 to 293 MWh pa	4
03	293 to 732 MWh pa	5
04	732 to 2,196 MWh pa	5

Large NDM

EUC Band	AQ Range	No. of Models
05	2,196 to 5,860 MWh pa	5
06	5,860 to 14,650 MWh pa	5
07	14,650 to 29,300 MWh pa	5
08	29,300 to 58,600 MWh pa	5
09	> 58,600	1

Small NDM EUC Selection

EUC Band	EUC Type	No. of Models	Supply Point Count (% of Small NDM)	AQ (% of Small NDM)
"01BND"	Domestic Non-PPM	1	88%	75%
"02BNI"	I&C Non-PPM	1	0.6%	5%

Large NDM EUC Selection

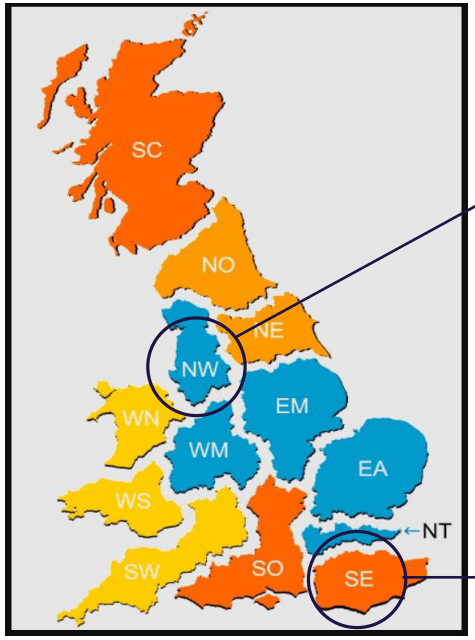
EUC Band 01	EUC Type	No. of Models	Supply Point Count (% of Large NDM)	AQ (% of Large NDM)
"05B"	I&C Non-PPM	1	65%	29%

Area 1: Trial LDZ Selection

It is proposed the following LDZs will be focussed on:

- North West (NW) represents a Northern LDZ and South East (SE) represents a Southern LDZ which is in line with the Workgroup's preference

The proposed EUC selections for both LDZs compare well as a representation of the total population – see tables below:



North West LDZ

EUC Band	EUC Type	Supply Point Count LDZ (Pop'n)	AQ LDZ (Pop'n)
"01BND"	Domestic Non-PPM	87% (88%)	75% (75%)
"02BNI"	I&C Non-PPM	0.6% (0.6%)	5% (5%)
"05B"	I&C Non-PPM	64% (65%)	26% (29%)

South East LDZ

EUC Band	EUC Type	Supply Point Count LDZ (Pop'n)	AQ LDZ (Pop'n)
"01BND"	Domestic Non-PPM	88% (88%)	77% (75%)
"02BNI"	I&C Non-PPM	0.6% (0.6%)	5% (5%)
"05B"	I&C Non-PPM	71% (65%)	36% (29%)

Area 1: Data Availability for Trial EUCs and LDZs

- It is proposed that the daily gas consumption data used in the existing EUC Demand Modelling is utilised in the exploratory analysis. This data has the benefit of being 'clean' (having been previously validated) and effectively available 'off the shelf'

- The tables below display the number of validated supply points available from the past 6 analysis periods (typically April to March). There appears to be sufficient data here to use as an input to trialling different approaches ?

North West LDZ

EUC Band	EUC Type	Modelling 2016 2015/16	Modelling 2017 2016/17	Modelling 2018 2017/18	Modelling 2019 2018/19	Modelling 2020 2019/20	Modelling 2021 2020/21	Total
"01BND"	Domestic Non-PPM	209	185	170	199	327	443	1533
"02BNI"	I&C Non-PPM	159	126	471	273	452	334	1815
"05B"	I&C Non-PPM	152	139	126	118	122	117	774

South East LDZ

EUC Band	EUC Type	Modelling 2016 2015/16	Modelling 2017 2016/17	Modelling 2018 2017/18	Modelling 2019 2018/19	Modelling 2020 2019/20	Modelling 2021 2020/21	Total
"01BND"	Domestic Non-PPM	207	196	177	223	323	427	1553
"02BNI"	I&C Non-PPM	162	164	448	300	465	539	2078
"05B"	I&C Non-PPM	139	145	157	145	162	121	869

Update on Progress:

Data Requirements

Success Criteria and Measures

System Set-Up

Areas 1 to 3: Data Availability - No Update*

WEATHER

Daily Values (back to 10/1960)

- Composite Weather Variable (CWV)
- Composite Weather (CW)
- Effective Temperature (ET)
- Actual Temperature (AT)
- Actual Wind Speed (AW)
- Actual Solar Radiation (AS)

- Weather Type*

**Investigating acquiring "Weather Type" e.g. "Cloudy", "Snow" etc*

Hourly Values (back to 10/1960)

- Actual Temperature
- Actual Wind Speed
- Actual Solar Radiation
- Actual Precipitation*

**Gap for 1st October 2019 to 1st May 2020*

NDM GAS CONSUMPTION (DAILY)

Source - Demand Modelling:

- April 2016 to March 2017
- March 2017 to March 2018
- April 2018 to March 2019
- April 2019 to March 2020*
- April 2020 to March 2021*

Source - Demand Model Performance:

- October 2016 to September 2017
- October 2017 to September 2018
- October 2018 to September 2019
- October 2019 to September 2020*

**COVID-19 impacted period*

Any additional consumption data provided to CDSP / Workgroup would require validation prior to using in any analysis

RECONCILIATION DATA

Class 3 Reads
Class 4 Reads

SUPPLY POINT ATTRIBUTES

- Geography (LDZ, Post Code Outcode)
- Load (AQ – UK Link, Sample Derived)
- Consumer Type (Market Sector Code)*
- Meter (Mechanism, Payment Method)

* SIC Codes could be purchased to help provide additional consumer type info.?

CALENDAR

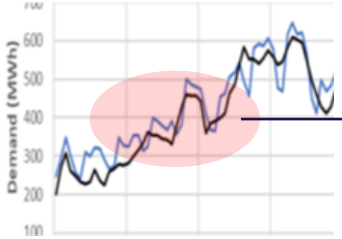
- Weekday / Weekends
- Holidays
- Seasons
- Month
- "Shoulder periods"

Notes:
The above represents our current view of data required, this may be added to as we learn more about the approaches to be applied
Green Items available and minor prep-work needed
Red items requires work to retrieve and/or prepare ready for use

Area 1: Success Criteria and Measures

- Area 1 investigation will be measured against the following Success Criteria:
  1) Reduce Demand Modelling Error, particularly during higher volume period (October to March)
  2) Reduction in Unidentified Gas volumes ('Temporary')
  3) Minimise impacts to simulated peak demand, thereby maintaining Peak Load Factors and SOQ levels

1.Demand Modelling Error



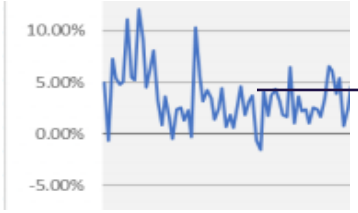
Demand Modelling Error will always exist and is caused primarily by 2 factors:

i) Imperfections in the Demand Model
ii) AQ Inaccuracy

AQ Inaccuracy (i.e. out of date) can be mitigated via regular meter read submissions

Measures	Benefit	Assessment Period
Mean Absolute Percentage Error (MAPE)	Provides a view of overall ACCURACY of the Model for validated period	Annual, Season, Month, Weekday
Mean Percentage Error (MPE)	Provides a view of BIAS within the Model for validated period	

2.Temporary UIG Volumes



Demand Modelling Error is one of the contributory factors of Temporary UIG

Improvements which reduce Modelling Error should also be seen in a reduction in UIG volumes (+ or -)

Measures	Benefit	Assessment Period
Average UIG (%)	Provides a view of the change in BASE level UIG volumes due to reduced Demand Modelling Error	Annual, Season, Month, Weekday
Temporary UIG Volume Change		
Day to Day Change in UIG	Provides view of change in VOLATILITY in UIG	

3. Peak Load Factors

Demand Models and their relevant characteristics (e.g. Weather Sensitivity) are a key input into the process which determines Peak Day Demand (SOQs)

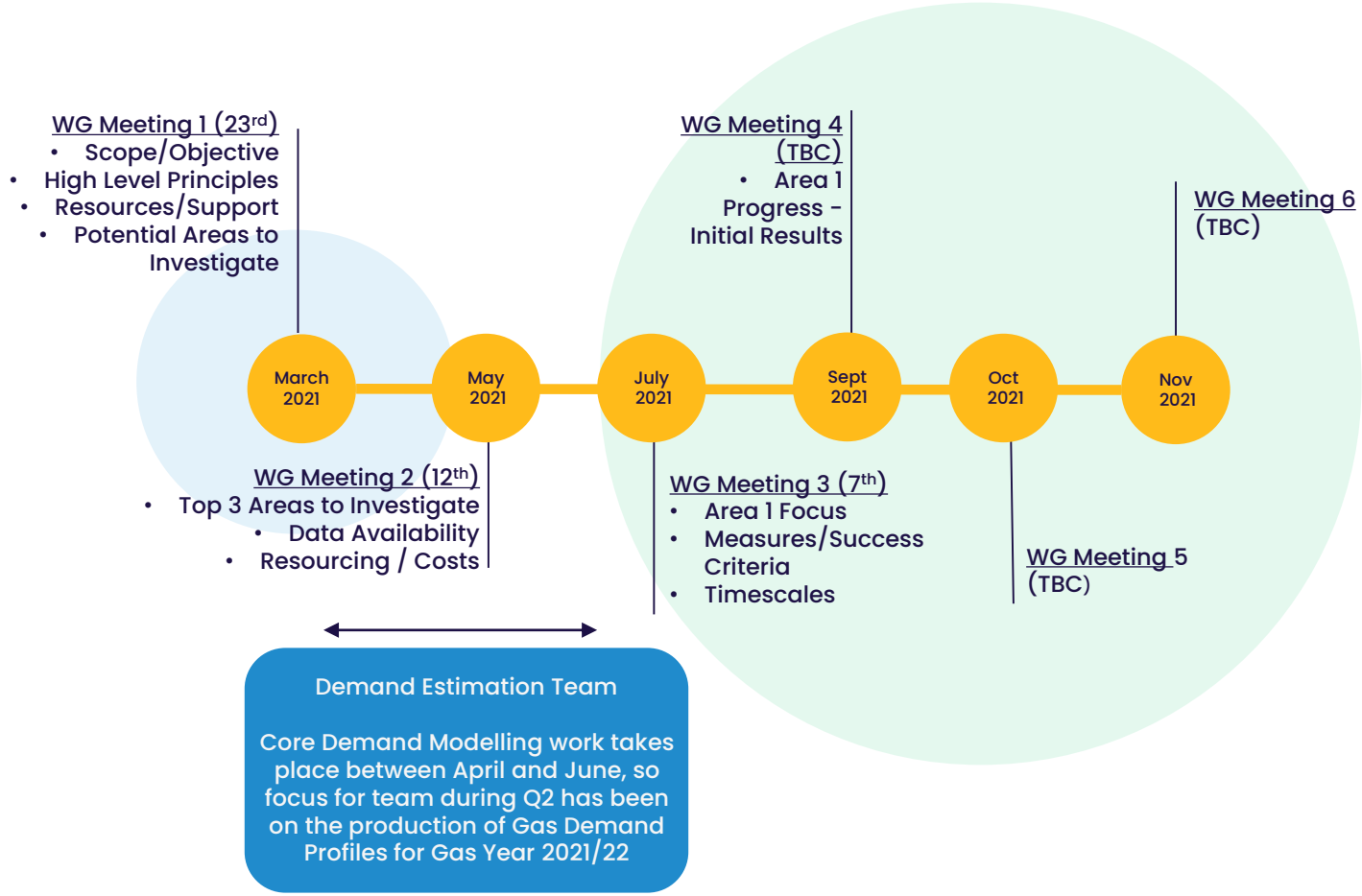
Measures	Benefit	Assessment Period
Indicative Load Factor (ILF) Change	The ILF provides an estimated view of overall weather sensitivity for the Demand Model	Annual

Areas 1 to 3: Resources – No Update*

- Systems / Set-up:
  - We plan to utilise access to existing SAS Platform/Products in the form of:
    - Enterprise Miner for performing advanced analytics such as clustering, ARIMA
    - Existing Demand Estimation modelling system for assessing changes within current approach and replication of demand attribution calculations (needed to review impacts to temporary UIG)
  - We believe there is significant work required in 'setting up' before proceeding with analysis – i.e. documenting approach/measures, collating relevant data, infrastructure/tooling set-up
- Costs:
  - Current understanding from our SAS 'system manager' is that our existing service provision will support the analysis we wish to perform but this may need to be reviewed as we begin the analysis, particularly as and when we need to perform at scale
- Resources / Expertise:
  - Current plan is to use 'In House' Demand Estimation Team resource to support the Workgroup analysis, which is likely to increase as peak BAU activities complete in mid-July
  - Certain aspects will be new to the Demand Estimation team and so support/advice from Workgroup members welcome (which we have already drawn upon)
  - UIG Task Force Findings will also provide important reference material

Timeline

Workgroup 0754R Timeline



Next Steps

Next Steps

Develop Test and Evidence collection processes

Area 1: Document approach(s)

Include any requirements from meeting 3 and task group findings

Area 1: Develop model(s) Refine models (as needed)

Evaluate test and evidence

Meeting 4 preparation

Thank you