

UNC Workgroup 0754R

12/05/2021

The logo for Xserve, featuring a stylized 'X' composed of two overlapping blue shapes, followed by the word 'serve' in a light blue sans-serif font.

Provided by:

The logo for Correla, consisting of two overlapping circles, one blue and one yellow, followed by the word 'correla' in a dark blue sans-serif font.

Contents

- Workgroup Meeting 1 Recap (23rd March 2021)
 - Background / Rationale
 - Discussion Points / Actions
- Action 0302
 - Review of Areas for Investigation and resources required
- 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:
 - LDZ: Local Distribution Zone
 - UIG: Unidentified Gas
 - DM: Daily Metered
 - NDM: Non-Daily Metered
 - ALP: Annual Load Profile
 - DAF: Daily Adjustment Factor
 - PLF: Peak Load Factor
 - EUC: End User Category
 - CWV: Composite Weather Variable
 - SNCWV: Seasonal Normal Composite Weather Variable
 - WCF: Weather Correction Factor
 - DESC: Demand Estimation Sub Committee
 - CDSP: Central Data Services Provider
 - AUG: Allocation of Unidentified Gas Expert
 - UNC: Uniform Network Code

Meeting 1 Re-cap (23rd March 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 1 of 754R were.....

Scope / Objective:

- Initial proposal and scope of Workgroup was *“To investigate use of advanced analytics to improve the NDM Demand Modelling”*
- Workgroup decided to extend this scope to allow other ideas to be considered – revised objective *“Investigate improvements to NDM Demand Modelling”*

Resources / Support:

- Wider industry resources not readily available to perform analysis but can provide advice and data if needed
- Preference was for CDSP (Correla on behalf of Xoserve) to be performing analysis. Results more likely to be accepted if carried out by impartial industry participant

Approach:

- The Workgroup should proceed with analysis as an “academic exercise” and not be distracted at this stage with potential implementation issues, unless identified as obvious ‘non-starters’ by the Workgroup
- More focus to be directed on the Winter period when volumes are at their highest

Actions:

- **Action 0301:**
RH and MP to seek input from National Grid NTS and/or Xoserve representatives involved with Gemini, regarding impacts of changes to Gemini on this Request Workgroup 0754
- **Action 0302:**
Xoserve (MP) to identify a) what the cost would be for them to carry out an initial benefit analysis in relation to SND_t, and b) what resources they would need in terms of data, time or funding to investigate the use of advanced analytics to trial alternative approaches for producing more accurate EUC demand models
- **Action 0303:**
Workgroup to review the UIG taskforce recommendations to review before the next meeting

Action 0302:
Areas to Investigate –
Resources Required

Potential areas for Workgroup to Investigate

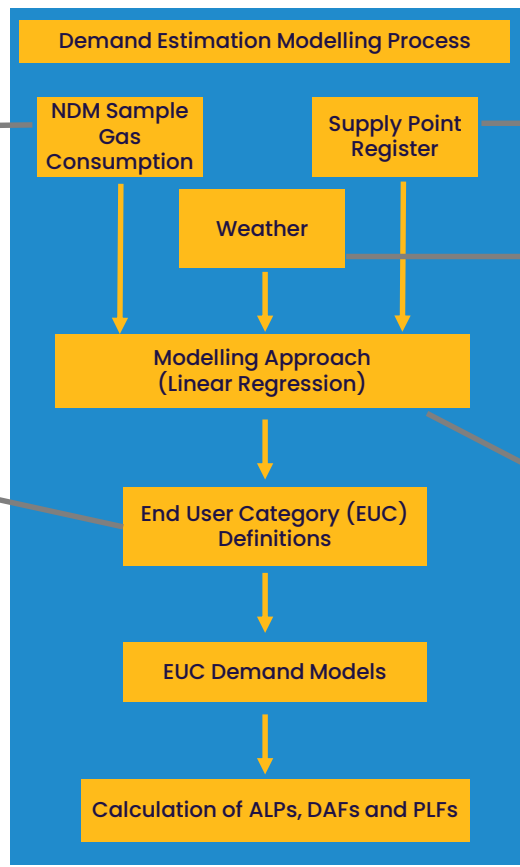
Investigate use of advanced analytics to...

1. Improve the validation processes of NDM sample gas consumption to identify erroneous supply meter points ([UIG TF 13.2.8](#))
2. Improve the 'infilling' of missing data

Investigate use of advanced analytics to review the appropriateness of the existing EUC definitions

UIG TF Reference links on this slide relate to the UIG Task Force Findings which should provide useful reference material for this Workgroup

Health Warning:
It should be noted that the % improvements in UIG quoted in the material will have been based on a more comprehensive adoption of M/L techniques as opposed to the industry's 'compromise' position of retaining ALPs & DAFs (i.e. "option 2")



Investigate use of additional data items on Supply Point Register for use in derivation of EUC definitions and production of Demand Models

Investigate use of advanced analytics to enhance further the weather vs demand relationship ([UIG TF 13.2.5](#))

Investigate use of advanced analytics to trial alternative approaches for producing more accurate EUC demand models ([UIG TF 13.2.6](#) & [13.2.7](#))

CDSP Approach

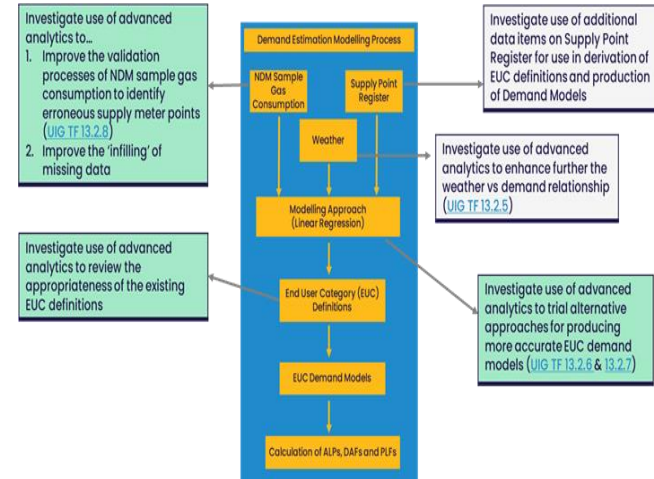
What we plan to do:

- Our main objective shall be in line with the primary intent of the Workgroup proposal which is to investigate how advanced analytics can be used to improve NDM Demand Modelling
- We intend to focus on the 3 areas to investigate highlighted green (Fig.1) and will address them in this order
 - Area 1: Trial alternative approaches to deriving SND_t
 - Area 2: Improve Validation Processes
 - Area 3: Review End User Category definitions
- During the investigations we shall look to identify any improvement ideas not related to advanced analytics and those generated by Workgroup suggestions and feedback

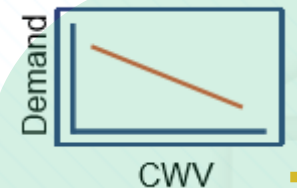
How:

- For each area, where appropriate we shall utilise advanced analytic investigations focusing on 1 or 2 LDZs and report back to the Workgroup before deciding on wider coverage (i.e. more LDZs)
- Success criteria will be primarily based on 1) reducing modelling error 2) subsequent reduction in UIG (temporary) and 3) minimal impacts to simulated peak demand – statistical measures to be confirmed with Workgroup

Fig. 1



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.

$$(SND_t = P_t * (C_1 + C_2 * SNCWV_t), WSENS_t = P_t * C_2)$$

P_t is a multiplicative factor which is set equal to 1 for a non-holiday Monday to Thursday and potentially a different value on a weekend (Friday, Saturday and Sunday) 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:

$$ALP_t = \frac{SNDE_t}{\frac{\sum_{t=1}^N SNDE_t}{N}}$$

where

ALP_t is the ALP on day t ;

$SNDE_t$ is Seasonal Normal Demand for the EUC for day t ;

N is the number of days in the Gas Year;

Daily Adjustment Factor

The DAF_t shall be determined as:

$$DAF_t = \frac{WVCE_t}{SNDE_t}$$

where

DAF_t is the DAF on day t ;

$WVCE_t$ is the value of the Weather Variable Coefficient in the Demand Model for the EUC (i.e. the sensitivity to weather);

$SNDE_t$ is the value of the seasonal normal demand for the EUC.

Peak Load Factor

The Peak Load Factor (PLF) for the EUC is calculated as:

Aggregate AQ from the EUC model

1 in 20 peak demand from the EUC model * 365

where

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 February 29th).

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

Area 2: Improve Validation Processes

- Prior to performing any demand modelling its clearly important to quality check the data inputs.
- The current checks aim to strike the balance of ensuring erroneous data is removed but also ensuring sufficient data is available for modelling. The current validation rules are shown in Fig.2.

- Where sample data streams are generally complete but are missing a small amount of data an 'infilling' is applied e.g. weather correction of D-7 consumption.
- The latest validation results, recently presented at DESC, are shown in Fig.3

Fig.2

Appendix 2 – Daily Gas Consumption Data Validation

The following provides the proposed validation criteria for use against the Daily Gas Consumption Data in the 2021 Gas Demand EUC Modelling. Section 1 of the NDM Algorithms Booklet will contain further details of the validation process and outcomes

Small NDM: 0 to 2,196 MWh p.a.

Source	EUC Bands	Missing Days		Consecutive Zeros		Spike Ratios	
		Summer	Winter	Summer	Winter	Summer	Winter
Xoserve Managed sample (and any third party data)	01 and 02	15 or more	15 or more	N/A	33 or more	15:01	08:01
Network Managed sample (and any third party data)	02, 03	28 or more	28 or more	N/A	20 or more	13:01	05:01

Large NDM: >2,196 MWh p.a.

Source	EUC Bands	Missing Days		Consecutive Zeros		Spike Ratios	
		Annual	Winter	Annual	Winter	Annual	Winter
Network Managed sample (and any third party data)	05, 06, 07 and 08	40 or more	20 or more	N/A	20 or more	08:01	N/A

Where:

Summer period is defined as 1st April 2020 to 30th September 2020.

Winter period is defined as 1st October 2020 to 31st March 2021.

Annual period is defined as 1st April 2020 to 31st March 2021

Fig.3

TABLE A.1 - VALIDATION SUMMARY OF DAILY GAS CONSUMPTION DATA

Analysis Period	01/04/20 to 31/03/21
START: MPRs with at least 300 'Daily Gas Consumption Data' records present within the analysis period	42,961
REMOVED: MPRs not on SAP-ISU	58
REMOVED: Exceeds Missing Read Tolerance	7,988
REMOVED: Exceeds Consecutive Zero Consumption Tolerance	2,496
REMOVED: Exceeds Maximum vs Average Consumption Tolerance (Spikes)	1,754
REMOVED: AQ Ratio Failures	3,510
REMOVED: Winter vs Summer Consumption Ratio Tolerance (Band 09B sites)	29
REMOVED: Winter Annual Ratio (WAR) less than 20%	111
REMOVED: Other (i.e. Scottish Independent LDZs; AQ Too Small; PrePayment above Band 02)	4,625
REMOVED: TOTAL	20,571
PA SSED VALIDATION (Pre Stratification)	22,390
REMOVED: MPRs not required following Stratification	1,970
PA SSED VALIDATION (Post Stratification)	20,420

Area 2: Improve Validation Processes

Objective:

Explore use of advanced analytic techniques to improve validation routines prior to modelling

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
<p>Investigate latest data cleansing techniques/scripts</p> <p>'Uncertainty Estimation' UIG TF 13.2.8</p> <p>Compare current post validation results to revised methods - both 'infill' approach and ability to identify suspicious demand patterns</p>	<p>Daily Gas Consumption</p> <p>Daily Weather</p> <p>Supply Point Attributes (AQ, MSC, Correction Factor)</p>	<p>SAS Demand Estimation Modelling</p>

Area 3: Review End User Category (EUC) definitions

Fig.4

EUC Band	AQ Range	No. of Models	Supply Point Count (% of Total)	AQ (% of Total NDM)	Winter T/Put (%)*	Summer T/Put (%)*
01	0 to 73.2 MWh pa	4	24,441,667 (99%)	74	74.67	64.85
02	73.2 to 293 MWh pa	4	194,406 (0.8%)	5.8	6.2	6.79
03	293 to 732 MWh pa	5	41,148 (0.17%)	4.16	4.51	4.9
04	732 to 2196 MWh pa	5	17,381 (0.07%)	4.63	4.92	5.68
05	2196 to 5860 MWh pa	5	4,147 (0.02%)	3.14	3.2	4.56
06	5860 to 14650 MWh pa	5	1,383 (0.01%)	2.79	2.53	4.51
07	14650 to 29300 MWh pa	5	540 (0%)	2.47	1.99	4.07
08	29300 to 58600 MWh pa	5	284 (0%)	2.57	1.82	3.98
09	> 58,600	1	26 (0%)	0.43	0.17	0.66

All EUC Observations:

- Band 1 represents largest EUC component (see Fig.5 for breakdown)
- In higher EUC Bands, e.g. Bands 07 and 08 a large no. of models are produced for small number of sites. Due to lack of sampling numbers, aggregations are often necessary across LDZs. These sites are more likely to require their own unique profile – candidates for Class 2?
- Nearly 2 TWh currently resides in Band 9 (as at April 2021) which should be Class 1 and will not be receiving an accurate allocation with the error contributing to UIG

* Winter (Oct-Mar) and Summer (Apr-Sep) Throughput relates to Gas Year 2019/20. This will vary from year to year depending on weather experienced

Band 1 Observations:

- One Demand Model (“01BND”) represents 2/3 of Winter Throughput and so even minor improvements could provide sig. benefits?
- The next largest EUC “01BPD” is (i) currently being allocated using a historic demand profile (2012/13) and (ii) uses unreliable data items for EUC assignment

Fig.5

EUC Band 01	EUC Type	No. of Models	Supply Point Count (% of Total)	AQ (% of Total NDM)	Winter T/Put (%)*	Summer T/Put (%)*
01BND	Domestic Non-PPM	1	21,729,297 (88%)	66.7	67.17	57.38
01BNI	I&C Non-PPM	1	554,068 (2.2%)	2.6	2.79	2.63
01BPD	Domestic PPM	1	2,154,826 (8.72%)	4.7	4.70	4.83
01BPI	I&C PPM	1	3,476 (0.01%)	0	0.01	0.01

Area 3: Review EUC Definitions

Objective:

Review the current line up of 39 End User Categories (EUCs) per LDZ and explore whether a more appropriate set of definitions and models 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
Cluster Analysis to identify 'more suitable' profiles	Daily Gas Consumption Daily Weather Supply Point Attributes (AQ, MSC)	SAS Enterprise Miner for exploration SAS Demand Estimation Modelling
Review reconciliation data for EUCs as a method for assessing error ? e.g. how does reconciliation % compare across EUC Bands	Reconciliation Data AUGE Reports	SAS
Re-run Demand Attribution assuming Higher EUC Bands (e.g. 7,8) were Daily Metered	Daily Gas Consumption Daily Weather UK Link AQ Energy - LDZ, DM & Shrinkage	SAS Demand Estimation Modelling

Areas 1 to 3: Data Availability

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:

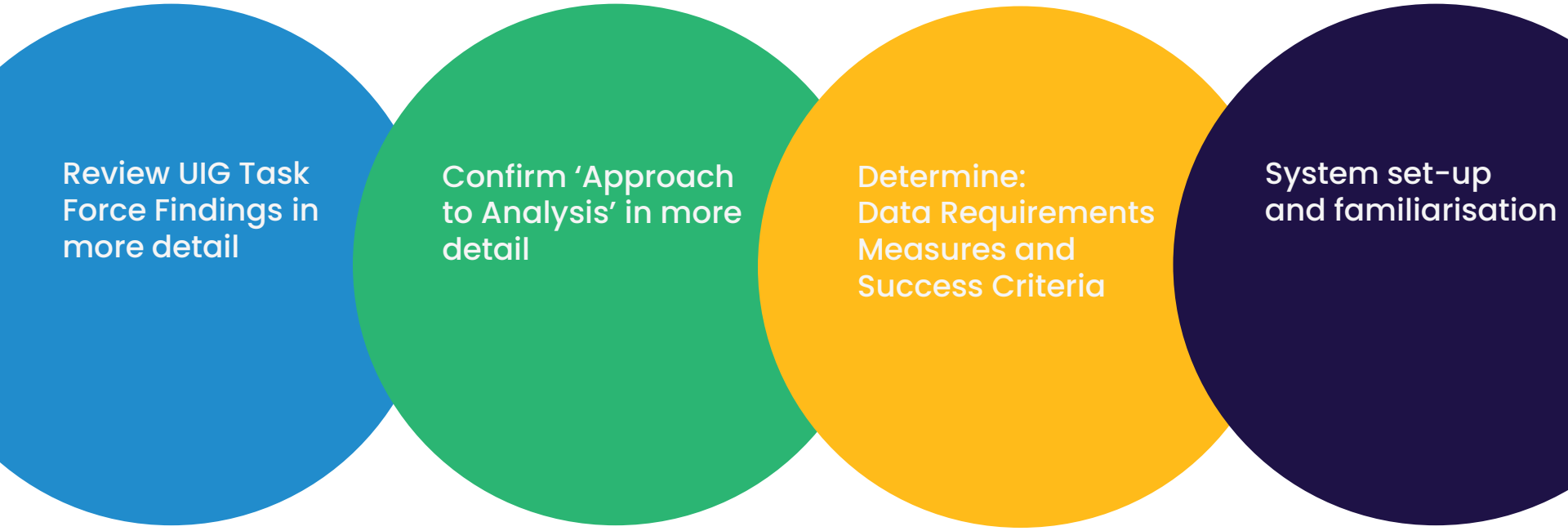
The above represents our current view of data required, this may be added to as we learn more about the approaches to be applied
Green Items available and minor prep-work needed
Red items requires work to retrieve and/or prepare ready for use

Areas 1 to 3: Resources

- **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

Next Steps

Next Steps – “Area 1”



Review UIG Task Force Findings in more detail

Confirm ‘Approach to Analysis’ in more detail

Determine: Data Requirements Measures and Success Criteria

System set-up and familiarisation

Thank you

