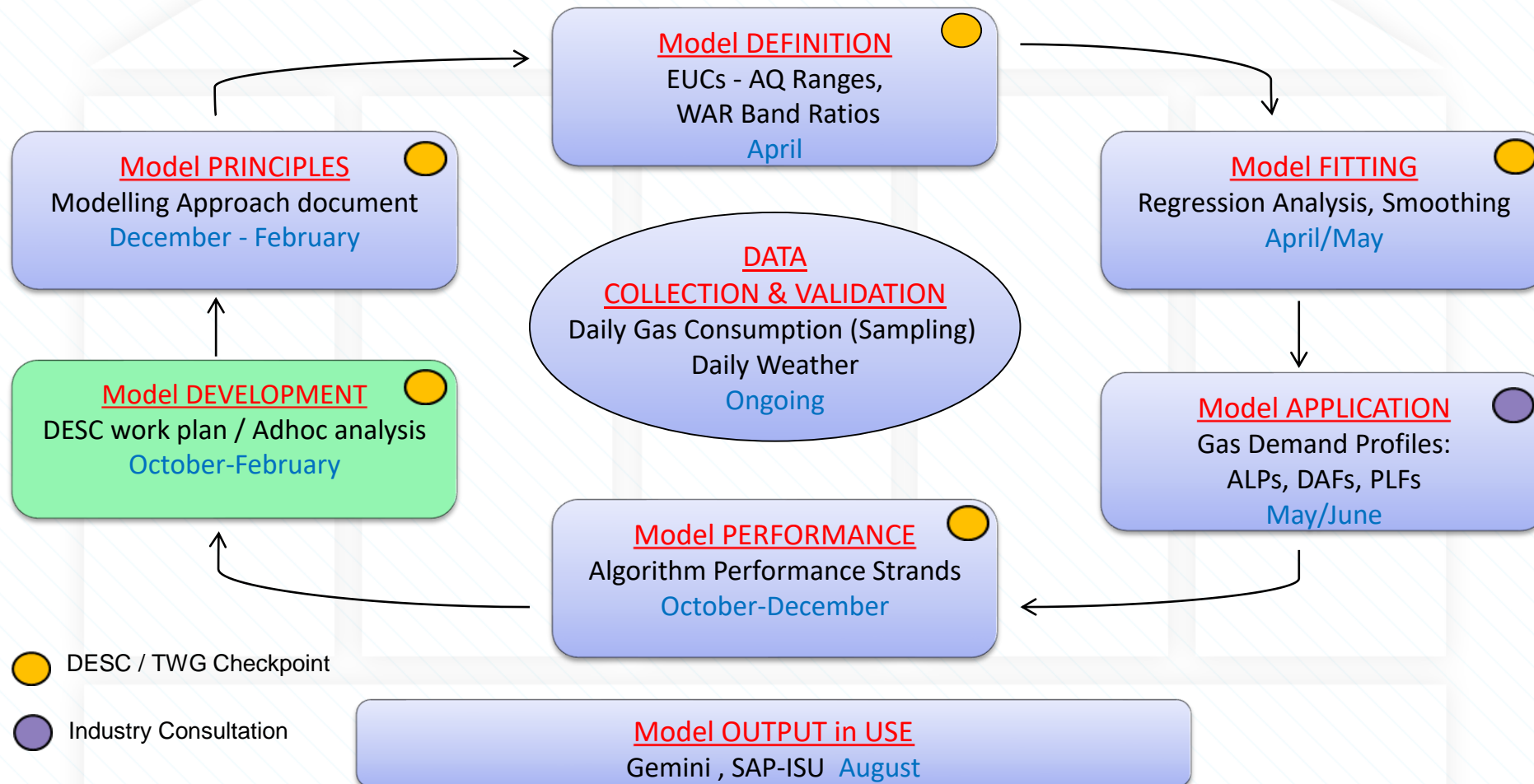# Demand Estimation Sub Committee

Updates on Ad-hoc Work Plan Items

24th February 2021

Overview – Demand Estimation

- An overview of the Demand Estimation process and output can be found here
- This presentation relates to the "Model Development" phase of the Demand Model cycle

Model DEFINITION
EUCs - AQ Ranges,
WAR Band Ratios
April

Model PRINCIPLES
Modelling Approach document
December - February

DATA
COLLECTION & VALIDATION
Daily Gas Consumption (Sampling)
Daily Weather
Ongoing

Model FITTING
Regression Analysis, Smoothing
April/May

Model DEVELOPMENT
DESC work plan / Adhoc analysis
October-February

Model APPLICATION
Gas Demand Profiles:
ALPs, DAFs, PLFs
May/June

Model PERFORMANCE
Algorithm Performance Strands
October-December

DESC / TWG Checkpoint

Industry Consultation

Model OUTPUT in USE
Gemini , SAP-ISU August

2

Objective:

- Objective for this agenda item is:

    – To provide an update on DESC's work plan items for Autumn 2020 / Winter 2021, as agreed last July

DESC Agreed Work Plan Items - Progress

- Table represents agreed work items over the Autumn '20 / Winter '21 period

Description of Work item	Status	Latest Update	DESC Meeting Date (to report / provide update)
Upgrades to CDSP's systems for handling and validating sample data submissions and managing weather data	In Progress	Development / Testing work ongoing	24th Feb'21
Seasonal Normal Review 2020 (completion of remaining activities)	Complete	Final update provided in October	5th Oct'20
Support NDM Algorithm Review	In Progress	UNC Review Group proposal	Standing Agenda Item for 2021 meetings
Algorithm Performance for Gas Year 2019/20 Strand 1 Weather Analysis Strand 2 Unidentified Gas (UiG) Analysis Strand 3 NDM Sample Analysis	Complete	Final document published on 30th December	7th Dec'20
Modelling Approach 2021 preparations (for Gas Year 2021/22)	In Progress	Email sent to DESC/TWG requesting comments	Draft for Review: 7th Dec'20 Final Approval: 24th Feb'21
Model Smoothing Methodology Review	In Progress	Analysis on track	24th Feb'21
Review of current Holiday code rules	On Hold	Modelling Approach document enhanced	24th Feb'21

Work Plan Progress

Upgrades to Systems for handling and validating NDM Sample Data submissions

Introduction

- Actual daily gas consumption data is a key input to the EUC Demand Modelling process which in turn produces the parameters used in the NDM demand allocation formula

- Historically this data was sourced from samples managed by Xoserve and the Distribution Networks only

- However, in April 2018 MOD654 was raised and implemented (as MOD654s) in March 2019

- This modification instructs all Shippers with a portfolio greater than 25,000 sites (35+ Shippers) to support the EUC Demand Modelling process by submitting daily gas consumption data to Xoserve

- This has greatly boosted the number of sample points available to Xoserve but has required upgrades to our processes to handle the volumes.

Current State Review

- The number of data source providers the Demand Estimation Team now handles has increased from 3 to over 40

- The current process uses multiple Microsoft Access databases which are being pushed to their limits given the large increase in data submissions

- This current set-up makes it difficult to maintain a robust audit trail and any necessary performance reporting, which can mean tricky to get a picture of whether we have sufficient data until sometimes its too late

- Reporting is even more important now, given requirements to update Performance Assurance Committee (PAC) as well as DESC

Future State Goals

- Design a simplified end to end process and system for managing all submissions of daily gas consumption data used to support Demand Estimation processes

- Select software which is more suitable to …
  - handling and storing large volumes of data
  - processing the data in much quicker timescales
  - provide better analysis functionality which can enable more data errors to be discovered

- Host all sample management activities of handling, loading, cleansing and validating in one system, thus reducing hand offs/inefficiencies and ensure any internal / external reporting is more reliable ('one version of the truth')

- Ensure the team can review data submissions and make judgements through the year on potential sampling gaps in 'real time'

- We are in the process of transferring these processes into SAS, along with several process improvements

What does this mean for you?

- The Demand Estimation Team will be able to monitor the sample position throughout the year and request sites to fill any gaps in EUCs accordingly

- The enhanced data analysis functionality should result in cleaner data being put forward to modelling, which will help contribute to more accurate profiles

- Processes will be hosted on a more stable platform and provide more opportunities to interact with other systems/data feeds such as Class 3 reads

- The new system will be more controlled which means the industry file format for third party submissions will need to be adhered to e.g. file format and naming convention. The vast majority of data submissions meets the criteria, however where this is not met, we shall need to return the files to sender for correction

- It will also allow for real time validation and reporting, which will hopefully allow for corrections to made well ahead of when the data is used

- It's planned that the system goes live in April to support this year's Modelling process

Progress to date

- Process for receiving, checking and loading all types of data files complete

% Complete	WBS	Task Name	LWI reference	Predecessors	Duration (days)	Start	Finish	Resource Names
	14.1	Analysis and Planning						n/a
100%	14.1.1	Start Process Flow				11/2/2020	12/4/2020	SB
100%	14.1.2	Develop Process Flow				12/30/2020	12/4/2020	SB
100%	14.1.3	Develop and Review Process Flow				12/1/2020	12/4/2020	KS
100%	14.1.4	Review Process Flow				12/2/2020	12/14/2020	SB / KS
100%	14.1.5	Share Process with Team				12/3/2020	12/3/2020	All
100%	14.1.6	Finalise Process Flow				12/7/2020	12/7/2020	All
100%	14.1.8	Naming Convention				12/2/2020	12/2/2020	SB
100%	14.1.9	Table Atrributes				12/3/2020	12/4/2020	SB
100%	14.1.10	Format of Fields: Data Dictionary				10/1/2020	10/1/2020	SB / KS
100%	14.1.11	Entity Relationship Diagram				10/1/2020	10/1/2020	SB / KS
	14.2	File Loads						n/a
100%	14.2.1	Extract files from box account into one area (Technolog, Shipper and Network)						KS
100%	14.2.2	File Format Check Code						SB / JL / KS
100%	14.2.3	Produce table for Shipper				1/6/2021	1/15/2021	SB / KS
100%	14.2.4	Produce table for Technolog				1/6/2021	1/15/2021	SB / KS
100%	14.2.5	Produce table for Network				1/6/2021	1/15/2021	SB / KS
100%	14.2.6	Produce validation file merge table				1/6/2021	1/15/2021	SB / KS

- We are now working on the various validations such as:
  - Duplicates
  - Negative Volumes
  - Aggregated Consumption
  - Day of the Week Check
  - Consecutive Zeros
  - Volume Spikes

- On track to utilise new system for sample data handling in April's data collection and validation

10

Conclusions

- The implementation of MOD654S has meant that Xoserve has started to receive data submissions from 35+ different parties

- To help us manage the increase, we needed to upgrade the processes and systems associated with the sample data

- As a result of this, it is essential that the data is received in the correct file format and has been checked for data errors prior to submission to Xoserve

- The introduction of a single system with more controls means we cannot accept files outside of designated file format, hopefully this is understood given the numbers of external parties now actively sending data to us
  - Note: We shall ensure you are notified asap when we have experienced any issues with the files. Where data is deemed to be unusable this will be included in notifications to PAC along with reports of failures to provide any data

- We need to work together to ensure that the EUC demand models we produce benefit the whole industry and are as accurate as possible.

Work Plan Progress

Upgrades to Systems for Managing Weather Data

Weather Handling System - Background

- As part of change XRN4772, the CWV Formula now uses 24 hourly weather observations and weightings for each weather variable to calculate daily weather values. As a reminder, the weightings applied to each observation have been manipulated such that this change is not evident in the resultant CWV value

- The addition of Solar Radiation and Precipitation (albeit currently dormant within the formula), as well as additional weather stations required, means a much greater volume of data is required to be stored and used to calculate historic and daily CWVs

- As such, there is a necessity to upgrade CDSP's systems for managing weather data. With the EUC Demand Modelling system, and the Sample Validation System both moving to SAS platforms, naturally this is also the chosen platform for a weather handling system

Weather Handling System - Benefits

- Previously mentioned systems (e.g. sample data) will have greater capability to interact with one another reducing duplication of work and minimising potential for any data errors to occur between systems

- The Demand Estimation team with have greater ability to monitor reactions in gas demand due to changes in weather conditions throughout the year

- Provide greater flexibility for future work streams relating to the calculation of CWVs, such as, how precipitation may influence gas demand, Algorithm Performance, CWV Parameter Optimisation, Seasonal Normal Reviews etc.

Weather Handling System - Update

Progress to Date:
- All applicable historic weather data spanning from 1st January 1960 onwards has been formatted and imported into the system (currently 14m+ records) to feed into CWV calculations
- This data is kept up to date based on the flow of weather data received from Xoserve's Weather Service Provider
- The core aspects of the system will be in place and support this year's Modelling process

Observations:
- Similar to analysis conducted on Solar Radiation, there are a number of missing records in the historic Precipitation data such as the WSSM datasets, there is therefore a necessity to devise an infill methodology to form a complete history for all applicable weather stations. We would welcome any DESC members thoughts or insights on such methodology

- Xoserve are currently in the process of obtaining additional weather data to 'fill in' any gaps in the recent history for applicable Weather Stations