## Spring 2008 NDM Analysis - Proposed Approach

### Background

Ahead of each annual NDM analysis, it is customary to prepare a note for DESC setting out the proposed approach to the next NDM analysis. This note deals with the overall approach to the analysis and the approach to model smoothing.

A detailed analysis of model smoothing was presented to the DESC meeting on 8th November 2007 and it was confirmed that model smoothing should continue to be applied in spring 2008. At this same meeting DESC also agreed that should it become necessary in spring 2008, due to limited sample strength, the data sets applicable to consumption bands 07 and 08 could be combined for both consumption band and WAR band EUC modelling in these consumption ranges.

This year there are no weather station changes expected ahead the start of gas year 2008/09. Therefore no consequential changes to CWV definitions are expected.

This note summarises the overall approach proposed for the spring 2008 NDM analysis. Please note significant changes from last years approach are shown in blue text.

### Specific Points of Detail

#### Model Smoothing

This section and the points below reflect the agreement reached at the aforementioned DESC meeting.

1. Year on year model smoothing will be used in the spring 2008 NDM analysis, in deriving the NDM proposals to be applied to gas year 2008/09.

2. In the absence of evidence of trends in the parameters of the year on year models, and taking into account the discussions at DESC in November 2007, averaging will be applied to the NDM models feeding into model smoothing.

3. The NDM models for three years will be used for model smoothing. The three years will be 2005/06, 2006/07 and 2007/08. For both the second and third of these three analysis years (i.e. the middle and most recent), all data sets cover a twelve month period (mid-March to mid-March or April to March) while for the first (i.e. the earliest) analysis year, 2005/06, the data sets applied to modelling EUCs above 293 MWh pa cover a 13 month period March to March; this is necessary to ensure that there is at least one Easter holiday weekend in the data sets for that year. Note that the mid-March to mid-March data sets of 12 months duration, which are applied to modelling EUCs below 293 MWh pa will always include at least one Easter holiday period in every year. Note also that in the third of these three analysis years (2007/08), the April to March period contains two distinct Easter holiday weekends (in April 2007 and in March 2008).

4. In applying smoothing, models from equivalent WAR bands in the three separate years will be averaged although WAR band limits change from year to year. This is the approach adopted for each NDM analysis since spring 1999 (i.e. all previous NDM analyses in which model smoothing was applied) and has been deemed by DESC as the most suitable. The analysis will also retain the current period (December to March) in the definition of the WAR values for the spring 2008 analysis.

5. The approach to model smoothing will be at the level of the underlying demand models, as was the case in the previous analyses. Further details are attached as an appendix to this note.

6. In line with previous commitments, it is proposed that the approach to model smoothing should be reviewed in full again by DESC during the autumn in two years time, following finalisation of the NDM algorithms for 2009/10.

#### Model Re-runs

1. To assist in any investigation of trends, all three years (i.e. 2004/05, 2005/06 and 2006/07) used in the spring 2007 implementation of model smoothing will be re-run to correctly take into account the holiday periods applicable to the spring 2008 NDM analysis.

2. In order to facilitate the possible introduction (should the analyses indicate merit in doing so) of a new consumption band breakpoint at 1465 MWh pa in terms of EUC definitions, appropriate back-runs of these consumption band aggregations will be undertaken.

3. Only the re-runs from the 2005/06 and 2006/07 data sets will be used (along with the new data sets for 2007/08) in model smoothing, making up the three years of data applied in the spring 2008 analysis.

4. For EUCs in consumption ranges above 293 MWh pa, the data sets will cover the 12 month period April to March in 2006/07 and 2007/08. For these EUCs in 2005/06, the data sets will cover the 13 month period March to March (to ensure the presence of at least one Easter holiday weekend in that year's data sets).

5. For EUCs for the consumption ranges 0-73.2 MWh pa and 73.2-293 MWh pa, the analyses will in the case of all three years be undertaken on data sets covering 12 months. These data sets run from mid-March to mid-March in each analysis year and therefore always contain at least one Easter holiday weekend.

6. The set of holiday days applied to the analyses will be the union of the holidays applying to England and Wales on the one hand and Scotland on the other. This approach has been used since the adoption of model smoothing in spring 1999 and continues to be appropriate because EUC sample data from geographically adjacent LDZs are usually aggregated to allow some EUCs to be modelled. Both population and sample disposition are such that this aggregation of data is essential to enable modelling of all EUCs in all LDZs. The disposition of holiday codes and the actual holiday factor values (if any) that are applied will be derived from the modelling and will, of course, be as indicated by the characteristics of the various applicable data sets themselves.

**Modelling Details**

1. The general modelling approach to be adopted for the spring 2008 analysis will be the same as that applied in spring 2007. This approach is detailed in the flowcharts on pages 29 and 30 of the spring 2007 NDM Report. A broad outline of the approach is reproduced below :

    a. Exclude warm weather data and summer data (ie. June to September) and fit a line to the remaining data. Any flat models are detected and re-run with all the data.

    b. Warm weather data (for exclusion) is defined in this context as the warmest $2^o$ of data (ie. that for which the CWV is greater than Max. CWV - $2^o$).

    c. Assess the excluded summer data against the line fitted in step a to establish whether a summer reduction is required. It is proposed that the current condition of a 5% bar before any summer reduction is considered to apply is retained.

    d. Reintroduce the summer data into the data set (after inflating by any summer reduction identified in step c; if no summer reduction is identified then there would be no inflation). Fit a line to the augmented data set, excluding the warmest $2^o$, to establish whether a cut-off is appropriate, considering potential cut-offs in the range 0.5 to 4 degrees below the maximum value of the composite weather variable. It is proposed that the criterion applied from spring 2001 onwards, of a 20% improvement in the mean square residual over that obtained by using the straight line alone, is retained in assessing whether or not there should be a cut-off.

    e. If a cut-off is not required, then reintroduce the warmest $2^o$ of data and fit a line to the entire data set.

    f. Model smoothing considers three years' models and the application of summer reductions or not to the smoothed model is dependent on all of the years contributing to the smoothed model. Thus it is possible that the smoothed model will not incorporate a summer reduction, in spite of a summer reduction being identified for one (or more) of the individual years. To cover this eventuality it is necessary in each year's modelling to produce models with and without summer reductions. The model without summer reductions will be produced by including summer data (except for the warmest $2^o$) in the regression in step a above, and fitting a cut-off if necessary, as in steps d and e above.

2. As previously agreed and implemented from the spring 2002 NDM analysis onwards, weekend effects for the "01B" EUCs will be modelled using the same "variable weather sensitivity" form of model used for all other EUCs. (This form of the model is set out in Appendix 3, on page 21 of the spring 2006 NDM Report.) As in previous years, holidays will not be excluded from the regression models for "01B" EUCs.

3. The data applicable to the analysis year 2007/08 will not have been analysed previously, and so, investigation of the most appropriate data aggregations, determination of WAR band limits, etc., will be undertaken with respect to this data set.

4. The models for all EUCs will allow the possibility of summer cut-offs and summer reductions being applied. Note however that cut-offs will not be applied to the models derived for consumption bands up to 293 MWh pa (i.e. the "01B" and "02B" EUCs), for the spring 2008 analysis. This amended approach was agreed by DESC in December 2003, with a view to mitigating summer scaling factor instability and was also applied to all previous NDM analyses from spring 2004 onwards

5. In any single LDZ, the same definition of CWV will be used for all runs (i.e. for all EUCs in that LDZ and for all years of data).

6. Weekend, holiday and summer reductions will be calculated (where appropriate) as the average of the percentage reductions estimated for the three individual years' models; where applicable the CWV cut-off (at which models cease to be weather sensitive) will be the simple average of the three separate estimates. (If for one or two of the three years there is no CWV cut-off, the maximum value of the CWV will be substituted as the cut-off for those years.) Further details are provided in the Appendix attached hereto.

7. As set out in the Appendix, the key aspect of averaging the models will be to average the ratio of the slope to the constant term, from each year's model. These ratios are equivalent to the reciprocals of the CWV intercepts.

8. Prior to the averaging, any models giving non-negative slopes on initial analysis (excluding the warmest weather from the regression), will be re-fitted to the entire data set. Any positive slopes remaining will be set to zero. This has become established practice.

9.  In line with previous practice, a single EUC will be proposed in each LDZ for the consumption range 0-73.2 MWh pa. Also, in accordance with previous practice, domestic only data sets will be applied to modelling this consumption range. Following the detailed investigation (reported to DESC on 8th November 2007) of the modelling of this consumption range as two sub bands split at 20 MWh pa and at 30 MWh pa, with and without non-domestic supply points included in the upper sub-band, it was shown that given the available sample strength no statistical improvement in the representation of the consumption range was obtained by either sub-bands or the inclusion of non-domestic supply points. Note that it has been agreed at DESC that this investigation of sub-band analyses will be repeated in autumn 2008 and bi-annually thereafter.

10. For the "01B" EUCs, a pragmatic approach will be taken with respect to weekend effects taking into consideration both observed weekend scaling factor patterns and the prevalence or otherwise of non-positive and/or non-statistically significant (at the 95% level) weekend effects arising from the individual years' models over the weekend days (Friday, Saturday, Sunday). This is consistent with the approach taken during spring 2007 and previous years' NDM analyses.

11. For large NDM (i.e. above 2196 MWh pa), the consumption band break points by which large NDM EUCs are defined will remain the same as in spring 2007 (and all previous years) as required by UNC Clause H1.7.5. However, if sample numbers prove to be insufficient in the most recent year's data sets, it has been agreed with DESC that the samples applicable to consumption ranges 14650 - 29300 MWh pa and 29300 - 58600 MWh pa (EUC bands 07 and 08, respectively) may be combined to obtain samples of adequate strength for more robust demand modelling.

    The data may be combined in this way for either the consumption band EUCs or for the WAR band EUCs or for both, depending on whether sample deficiencies require it. Even if data is combined in this way, separate EUCs will be defined for consumption band and WAR band EUCs in the consumption ranges 14650 - 29300 MWh pa and 29300 - 58600 MWh pa.

    However, if data has been combined, the underlying demand models applicable to the most recent year of data will be the same for some of these EUCs. For the avoidance of doubt, previous years' data sets for these EUCs, included in deriving the final smoothed model, will be at the level of aggregation that applied at the time (i.e. bands 07 and 08 are separately modelled). Thus, despite any aggregation of data in the most recent year's data sets, the derived factors (i.e. ALPs, DAFs and load factors) will in general be different for each of these EUCs.

## Exploratory Analyses

As with previous years including spring 2007, it is proposed that the exploratory NDM analyses will focus on confirming EUC definitions (small NDM only) and establishing the most appropriate levels of aggregation to apply to the data sets for the various EUC analyses.

Additionally, as in all previous years, it is proposed that exploratory analyses will be undertaken to establish whether to analyse the consumption range 293-2196 MWh pa as a single data set or to undertake the analysis over this consumption range in more than one set. The alternatives to be explored are: 293-732 MWh pa and 732-2196 MWh pa, which was the approach adopted, for these consumption band EUCs only, from the spring 2002 analysis onwards, or including a further consumption band breakpoint at 1465 MWh pa to give three consumption band ranges: 293-732 MWh pa, 732-1465 MWh pa and 1465-2196 MWh pa.

Similarly, as in all previous years, it is proposed that exploratory analyses will be undertaken to establish whether to continue to analyse the consumption range 73.2-293 MWh pa as a single data set.

In line with previous practice, WAR band EUCs over the consumption range 293-2196 MWh pa will be based on the overall range, which should then enable analysis by individual LDZ instead of LDZ groupings.

Note that, apart from the requirements of potential new EUC definition break-point at 1465 MWh pa, whatever the outcome of the analysis of the 2007/08 NDM sample data, the models from the previous two years, used for smoothing purposes, will be based on the aggregations applied in those years' analyses.

## Derived Factors

1.  The DAFs for gas year 2008/09 will continue to be computed based on aggregate NDM demand.

2.  The load factor computation for each EUC will be based on the smoothed model. This is the same approach as adopted for all NDM analyses since spring 1999.

4.  In calculating DAF values in the case where the smoothed model has a cut-off, the reduction in the magnitude of weather sensitivity will be phased in as described on page 70 of the spring 2007 NDM Report. This approach has been in place since its introduction at the time of the spring 1997 NDM analysis.

5.  In the context of the non-application of cut-offs to EUC models in consumption range 0-293 MWh pa, and as agreed by DESC in December 2003, the values of ALPs for EUCs in this consumption range will be constrained to be never less that 1% of their maximum values. Note that this is a safeguard against a theoretical possibility of negative ALPs arising (in the profiles proposed for all gas years since 2004/05 it has not been necessary to invoke this constraint).

**Fallback Position**

For clarification, it should be noted that should the final NDM proposals made for gas year 2008/09 be formally disapproved as a result of an application for disapproval made to and upheld by Ofgem, then the fallback position that would apply is that EUC definitions and derived factors (ALPs, DAFs and load factors) would be applied to gas year 2008/09 based on the underlying EUC demand models from the spring 2007 NDM analysis and the forecast aggregate NDM demand model for 2008/09 made in spring 2007 (as amended for the changes in the day of the week and holidays applicable in 2008/09).

**Reporting**

The parameters for the smoothed models will be published, in an Appendix to the spring 2008 NDM Report.

As usual, all model parameters (for each of the three years feeding into model smoothing) will be provided in electronic form.

For all final smoothed EUC models, information (i.e. values of factors and flags where these apply to each model) pertaining to: summer cut-off, summer reduction, non-holiday weekend effects, and holiday effects will be provided in electronic form.

As in the spring 2007 NDM Report, all CWV intercepts (for each year's models and for the smoothed model) will be included in the report.

The performance evaluation appendix of the spring 2008 NDM Report will continue to have the now customary three strands of information (WCFs & SFs, RVs and NDM sample analysis).

## Appendix - EUC Model Smoothing

The key stages of the end user category (EUC) model smoothing process are explained below. This is unchanged from previous practice.

Produce models for the EUC based on the data for each of the last three years. In the case that summer reductions have been applied in an individual year, two versions of the EUC model for that year exist, one with summer reductions and one without summer reductions. Where summer reductions are applied, the magnitude of these reductions is expressed in terms of a summer multiplier applied to the fitted daily demands over the non holiday days from the spring bank holiday period to the last weekend in September. For example, a summer multiplier of 0.850 means that fitted demands are reduced by 15% over this period. If no summer reductions are applied, the summer multiplier takes a value of 1.

Decide whether to apply summer reductions to the final smoothed model. The criterion applied in making this decision is as follows. The summer multipliers for the three individual year models for the EUC are averaged. If this average summer multiplier is less than the critical value of 0.9 (a 10% reduction), summer reductions are applied in the smoothed model; the summer multiplier for the smoothed model is this average value. If the average summer multiplier is greater than or equal to the critical value, summer reductions are not applied to the smoothed model.

For example, for an EUC with summer multipliers of 1.000 (i.e. no summer reductions), 0.800, and 0.850 in the individual years, the average summer multiplier is 0.883. This is less than the critical value of 0.9, so a summer reduction is applied to the smoothed model.

This decision process allows a unique EUC model to be selected for each individual year. If summer reductions are to be applied in the smoothed model, the version of each individual year's model with summer reductions (if such a version exists) is selected. Otherwise, the version without summer reductions is selected for each individual year.

At this stage, the decision as to whether to set weekend effects to zero is taken.

The selected individual year models for the EUC are standardised, by dividing through by the constant for that individual year. This gives a model for each year (yr) of the form:

$$D_t(yr) = 1 + C_2(yr)*CWVt + C_3(yr)*Fri + C_4(yr)*Sat + C_5(yr)*Sun$$

This standardisation ensures that all three individual year models give the same normalised daily demand value (i.e. 1.0) for a non-holiday Monday to Thursday at 0° CWV. This ensures that equal weight is given to each individual year in the smoothing process.

Each individual parameter of the initial smoothed model for the EUC is calculated by averaging the values of the parameter over the three individual years.

For example,        $C2(smoothed) = \{C2(yr. 1) + C2(yr. 2) + C2(yr. 3)\}/3$

The constant (which is 1 in the standardised model) and the slope of the smoothed model are then multiplied by the constant term of the original (unstandardised) model for the most recent year. Note that this step has no effect on the NDM profiling or capacity estimation parameters, but it gives model parameters of the same scale as that of the model for the most recent individual year.

The multiplicative day of week/holiday factors ($P_t$ as described in Appendix 3 of the spring 2007 NDM Report) are calculated for the smoothed model for the EUC. These are calculated for each day as averages of the corresponding values in the three individual years' models.

A decision is made as to whether to apply a composite weather variable cut-off to the smoothed model for the EUC. Application of a CWV cut-off has the effect of causing the fitted demand to level off for values of CWV above the cut-off. The criterion used in making the decision is as follows. The value of the CWV cut-off is estimated for each year's model. If no cut-off is required, the cut-off value for that year is set to the maximum CWV for the LDZ. The three individual years' CWV cut-offs are then averaged. If this average value is less than the maximum CWV for that LDZ, a CWV cut-off is set at this value in the smoothed model. Otherwise no CWV cut-off is applied to the smoothed model. Note however that cut-offs will not be applied to the models derived for consumption bands up to 293 MWh pa (i.e. the "01B" and "02B" EUCs), for the spring 2008 analysis. This amended approach was agreed by DESC in December 2003, with a view to mitigating summer scaling factor instability and has been applied to all NDM analyses since spring 2004.

The ensuing form of model is used in the calculation of NDM profiling parameters and capacity estimation parameters.

A form of the smoothed model is also produced with additive weekend effects. The averaged standardised parameters for each day from Friday to Sunday are multiplied by the constant term of the original unstandardised model for the most recent year, to give additive weekend effects for the smoothed model. This gives a smoothed model of the form:

$$D_t = C_1 + C_2*CWVt + C_3*Fri + C_4*Sat + C_5*Sun$$

$C_1$ has the same value as the constant term of the EUC model for the most recent year. This is a simple form of the smoothed model because it does not embody such features as holiday effects, summer cut-offs and summer reductions. The parameter values for this form of model will be shown in the spring 2008 NDM Report, for consistency with previous years' reports.