NTEM Dataset Version 7.2

Summary of Quality Assurance

Department for Transport

February 2017

Notice

This document and its contents have been prepared and are intended solely for Department for Transport's information and use in relation to Task SO7417: National Trip End Model Dataset Update.

Atkins Ltd assumes no responsibility to any other party in respect of or arising out of or in connection with this document and/or its contents.

This document has 16 pages including the cover.

Document history

Job number: 5138121			**Document ref: SO7417/3**			
Revision	Purpose description	Originated	Checked	Reviewed	Authorised	Date
Rev 1	Draft issued for comment	DW	CLFL	CS		21/02/17
Rev 2	Minor revisions	DW	AKS	CLFL	PR	27/02/17

Table of contents

1. Introduction 4
1.1. Purpose 4
1.2. Overview of NTEM forecasting suite 4
1.3. Principles of QA 5
1.4. Structure of document 5

2. QA of model inputs 6
2.1. Scope 6
2.2. Comparison checks 6
2.3. Process checks 7

3. QA of model processes 8
3.1. Scope 8
3.2. Software checks 8
3.3. Model run checks 9

4. QA of model outputs 11
4.1. Scope 11
4.2. Sense checks 11
4.3. Comparison with NTEM 6.2 13

5. Outcomes 15

1. Introduction

1.1. Purpose

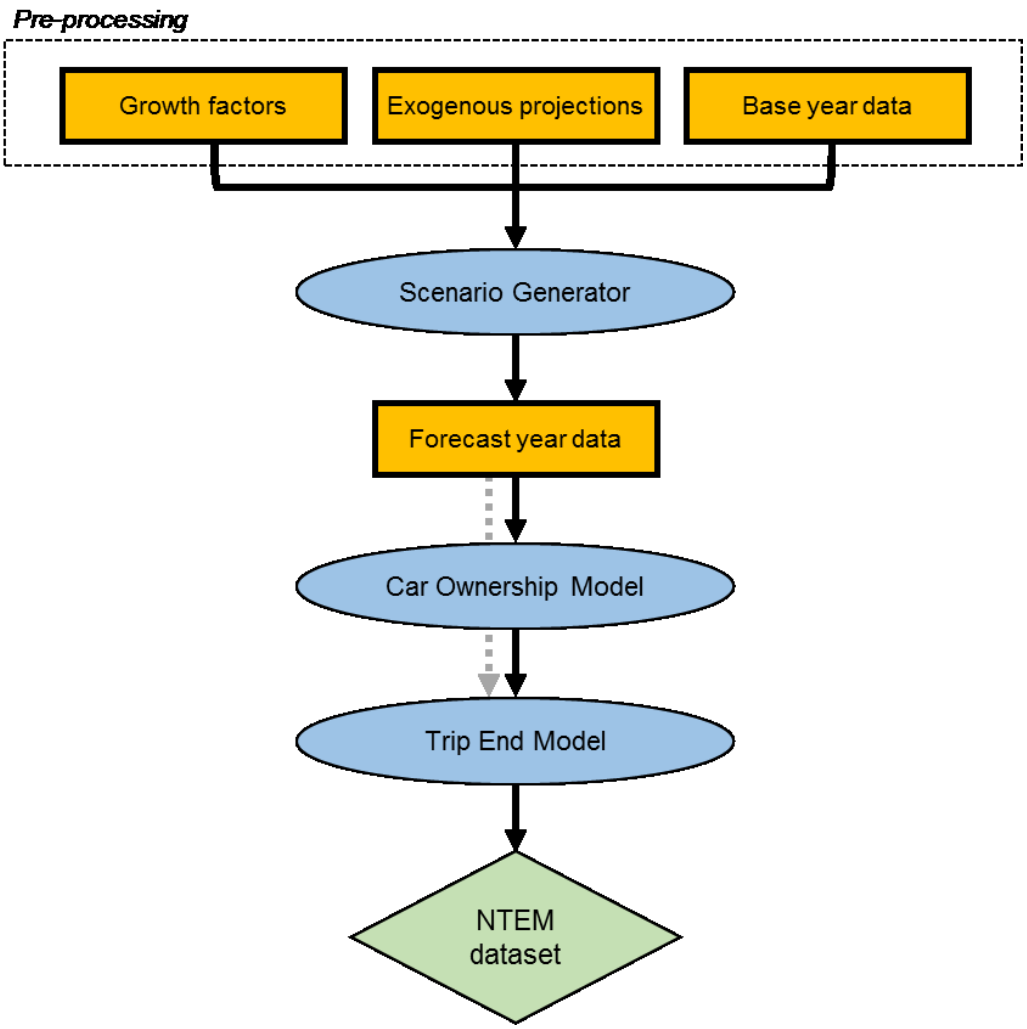
1.1.1. This report describes the quality assurance (QA) framework applied to the production of the NTEM 7.2 dataset. This report is divided into five sections.

1.1.2. The introductory section (Section 1) describes the principles of the QA framework. It also gives an overview of the NTEM modelling process. The remaining sections describe the quality assurance undertaken to verify model inputs, processes and the outputs of the NTEM forecasting suite.

1.2. Overview of NTEM forecasting suite

1.2.1. An overview of the process involved in the production of NTEM forecasts can be seen in Figure 1-1. The main model processes cover the *Scenario Generator* (ScenGen), which produces demographic forecasts, the *National Car Ownership Model* (NATCOP), which produces forecasts of car ownership, and the *National Trip End Model* (CTripEnd), which forecasts segmented trip ends.

Figure 1-1: Basic steps in generating the NTEM 7.2 dataset



1.2.2. Each step requires preparation of input data, coding, and ultimately running model software. Whilst the main processes are automated, there are a number of manual steps that must be undertaken. Full details of the process undertaken are covered in NTEM Planning Data Guidance Note v7.2[1].

1.2.3. Results from the NTEM forecasting suite are ultimately made available through the Department for Transport's (DfT's) TEMPro software. For information on this, please see the TEMPro documentation[2].

1.3. Principles of QA

1.3.1. The DfT's *Quality Assurance of Analytical Modelling*[3] defines a model as typically having three parts:

- Inputs – in the form of data and assumptions;
- A processing component – often through calculations; and,
- Outputs.

1.3.2. Whilst the NTEM forecasting suite is a series of models, these basic components are still relevant. The principles of QA applied here are based on ensuring model accuracy and reliability – that the model produces results that are as intended given the inputs provided.

1.3.3. Outputs from NTEM span 7,700 model zones, 15 trip purposes, 6 time of day categories, 6 mode categories, 4 planning data types and 3 person types. Output tables from NTEM contain more than 20 million cells of trip data and nearly 1 million cells of planning data. Across all model years this represents nearly 200 million output cells.

1.3.4. As the dataset is large, a robust and proportionate QA process requires a sufficient level of detail in order to provide users with confidence in the outputs. In particular, QA of the NTEM dataset should reflect the ultimate use in TEMPro. It is recognised that it is not plausible to scrutinise every data element. The general principles of the QA framework should:

- Reflect the typical level of segmentation required (e.g. area, mode, purpose, time of day);
- Reflect the spatial granularity of the input data; and,
- Seek to identify and explain outliers and anomalous results.

1.3.5. A robust and proportionate quality assurance framework has been applied to demonstrate the care, attention and due diligence undertaken in production of the NTEM dataset, and to provide confidence in the final dataset.

1.3.6. It is important to recognise the uncertainty associated with forecasting in transport modelling. This is separate from model quality assurance, and is described in the DfT's *Transport Analysis Guidance*, particularly TAG Unit M4.

1.4. Structure of document

1.4.1. The remaining sections of this document are structured around the QA undertaken on the NTEM 7.2 dataset. The sections are:

- QA of model inputs;
- QA of model processes (comprising software and model runs);
- QA of model outputs (comprising sense checks and comparisons of data available to TEMPro users); and,
- Outcomes of the QA process.

[1] For ScenGen processes and forecasts, and full NTEM forecasts see NTEM 7.2 Planning Data Guidance Note

[2] Available on the TEMPro website

[3] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/350904/qa-modelling-guidance_pdf.pdf

2. QA of model inputs

2.1. Scope

2.1.1. The NTEM forecasting suite, consisting of the Scenario Generator, National Car Ownership Model, and National Trip End Model, requires a variety of input data from various sources.

2.1.2. The National Car Ownership Model (NATCOP) was re-estimated to a 2011 base year. This process, alongside the quality assurance undertaken, is detailed in the estimation report[4].

2.1.3. Trip rates and associated parameters for use in the National Trip End Model (CTripEnd) were estimated as part of the *Provision of Travel Trends Analysis and Forecast Model Research* project, undertaken by AECOM and Atkins on behalf of the DfT. Quality assurance of these is documented in the associated report[5].

2.1.4. The quality assurance documented here relates to the Scenario Generator (ScenGen) input data. This data is described in detail in NTEM 7.2 Planning Data Guidance Note. Checks are broadly split into *comparison checks* and *process checks*.

2.2. Comparison checks

2.2.1. The major inputs relate to base and forecast year data for population and employment. For the base year (2011), Census data was processed to provide the input data required by NTEM. This included applying a mid-year adjustment, and removal of communal residents (who are not included in the NTEM model). Full details of this process are given in Section 3 of NTEM 7.2 Planning Data Guidance Note.

2.2.2. After processing of base year data inputs, these were compared to raw Census data at a disaggregate level. These checks ensured that the distribution of population, households, dwellings and jobs is consistent with the 2011 Census after processing. These checks are summarised in Table 2-1.

2.2.3. Further checks, covering base, forecast and parameter data, included comparison to NTEM 6.2 input data. This provided a general sense check to ensure data is of the appropriate magnitude, and provided insight into why NTEM 7.2 model results are different to NTEM 6.2.

Table 2-1: Input data comparison checks

Check	Details	Status
Granular comparison with Census data	The Census is the major source of data, and as such inputs to NTEM (which have been processed) were compared at a spatially disaggregate level (e.g. through scatter plots), to identify any errors.	Completed No issues
Trend comparison with NTEM 6.2 (at district level)	Forecast growth rates of each input were compared and sense checked, considering both the distribution of growth, the range of growth expected, and the growth of individual model areas – whilst differences are expected, this serves as a sense check and also provides explanation for ultimate model outputs.	Completed No issues

[4] Estimation of the National Car Ownership Model for Great Britain, 2011 base (RAND Europe, 2016)

[5] Provision of Travel Trends Analysis and Forecasting Model Research: Quality Assurance Report (AECOM and Atkins, 2016)

Check	Details	Status
Parameter sense check and comparison with NTEM 6.2	Comparison against NTEM 6.2 – provided a general sense check that parameters are of the appropriate magnitude; suitability of these parameters (and model inputs more generally) is considered through QA of model outputs.	Completed No issues

2.2.4. Scrutiny of the model outputs (described in Section 4) provides an additional check on the quality of model inputs and parameters.

2.3. Process checks

2.3.1. As an initial step, workbooks used for producing inputs in the required dimensions were checked by a team member not involved in set up. Since these processes (particularly for input data) can span several workbooks and involve multiple data sources, these checks are documented in a single place for each dataset.

2.3.2. Forecast dwellings are transcribed from various published documents. There is therefore a risk of manual error of transcription. Separate from this QA process, the DfT commissioned an independent audit of dwellings inputs, and we have implemented the resulting recommendations. This is alongside further checks on the process of deriving inputs, and on the scrutiny of model outputs (Section 4).

2.3.3. Whilst data relating to forecast population and households is taken directly from published ONS projections[7], forecast employment involves a more complex process. A parallel processing workbook has been created, confirming model inputs are as intended.

Table 2-2: Input data process checks

Check	Details	Status
Independent check of process	Workbooks were checked by a team member not involved with set up, validating that processes have been carried out correctly. This requires replication of results through repeat implementation or sample manual checks.	Completed No issues
Construction of verification workbook	As the processes required in the derivation of forecast employment are more complex than other forecast data, a verification workbook has been created on a subset of data, confirming the process applied is correct.	Completed No issues

2.3.4. Quality assurance of input data, documented in Table 2-1 and Table 2-2, has been completed and no issues were found.

[7] An extrapolation process is applied to extend projections to 2051; see Section 4 of NTEM 7.2 Planning Data Guidance Note for further details

3. QA of model processes

3.1. Scope

3.1.1. The processing component of NTEM comprises the modelling software, and the process of running the models with the desired input data.

3.1.2. The NTEM modelling suite comprises three software elements which are run individually. In the case of NATCOP, the software was updated to incorporate a new base (2011) and new estimation. This is documented in the associated estimation report[8] and developers documentation[9].

3.1.3. For the ScenGen and CTripEnd components, the software was updated to run on modern operating systems, with the code base updated from VB5 to VB.NET. Furthermore, the software was extended to allow flexible definition of age bands, and to provide additional outputs.

3.1.4. As the NTEM forecasting suite contains a number of manual processes (as shown in Figure 1-1), further checks are required in order to ensure model runs were set up appropriately, giving reference to the correct input data files.

3.2. Software checks

3.2.1. The mathematical processes undertaken by the ScenGen and CTripEnd models were not altered as part of the NTEM 7.2 dataset update. However, it was necessary to update software to allow running on newer operating systems. Furthermore, software was updated to extend functionality, improve the user interface (UI), and provide a greater range of outputs.

3.2.2. As part of this development, general model functionality was tested to ensure the UI worked as expected, including data import, export and processing functions.

3.2.3. In order to ensure that no errors had been introduced to the methodology as a result of upgrading and extending the source code, model results for a previous NTEM dataset (v6.2) were recreated using the updated software (incremental testing). These checks are summarised in Table 3-2.

3.2.4. Checks covered detailed scrutiny of disaggregate outputs, and, in the case of ScenGen, comparison of the model log files. Model results were found to match at both aggregate and disaggregate levels, reflecting the segmentation available in both models.

3.2.5. In some instances a small difference was found at the third decimal place. These differences have been investigated and explained by computational floating point differences.

3.2.6. Based on these incremental tests, we have concluded that the upgraded software adequately reproduces previous model outputs, and is functioning as intended.

[8] Estimation of the National Car Ownership Model for Great Britain, 2011 base (RAND Europe, 2016)

[9] NATCOP: Software developers note and QA (RAND Europe, 2016)

Table 3-1: Software checks

Check	Details	Status
ScenGen – model outputs	NTEM 6.2 model outputs were recreated and compared against originals – this includes detailed outputs for population, households and employment.	Completed No issues
ScenGen – model log file	The log file outputs are compared, ensuring the same information is reported.	Completed No issues
CTripEnd – model outputs	NTEM 6.2 model outputs were recreated and compared against originals – this includes segmented trip ends and demographic information including car ownership.	Completed No issues

3.3. Model run checks

3.3.1. Manual steps are required to set up and transfer data between models. Given this, there is risk of human error in generating the NTEM dataset.

3.3.2. Manual checks of model setup and run files were carried out to mitigate that risk. In all cases, these checks were carried out by a team member familiar with the software and model running process, but not involved in the initial set up. The model run checks undertaken are shown in Table 3-2.

3.3.3. Checks on the input data ensured that models have been set up correctly with the desired inputs, and that data has been transferred between model components in the correct way. Further checks on log files (where appropriate) demonstrated correct application of the input data, user-specified parameters and scenario definition.

3.3.4. Based on these checks, we have concluded that model set up and running has been carried out as intended.

Table 3-2: Model run checks

Check	Details	Status
ScenGen – base database	Database was populated with desired base year inputs and appropriate definitions.	Completed No issues
ScenGen – forecast database	Database was populated with desired forecast inputs and appropriate definitions.	Completed No issues
ScenGen – main database	Database was populated with appropriate definitions and consistent parameter selections.	Completed No issues

Check	Details	Status
ScenGen – log file check	Scenarios were appropriately defined, including previous scenarios (e.g. West Midlands 2021 should follow West Midlands 2016), parameter application is consistent and there are no error messages.	Completed No issues
NATCOP – log file	Model runs were coded correctly with the relevant (year-specific) ScenGen and parameter inputs.	Completed No issues
NATCOP – parameters	Parameter file were correctly set up with trends as desired (income factors and operating costs).	Completed No issues
CTripEnd – database setup (each year)	CTripEnd databases were set up properly for each year with consistent definitions, zone system and parameter selections (excluding trip rates).	Completed No issues
CTripEnd – application of trip rates	Variable trip rates were applied as intended (static from 2016 onwards).	Completed No issues
CTripEnd – data inputs	Year-specific data inputs from ScenGen and NATCOP were input correctly.	Completed No issues

4. QA of model outputs

4.1. Scope

4.1.1. A significant proportion of the quality assurance undertaken was focused on the outputs of the model. This is the data ultimately available to users through the TEMPro software.

4.1.2. Section 1.3 of this report set out the principles of this quality assurance. In particular, it is recognised that the scale of the NTEM dataset (more than 200 million output cells) does not allow for scrutiny of all outputs.

4.1.3. There are a number of geographic levels of data within NTEM. Zones are the most disaggregate level, followed by local authority districts, and study areas, which are comprised of the English regions, Scotland and Wales. For forecast years, input data is generally provided at the local authority district level. This is also the main level at which the model processes are conducted.

4.1.4. A robust and proportionate QA process requires an appropriate level of detail in order to give confidence in the final dataset. This should not be limited to trip ends, but should also include the planning data and segmentation available to users through TEMPro.

4.1.5. Whilst previous datasets (such as NTEM 6.2) should not be considered as a dataset to validate against, comparisons provide a useful insight into differences in planning data between two versions. There is no expectation or requirement for close comparison with the NTEM 6.2 dataset, but significant deviations should be explained (and confirmed) by the changes to model inputs.

4.1.6. Checks were therefore divided into two major categories – sense checks, and comparison checks with NTEM 6.2. Sense checks were used to scrutinise model outputs, confirm expectations regarding input data, and provide a prompt for further checking of input data. This requires knowledge of model processes and input datasets.

4.2. Sense checks

4.2.1. Sense checks were conducted on granular model outputs, relating to planning data and trip ends. In the main, this is concerned with the local authority district level, which is the spatial level of the majority of input data and processing within NTEM. These checks are summarised in Table 4-1.

4.2.2. A general sense check has been conducted on regional level outputs (planning data, trip ends) for each model year (2011 – 2051). A sensible trend is expected of increasing outputs (reflecting the input data), with the exception of trip ends, which should fall between 2011 and 2016 due to a decrease in trip rates. Furthermore, outputs should be in line with input data. This was confirmed through scrutiny of the data.

4.2.3. At a detailed spatial level, forecast NTEM population (output) is not expected to match inputs. The Scenario Generator process uses the forecast of additional dwellings in conjunction with ONS trend-based projections of population to redistribute population within a study area. As such, where the output population is significantly different to the input population, this should be explained through the trend in forecast dwellings[10]. This was confirmed for all model areas[11]. Furthermore, for all the areas where the growth in output population is more than 10 percentage points different from the growth in input population, input data was re-checked and confirmed.

[10] See Section 2 of NTEM 7.2 Planning Data Guidance Note for details of the forecast process

[11] At the local authority district level

4.2.4. The correlation between trip ends (productions and attractions) and their major drivers (population and employment respectively) was considered. A high correlation was found, in line with expectations. For areas where there is a noticeable deviation (i.e. productions or attractions are greater or fewer than expected based solely on the population or jobs present), further investigation was carried out. In the case of productions, this can generally be attributed to a high proportion of non-home-based trips (for example, associated with a large non-resident daytime population). In the case of attractions, this can generally be attributed to land use types not associated with jobs (for example, a large resident population).

4.2.5. As there is no representation of transport supply, mode and time of day split parameters are static in the model; changes in mode splits and time of day splits of model outputs are driven by demographic changes only. Model outputs were checked against this expectation, with mode and time of day splits stable through time, and in line with input model parameters (such as changing car ownership).

4.2.6. The growth in trip ends and planning data of individual districts was compared to the study area as a whole. Where these differ greatly, the reason for this is identified through checking *relative* growth of input data. This is generally attributed to relatively high or low inputs for population and dwelling projections, which is confirmed against input data.

4.2.7. There is a general expectation that model outputs – demographic and trip ends – should increase through time. Whilst this is usually the case, instances of negative growth have been identified and can be explained through comparison to input data. Particular instances[12] include a decline in population and a decline in workers. The former is associated with a decline in population from ONS population projections. The latter can be linked with a change in the age structure of the resident population, with a greater number of people aged 75+.

4.2.8. Finally, the changing age structure of the population is considered. This compares the age structure of input population (the aforementioned ONS projections) with output population (NTEM demographic forecasts). Due to the NTEM process (including reallocation of population and removal of communal residents), it is not expected that output age structure will match exactly. However, where there is a noticeable shift in the age structure which will influence trip making (for example, from working age to 75+ age segments) this should be evident, and is confirmed.

Table 4-1: Output sense checks

Check	Details	Status
Regional level outputs (trip ends, planning data) for each model year (2011 – 2051)	A sensible trend should be evident (i.e. increasing population) and regional level outputs should be equivalent to inputs.	Completed No issues
Population reallocation at the district level (2016, 2036, 2051)	Significant reallocation of population should be explained and checked against a dwellings trend.	Completed No issues
Zone and district level outputs for planning and trip end data (all model years)	Outputs should be consistent (i.e. well correlated) with the main drivers of trip ends: ▪ Population and total productions; ▪ Employment and total attractions.	Completed No issues

[12] Trip productions decline from 2011 – 2016 due to an *implemented* change in trip rates.

Check	Details	Status
Mode and time of day split checks (2011, 2036)	Expectations around mode and time of day splits (i.e. dominant modes and times) should be confirmed at the district level. Variation within a region should be explained, for example PT mode shares being higher in urban areas.	Completed No issues
Growth within a region (all data, 2011, 2036)	For each region, growth in total trip ends and headline planning data by district should be compared against the regional average; where these differ greatly, the reasons should be sought.	Completed No issues
Negative growth (all data, 2011, 2016, 2036)	Whilst negative growth is neither unusual nor unexpected, these districts should be identified and confirmed based on the input data used.	Completed No issues
Changing population structure (2011, 2036)	Where there are major changes in population structure (i.e. age structure), these should be explained through comparison to input data, to ensure all processes can be explained.	Completed No issues

4.3. Comparison with NTEM 6.2

4.3.1. Comparing model outputs with the equivalent obtained from NTEM 6.2, and available to TEMPro users, provides an additional check. This is done within the context of input data comparison (as described in Section 2.2). These checks are summarised in Table 4-2.

4.3.2. The distribution of demographic outputs – population, households, jobs, workers – within a region was considered for an equivalent forecast year. An excellent agreement was found, both in terms of ranking and the relative size (i.e. the largest districts remain the largest). Where large absolute differences are apparent, this was confirmed through changes to input data. For example, from comparison of population inputs (ONS projections), some areas have a significantly lower forecast.

4.3.3. Model outputs (demographic forecasts and trip ends) were compared for three model years (2011, 2016, 2036) between the two datasets. A good correlation[13] was observed across all data with no noticeable outliers. This does not account for differences in the absolute model data (for example, accounting for reduced trip rates), but demonstrates a similar distribution of people and trips.

4.3.4. Furthermore, a comparison in trip growth was made between the two datasets. To isolate changes due to trip rates, growth between 2016 and 2036 was considered. Whilst a general linear trend is apparent, there is a high degree of variation. Crucially, however, changes in growth in trips are well explained by changes to input data and intermediate model outputs, including population and car ownership.

4.3.5. Given the importance for users, segmented growth rates for AM peak car driver trips were also considered. Again, differences are explained by a change to input data and intermediate model outputs. Where particularly large changes to input data are seen, source inputs for NTEM 7.2 have been checked to ensure validity.

4.3.6. At the zone level, growth in trip ends was also considered. Both datasets exhibit a normal distribution, with the majority of model zones displaying growth within one standard deviation of the mean. It is also evident that, despite the higher number of model zones, there are less instances of ‘extreme’ growth in the NTEM 7.2 dataset.

[13] R^2 in excess of 0.96 for all data (five data types across three model years)

Table 4-2: NTEM 6.2 comparison

Check	Details	Status
Compare absolute planning data for a forecast year (2036)	There should be a good level of agreement around the distribution of people, households and jobs at the district level within a region; where this is not achieved, the reasons should be identified and documented.	Completed No issues
Modelled year comparison (2011, 2016, 2036)	Compare output population, households, jobs, workers and total trip ends between NTEM 6.2 and NTEM 7. A reasonable correlation is expected. Major discrepancies should be identified through scrutiny of the data and explained.	Completed No issues
Compare growth in total trip ends, planning data and car ownership at a district level (2016 – 2036)	The pattern of district-level growth is generally expected to be similar between NTEM 6.2 and NTEM 7. If this is not the case (for example, if a district with relatively low growth in NTEM 7 had relatively high growth in NTEM 6.2), the reasons for this should be identified (i.e. large differences in input data).	Completed No issues
Segmented growth rates – weekday car driver AM trips (2016 – 2036)	Recognising this as a primary output for users – differences in growth rates should be explained (for example, through trends in car ownership, changing employment rate, or greater population) at the district level.	Completed No issues
Maximum and minimum zone growth – total productions and total attractions (2016 – 2036)	Distribution of growth rates by model zone should be compared between NTEM 6.2 and NTEM 7. The expectation is that the general pattern of growth should be similar and normally distributed (some differences are expected due to the large additional number of zones).	Completed No issues

5. Outcomes

5.1.1. A quality assurance process has been defined and agreed with regards to model inputs, model processes and model outputs relating to the NTEM dataset. This procedure was applied to the NTEM 7.2 dataset, with results reviewed.

5.1.2. Model inputs were quality assured through independent review, sense checks and comparison checks to raw data and NTEM 6.2 inputs.

5.1.3. Model processes were quality assured for both software and model set up. NTEM 6.2 results were reproduced with updated model software to ensure no errors were introduced. Model runs were independently scrutinised to ensure runs were set up correctly.

5.1.4. Model outputs were quality assured through detailed sense checks and comparison with NTEM 6.2, with model mechanisms confirmed and outliers explained. This has led to additional scrutiny of model input data for particular outliers, and therefore provides confidence in model inputs and parameters.

5.1.5. Based on the quality assurance carried out, and of the review of the evidence of this quality assurance, no issues have been found which cause us to question the reliability of the dataset.

© Atkins Ltd except where stated otherwise.

The Atkins logo, ‘Carbon Critical Design’ and the strapline
‘Plan Design Enable’ are trademarks of Atkins Ltd.