



Department
for Transport

Quality Assurance of Analytical Modelling

April 2013 (Updated September 2014)



Department
for Transport

Contents

| | |
|---|----|
| Foreword | 3 |
| Introduction..... | 4 |
| 1. Principles of Analytical Model Quality Assurance in DfT | 5 |
| 2. The DfT Model Quality Assurance Framework..... | 6 |
| The Model Control Environment | 7 |
| Model Accuracy and Reliability | 7 |
| Governance and Transparency | 7 |
| 3. DfT Quality Assurance Processes..... | 9 |
| Overall Approach | 9 |
| Processes for the Model Control Environment..... | 10 |
| Processes for Model Accuracy and Reliability - Model in Development | 13 |
| Processes for Model Accuracy and Reliability - Model In Use | 16 |
| Processes for Governance and Transparency..... | 18 |
| 4. Detailed Process Explanation and Guidance | 20 |
| Model Control Environment Detailed Process Guidance | 20 |
| Model Accuracy and Reliability in Development - Detailed Process Guidance | 23 |
| Model Accuracy and Reliability in Use - Detailed Process Guidance | 26 |
| Model Governance and Transparency - Detailed Process Guidance | 27 |
| Annex A: Glossary of Terms..... | 29 |
| Annex B: Example Model Governance Checklist | 30 |
| Annex C: Model Impact Matrix | 31 |
| Annex D: Model Complexity Matrix | 32 |

Foreword

The Department for Transport has a strong track record in placing analysis at the heart of decision-making. That record is something to be valued and, as Chief Economist, I want to ensure that the analytical community continues to remain at the forefront of policy delivery, helping the Department tackle the wide range of complex and fast-moving challenges that it faces. In order for that to happen however, our analysis needs to remain credible to those that rely on it and to stand up to the highest levels of external scrutiny. To support that aim, I am introducing this guidance for the quality assurance of analytical models in DfT.

Quality assurance should be embedded into the way that we approach, deliver and communicate analytical modelling. This is not the same however, as creating unnecessary delay and disproportionate process around the delivery of advice. Instead it is about creating the environment in which high levels of quality are demanded, valued and delivered efficiently. In conjunction with the analytical assurance framework, this guidance helps to create that environment and provides you with processes for delivering it. It will be for analysts and their policy partners to ensure that it happens and to take up the challenge of ensuring that analysis in DfT continues to enjoy the well-founded reputation for excellence that it has built up over the years.

Tracey Waltho

Director of Analysis and Strategy, Chief Economist

Introduction

1. This document is intended as high level guidance for the appropriate quality assurance and audit regime around analytical models in DfT. The consideration of an appropriate quality assurance approach is mandatory for all analysts that own and use models and their policy partners that rely on their outputs to aid decision making. This guidance is intended to help fulfil that duty.
2. This guidance forms a core component of the DfT Analytical Assurance Framework and will interact with the analytical strategies and governance arrangements that are mandated within that framework. The application of these guidelines will form part of the standards against which audits will be completed.
3. The overarching principles that can be found at the start of the documents make clear that the approach to quality assurance and auditing of analytical models, should be proportionate to the risk and impact of an error occurring. Annexes C and D of this document provide explicit guidance on judging the proportionate approach required. These matrices should be used to decide the level of QA appropriate to the analysis.
4. As these guidelines are intended to be applied across the full range of DfT business, there is necessarily a degree of flexibility built into the exact processes and approach that should be taken. However, basic measures to ensure that the environment within which a model is developed, used and stored are secure, should be routinely applied. The degree to which model owners then achieve higher levels of assurance around model accuracy and reliability should form part of the governance process and discussions.
5. Finally, the processes described within this note are not intended to be exhaustive. It is recognised that that specific professions or DfT areas may well apply additional quality assurance processes. Where that is the case, it is still expected that these processes will still be part of a QA regime that satisfies the overarching principles and approach that is provided by this document.
6. Where an analytical model is defined as 'business critical', it triggers a wider set of responsibilities and accountabilities. These are established in detail in the *DfT Analytical Assurance Framework*.

1. Principles of Analytical Model Quality Assurance in DfT

1.1 Model quality assurance is not an exact science but there are a number of guiding principles that should be considered when deciding the appropriate level of assurance. These principles are established below.

i) A wide range of Quality Assurance processes can be deployed and they should go wider than merely checking that a model calculates without errors.

1.2 Quality assurance involves ensuring that the model is securely managed, stored and operated, that the model performs accurately and reliably and that appropriate governance is in place and the model is suitably transparent.

ii) The level of assurance required should increase with the assessed impact of the model and consequently of a model error

1.3 In order to insert a degree of objectivity into the assessment of model impact, analysts and their policy partners can use the DfT Model Impact Matrix to generate an impact score for their modelling. These scores can be compared across a range of DfT models to test where a particular model sits on the impact spectrum. **See Annex C.**

iii) The more complex or innovative the approach, the higher the risk of model error, and therefore the greater the level of QA that should be undertaken.

1.4 Commensurately, the more complex a model is, the more QA (and thus time/resource) is likely to be required to achieve a given level of assurance. To provide a measure of complexity, DfT has a high level complexity matrix for assessment against. **See Annex D.**

iv) The approach to quality assurance will differ between a model in development and a model in use and should be explicitly considered.

1.5 While DfT develops some models as a one-off, several of our most important models are developed in stages, and then used many times over. Our QA regime around model development will therefore focus much more on ensuring the model framework is reasonable and performs accurately. In use, it will be ensuring that the inputs are correct

and the results reliable and the limitations and uncertainty around the results is effectively communicated.

v) Where model development and use is outsourced, the DfT model owner and policy customer should ensure that the appropriate QA regime is in place and record that fact.

1.6 DfT have a number of analytical processes which rely on the modelling of other parties to inform decisions. Where this is the case, DfT should satisfy itself that the information upon which it is making decisions has been subject to appropriate QA.

vi) Quality Assurance should be embedded within the governance structure of analytical models and/or the governance structure of a policy project.

1.7 The trade-off between time, quality and resources that is an integral part of any QA regime is a question of governance and what tolerance of risk is deemed acceptable. This trade-off should be made explicit and signed off as part of the governance process. For more information on the overarching framework for analytical assurance and governance in DfT, analysts and their policy partners should refer to the DfT Analytical Assurance framework.

vii) If an analytical model is defined as business critical, a specific set of roles, responsibilities and accountabilities are triggered.

1.8 The Macpherson Review of quality assurance of analytical modelling in government provided the following factors for determining whether a model is business critical:

- the modelling drives essential financial and funding decisions;
- the model is essential to the achievement of business plan actions and priorities; and/or;
- errors could engender serious financial, legal, reputational damages or penalties

1.9 If a model in development or use appears to meet these criteria, you should refer to section 3 of the DfT Analytical Assurance Framework for the process to follow.

2. The DfT Model Quality Assurance Framework

- 2.1** The DfT quality assurance framework requires those devising a QA regime to think in three strands. Model owners are expected to reach high standards for the first strand, whereas the second and third strands are likely to need a risk-based judgement agreed between model owner and policy partner.

The Model Control Environment

- 2.2** The first strand of model quality assurance; the model control environment relates to the conditions in which the model is developed, stored and used and the controls in place to ensure that the risks of malicious and accidental changes are minimised and the model functions as intended. There are a range of processes that can be applied to satisfy this strand and high standards should be routinely achieved.

Model Accuracy and Reliability

- 2.3** The second strand of model quality assurance, involves ensuring that the model results are accurate and reliable. Here we define:
- Accuracy as ensuring that the model calculates as intended and that those calculations are free from error.
 - Reliability as ensuring that the approach is reasonable and conforms with accepted practice and that the outputs intuitively make sense and fit to real-world data.
- 2.4** This is the area of quality assurance where there is the scope to consume the most resource and correspondingly to achieve higher levels of assurance. For example, an independent review of the model can be performed in a light touch way by a colleague or performed line-by-line by an external body. The judgement as to the appropriate level for these activities is likely to be the focus of discussion in the governance process and should deliver a proportionate level of assurance in line with the impact of an error and the complexity of the approach. There should be a clear distinction as to whether the model is being developed or is in use.

Governance and Transparency

- 2.5** The final phase of the DfT framework, relates to how the model is governed. Governance defines the set of processes, roles and responsibilities, which ensure; an effective environment for QA exists; that model risks, limitations and major assumptions are understood by the users of the model; and the use of the model output is appropriate.
- 2.6** Closely intertwined with the issue of governance, is the question of transparency. Making the results and methods of modelling as transparent as possible can be an effective mechanism for driving up standards of quality and the default should be to encourage transparency. However, it is accepted that there are a number of reasons why transparency may not be appropriate, such as commercial sensitivity. Where this is the case, it is sufficient to indicate whether transparency has been considered and why it is not felt appropriate.
- 2.7** The following section provides tables that summarise the processes that form part of each of the three stages along with an assessment of the risks they are attempting to mitigate. These are designed to facilitate discussions in deciding the appropriate level of QA to be used. There then follows a reference section that draws out in greater detail what is involved with each of the processes and provides some examples as to how they may be applied.

Proportionality: deciding the appropriate level of QA

- 2.8** The level of QA required is to be decided in the early discussions between analyst and policy/operational/delivery leads. This varies by model and depends on several factors detailed in Annexes C and D. The overall principle is that the greater the model's potential impact, and the more complex the model, the higher the level of QA required and the more time and resource for QA should be included in planning the analysis.

3. DfT Quality Assurance Processes

Overall Approach

- 3.1** This section of the QA guidelines presents a number of tables containing QA processes in each of the 3 strands of our assurance framework, compared against some the risks that they are designed to mitigate. These risks vary across the three strands although some risks appear in more than one strand.
- 3.2** The tables are designed to be used to aid discussion. By presenting the information in this way, it is intended that model owners and policy partners will be able to understand what QA process are available to mitigate a range of risks and design their QA approach accordingly.
- 3.3** When designing the appropriate QA approach, it is anticipated that high standards will be achieved in the Model Control Environment will routinely be achieved. The Model Accuracy and Reliability and Governance and Transparency strands are likely to be where greater variation occurs, in line with model impact and complexity.

Processes for the Model Control Environment

| Process | Sub-Process | Risk Being Mitigated | | | | | | |
|-----------------|--|--|--|---------------------------------------|----------------------------|--|---|---|
| | | Accidental, malicious or unilateral changes to Model | Model does not contain all changes it should | Errors in Model inputs and parameters | Errors in use of the model | Resilience of DfT and Modelling Capability | Inability to account for previous results | Modelling does not follow best practice |
| Access Control | Access restricted to a need to know basis | ✓ | ✓ | | | | | |
| | Population with access to models broadly represents users | ✓ | ✓ | | | | | |
| Change Control | Changes are subject to a proportionate approvals process before they are made. | ✓ | ✓ | | | | | |
| | Controls in place to prevent unauthorised or accidental changes | ✓ | ✓ | | | | ✓ | |
| Version Control | Keeping a control log of versions and changes made. | ✓ | ✓ | | | | ✓ | |
| | Naming conventions and version numbers | ✓ | ✓ | | | | ✓ | |

| Process | Sub-Process | Risk Being Mitigated | | | | | | |
|------------------------------------|--|--|--|---------------------------------------|----------------------------|--|---|---|
| | | Accidental, malicious or unilateral changes to Model | Model does not contain all changes it should | Errors in Model inputs and parameters | Errors in use of the model | Resilience of DfT and Modelling Capability | Inability to account for previous results | Modelling does not follow best practice |
| Back-Up and Recovery | Models are located on IT approved infrastructure with back-up and recovery processes | | | | | ✓ | ✓ | |
| | Previous operational versions of the model are kept in restricted areas to prevent changes to historic records | | | | | ✓ | ✓ | |
| Single Person Dependency | Consideration of whether more than one person should be skilled in the development and use of the model. | | | | ✓ | ✓ | ✓ | |
| User Guide and Succession Planning | Existence of a user guide for the model | | | | ✓ | ✓ | | |
| | Consider succession planning | | | | ✓ | ✓ | ✓ | |

| Process | Sub-Process | Risk Being Mitigated | | | | | | |
|-------------------------|--|--|--|---------------------------------------|----------------------------|--|---|---|
| | | Accidental, malicious or unilateral changes to Model | Model does not contain all changes it should | Errors in Model inputs and parameters | Errors in use of the model | Resilience of DfT and Modelling Capability | Inability to account for previous results | Modelling does not follow best practice |
| Documentation Standards | Data, methods, assumptions and parameters in the model are documented. | | ✓ | ✓ | ✓ | ✓ | | |
| | Model developers and users are trained in the modelling tools, techniques and controls | | | | ✓ | | | ✓ |
| Skills and Experience | Understanding of how modelling suite fits together | | | | ✓ | | | |
| | Previous operational versions of the model are kept in restricted areas to prevent changes to historic records | | | | | ✓ | ✓ | |

Processes for Model Accuracy and Reliability - Model in Development

| Process | Sub-Process | Risk Being Mitigated | | | | | | |
|---|--|---------------------------------------|----------------------------|---|---|--|---|--|
| | | Errors in Model inputs and parameters | Errors in use of the model | Inability to account for previous results | Modelling does not follow best practice | Errors in the way the model calculates | Undue precision attributed to model outputs | Model lacks credibility or cannot be explained |
| Developed in line with model life-cycle | All stages of the model life-cycle are considered and appropriate time is given to each stage: model specification, build and test | ✓ | ✓ | | ✓ | | | ✓ |
| | Consideration of alternative approaches | ✓ | ✓ | | ✓ | | | ✓ |
| Input Validation | Measures to check accuracy and reliability of input data | ✓ | | ✓ | | | | |
| | Log of all inputs and sources | ✓ | | ✓ | | | | |
| | Inputs and assumptions are signed off | ✓ | | ✓ | | | | |
| Developer Testing | Review of all unique formulae or use of audit software | ✓ | ✓ | | | ✓ | | |
| | A walkthrough of the model – checking and testing of code, formulae and linkages | ✓ | ✓ | | | ✓ | | |

| Process | Sub-Process | Risk Being Mitigated | | | | | | |
|--|---|---------------------------------------|----------------------------|---|---|--|---|--|
| | | Errors in Model inputs and parameters | Errors in use of the model | Inability to account for previous results | Modelling does not follow best practice | Errors in the way the model calculates | Undue precision attributed to model outputs | Model lacks credibility or cannot be explained |
| Developer Testing (contd) | A detailed review of model logic eg. checking the actual flow of data through the model against a stylised example of how data is perceived to be flowing through the model | ✓ | ✓ | | | ✓ | | |
| | Cross-check of model outputs against an alternative set of data or information | ✓ | | ✓ | | ✓ | ✓ | ✓ |
| | Parallel Model Build | ✓ | ✓ | | ✓ | ✓ | | ✓ |
| | Parameters in the model are fitted to real-world data | ✓ | | | | | | ✓ |
| | Model forecasts are checked against observed information | ✓ | | | | | | ✓ |
| Communication of Model Limitations and Uncertainty | Sensitivity testing of key model assumptions | | | ✓ | | | ✓ | ✓ |
| | Scenario testing of a group of model assumptions | | | ✓ | | | ✓ | ✓ |

| Process | Sub-Process | Risk Being Mitigated | | | | | | |
|--|--|---------------------------------------|----------------------------|---|---|--|---|--|
| | | Errors in Model inputs and parameters | Errors in use of the model | Inability to account for previous results | Modelling does not follow best practice | Errors in the way the model calculates | Undue precision attributed to model outputs | Model lacks credibility or cannot be explained |
| Communication of Model Limitations and Uncertainty (contd) | Communication of limitations | | | | | | ✓ | ✓ |
| | Communication of model uncertainty | | | | | | ✓ | ✓ |
| Independent Review | Review of model development or results by someone other than the developer – could take the form of any of the ‘developer tests’ or another agreed method. | ✓ | ✓ | ✓ | ✓ | ✓ | | |
| | High level sense-check | | | | | | ✓ | ✓ |

Processes for Model Accuracy and Reliability - Model in Use

| Process | Sub-Process | Risk Being Mitigated | | | | | | |
|-----------------------|--|---------------------------------------|----------------------------|---|---|--|---|--|
| | | Errors in Model inputs and parameters | Errors in use of the model | Inability to account for previous results | Modelling does not follow best practice | Errors in the way the model calculates | Undue precision attributed to model outputs | Model lacks credibility or cannot be explained |
| Input Validation | Measures to check accuracy and reliability of input data | ✓ | | ✓ | | | | |
| | Log of all inputs and sources | ✓ | | ✓ | | | | |
| | Inputs and assumptions are signed off | ✓ | | ✓ | | | | |
| Testing of Model Runs | Review of any changes since last use, for example compare outputs to expected size or sign | | | ✓ | | | | ✓ |
| | Parallel Model Run | ✓ | ✓ | ✓ | | ✓ | | ✓ |
| | Cross-check of model outputs against an alternative set of data or information | ✓ | ✓ | ✓ | | ✓ | | ✓ |
| | Use of error traps or diagnostics | ✓ | ✓ | ✓ | | ✓ | | ✓ |
| | Test convergence of iterating models | ✓ | ✓ | | | ✓ | | ✓ |

| Process | Sub-Process | Risk Being Mitigated | | | | | | |
|---|--|---------------------------------------|----------------------------|---|---|--|---|--|
| | | Errors in Model inputs and parameters | Errors in use of the model | Inability to account for previous results | Modelling does not follow best practice | Errors in the way the model calculates | Undue precision attributed to model outputs | Model lacks credibility or cannot be explained |
| Testing of Model Runs (Continued) | Review of model parameters to ensure they remain fit for purpose | ✓ | | | | | | ✓ |
| | Model forecasts are checked against observed information | ✓ | | | | | | ✓ |
| | Automation of model run process | | ✓ | | | ✓ | | |
| Communication of Model Limits and Uncertainty | Sensitivity testing of key model assumptions | | | ✓ | | | ✓ | ✓ |
| | Scenario testing of a group of model assumptions | | | ✓ | | | ✓ | ✓ |
| | Communication of limitations | | | | | | ✓ | ✓ |
| | Communication of model uncertainty | | | | | | ✓ | ✓ |
| Independent Review | Review of model results by someone other than the user – could take the form of any of the ‘developer tests’ or another agreed method. | ✓ | ✓ | ✓ | ✓ | ✓ | | |
| | High level sense-check | | | | | | ✓ | ✓ |

Processes for Governance and Transparency

| Process | Sub-Process | Risk Being Mitigated | | | | | | |
|------------|---|---|---|---|--|------------------------------------|--|--|
| | | Inability to account for previous results | Modelling does not follow best practice | Undue precision attributed to model outputs | Model lacks credibility or cannot be explained | Model fails to satisfy requirement | Model results are not used appropriately | Modelling is not accepted by external stakeholders |
| Governance | Shared understanding of modelling requirements between analyst and policy partner | | | ✓ | ✓ | ✓ | ✓ | |
| | Procedures in place for the flow of information | | | | ✓ | | ✓ | |
| | Clear process for the internal challenge of results | ✓ | | | ✓ | | ✓ | |
| | Clearance of results from the SMO | | | | ✓ | | ✓ | |
| | Uncertainty in the modelling output is conveyed to decision-makers | | | ✓ | ✓ | | ✓ | |
| | Fitness for purpose of model under periodic review | | ✓ | | | ✓ | | ✓ |

| Process | Sub-Process | Risk Being Mitigated | | | | | | |
|--------------|--|---|---|---|--|------------------------------------|--|--|
| | | Inability to account for previous results | Modelling does not follow best practice | Undue precision attributed to model outputs | Model lacks credibility or cannot be explained | Model fails to satisfy requirement | Model results are not used appropriately | Modelling is isolated from external scrutiny |
| Transparency | Results are in the public domain | | | | | | | ✓ |
| | Model exposed to external challenge through stakeholder engagement, report publication, planning inquiry system, judicial review | | | | | | | ✓ |
| | Methods are in the public domain | | | | | | | ✓ |
| | Useable model is in the public domain | | | | | | | ✓ |
| | Model not in public domain but perform model runs for stakeholders. | | | | ✓ | | | ✓ |

4. Detailed Process Explanation and Guidance

- 4.1 The information contained within this section is intended to provide a little more clarity around the specific processes that can be adopted. It is not intended to be exhaustive or limiting, as specific areas and model types are likely to have a greater depth QA processes than can be captured in this general guidance.
- 4.2 Where individuals or business areas have examples of specific processes, some useful resource or experience of application, we can adapt the guidance to capture that, with the intention of spreading that knowledge and good practice across DfT.

Model Control Environment Detailed Process Guidance

Access Control

- 4.3 Access control is a process that should be relatively easy to design and is aimed at ensuring that changes are not maliciously or accidentally by external parties. It may be enacted by placing the model within a specific area of the corporate file plan, which only the immediate team can access. Alternatively, it may involve password protection and limiting the distribution of that password to a list of users that is logged and regularly reviewed.

Change Control

- 4.4 Changes made to the model should be subject to a proportionate approvals process before they are made. This process should be clear to all those using or developing model. It may be that relatively minor changes such as the updating of a data input can be delegated down, but the change should be recorded.
- 4.5 It is also important to consider whether controls are needed to prevent changes to the structure of the model or its inputs by, for example, the use of a read-only version.

Version Control

- 4.6** Version control is a simple process to embed within the development and use of a model and is fundamental to ensuring that the version in question is the one that it is believed to be. For example, a document may be circulated for comment and involve multiple comments made by a range of people simultaneously. Version control of a models acts to ensure that a record is kept of all 'comments' that have been made on the model. It may also be the case that there are multiple versions of a model performing slightly different functions to different policy areas. Version control is vital to ensuring such situations are documented.
- 4.7** The approach to version control can be relatively informal, perhaps through the use of consistent date stamping on file names. Where models are important or complex, with lots of inter-connected files and programs, it is advisable to keep a log of all versions, complete with the date, name and time that a new version was created. That log should record what is different about the new version and perhaps how this has altered the main output.
- 4.8** Whether formal or informal, there should be a standard approach to version control that is consistently understood and applied by all model users. The log of versions may be embedded within a spreadsheet, as a cover sheet, or it can be in a separate document or template.

Back-Up and Recovery

- 4.9** To ensure that DfT is resilient to IT or infrastructure failure, all models should be able to be restored if the primary version of the model or access to IT system on which the primary version of the model is stored is disrupted.
- 4.10** To enable a proper audit trail of past decision-making, consideration should also be given as to the storage of previously live version of the models, upon which decisions were made. It also helps if subsequent errors are found, in that analysis can be completed on previous model version to understand the size of any error. Storage of previous 'live' versions of a model will also aid any future ex post evaluation.

Single Person Dependency

- 4.11** Where the development, use or interpretation of a model is reliant on a single-person, this imports a significant amount of risk into the use of analysis and decision-making based on that modelling. This may be an acceptable risk where the modelling is straightforward or of limited impact. Where it is more complex and high impact, then it is likely that the risk will need to be mitigated. One element of the risk may relate to the availability of the single person upon which the model relies and consideration should be given to building capability amongst a wide enough number of staff to ensure the Department has contingency plans in place.
- 4.12** The second risk however, is that the use of a single developer or user, through no fault of their own, may increase that chances that errors in the

model or model approach go unchallenged. This second risk, may point towards the need for higher levels of independent review when the model has been developed or is used by a single person.

User Guide and Succession Planning

- 4.13** Each DfT model should contain either a user guide or instructions within the interface that, as a minimum, enable an 'intelligent new user' to produce new model runs without any further input. More comprehensive user guides can prove resource intensive, and the precise format and content should take into account how frequently the model may be used in the future, how complex the model is to use and how wide the scope for error is in the operation or interpretation of the model. Where the model has paid for external contractors to develop a model, a comprehensive user guide should be seen as a way to protect that investment.
- 4.14** Strong QA practice would also include succession planning, where someone with the right skills and experience to take over the primary model user role is identified. Ideally this would encompass a period of handover to properly facilitate the transfer of the modelling function and capability.

Documentation Standards

- 4.15** Here we define documentation standards as the documentation of methods, their source and the reasoning for them. This should be a reference source, to allow new developers to understand the model's construction, and enable them to undertake new development work or create a new version.
- 4.16** All users and developers should have an understanding of what is expected. This may include formal documentation of elements such as model procedures, controls, structure and use. There should also be a standard approach to recording all data sources and assuring that model parameters and assumptions are transparent and where possible, evidence-based.

Skills and Experience

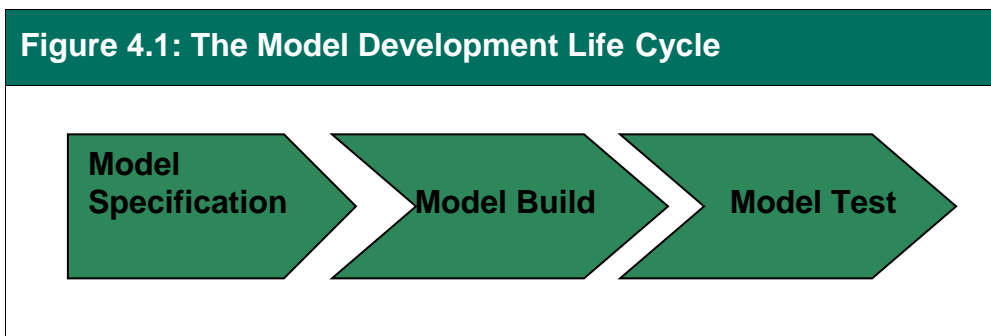
- 4.17** Model developers and users should have the right level of skill and experience in developing and using the model or model techniques. For complex models it may be appropriate for the department to ensure sufficient skills to operate and update a model but use external expertise for development.
- 4.18** If the appropriate capability does not exist within DfT it may be necessary to build the capability through the use of a training and development plan that formally considers the appropriate skill mix and sets out how any gaps are to be filled.

- 4.19** Where there are multiple models interacting, there should be at least one member of staff that understands how they all fit together and has oversight of their interaction.

Model Accuracy and Reliability in Development - Detailed Process Guidance

Developed in line with Model Life Cycle

- 4.20** The development of an analytical model can be thought of as a three stage life-cycle:



Model Specification

- 4.21** The life cycle should start with a common understanding of what the model is needed for; what questions the model will answer; what the risks and limitations around the different possible approaches are; what the timescales are and what complexity and quality is required. The model impact and complexity matrices can provide a useful point of reference for these latter questions.

Model Build

- 4.22** During the model build stage, the model owner or developer should be ensuring that the model is developed or used in line with the agreed specification and that any deviations or problems are communicated.

Model Test

- 4.23** The final stage in the model life-cycle comes when the outputs are tested against requirements to ensure that they answer the question that was set and that any limitations and uncertainty is understood. Mutual sign-off between model owner and policy customer should be facilitated by the SMO.

- 4.24** It is a reasonable rule of thumb that each stage should be given roughly equal weighting in terms of time and resource. The specification of the model will integrate heavily with the governance structure and should form the basis of the mutual sign-off of requirements between the policy partner and the model owner. It will also repeat itself at the final stage as

the product is signed off by model owner and policy partner. This process should be facilitated or overseen by the model SMO.

Input Validation

- 4.25** Input validation provides assurance that the data inputs and assumptions entering into the model and upon which the model calculations are made, are accurate and reliable. By accurate, we mean that the inputs are in the correct format and do not contain any errors. By reliable we mean that they are the correct inputs for the question in hand, for example the correct measure of GDP or inflation.
- 4.26** As well as data inputs, the model is also likely to have parameters. These are structural properties of the model, such as elasticities, might be fixed for many runs of the model. These parameters will also need to be kept under review as part of the input validation process.
- 4.27** Once data and assumptions have been entered into the model, good practice is formally log their sign-off, perhaps in a separate register.

Developer Testing

- 4.28** Developer testing is a catch-all term for a wide range of processes and tests that can be undertaken to ensure that the model is reliable. This element of the QA regime, perhaps interacting with the independent review, is likely to prove the most time and resource heavy part of the process. There are also likely to be a number of different processes that can be applied and the examples given in the framework are not intended to be exhaustive – it will be for individual areas and model owners to identify the processes that are appropriate to their modelling.
- 4.29** Further to identifying the available range of developer tests, it is for model owners and policy partners to agree the appropriate set of tests, addressing the trade-off between time, quality cost and resulting risk that is deemed acceptable to them. This dialogue should be recorded in the model governance.

Communication of Model Limitations and Uncertainty

- 4.30** The results from an analytical model are more often than not, only going to be one part of the overall decision-making process. For decision-makers to gain a richer understanding of how large a role analysis should play, analysts need to communicate the limitations and uncertainty in the modelling. To inform those communications analysts should consider:
- **Sensitivity testing** – how sensitive are the final results to plausible changes in individual assumptions.
 - **Scenario testing** – how sensitive are the final results to plausible changes in a group of assumptions.

- **Uncertainty in the modelling** – the range of statistical uncertainty that exists around a result, perhaps caused by the uncertainty in the forecasts of inputs or model parameters.
- **The limitations of the model** – what are the questions that the model cannot or should not be used to inform the answers to.
- **Risks from the QA approach** - what are remaining risks given the mix of QA that has and has not been conducted, with reference to model impact and complexity?

Independent Review

4.31 Independent review is a strong element of any model QA regime but is also likely to require more time and resource to undertake, the more thoroughly it is completed. The degree of independent review that is desirable should be considered at the start of an analytical project and time and resource built into the project plan to allow for it. As a project evolves, it may be that the modelling turns out to be more risky, either through complexity or profile than anticipated. In such circumstances, it is likely that the degree of independent review that is desirable would increase.

4.32 Irrespective of the type of review that is chosen, it will always make sense to include have someone other than the model developer or user performing a high level sense check of the final results.

4.33 There are numerous ways in which a review can be undertaken and the list below is intended to provide some of these. It is not the case that 'Independent Review' stands as a QA process divorced from the range of processes outlined in the tables in Section 3. In practice, it is likely that an Independent Review will entail conducting a range of the tests and process outlined in that section, independently.

4.34 The exact nature of the review and the checks and assurance that you are seeking to achieve, should be made explicit in a Terms of Reference. Some review types, such as Internal Audit, may have set procedures and tests, other reviews may be slightly more academic in nature, challenging the robustness of the approach. It is likely that the mechanism for ensuring independent review may vary from model to model.

- **Internal Review**

Internal review is usually defined as being undertaken by someone other than the modeller or model developer, but internal to the organisation. At its lowest level, this may range from high level sense checks, to more detailed interrogation of a model.

- **External Review**

External review may of itself be no more stringent than an internal review, but it is defined as being by someone outside of DfT. This could be a professional services firm, another government department or an academic institution.

- **Peer Review**

Peer review tends to be a sub-set of external review, usually associated with a more academic approach to testing the overall quality of the approach taken and whether it fits with accepted best practice. The review tends not to focus heavily on the checking of calculations, but may, for example, suggest whether the overall approach to forecasting demand for travel or calculating wider economic impacts is valid.

- **Internal Model Audit**

Internal audit is likely to perform a different function to a review and will involve the use of a professional audit firm or DfT's own internal audit team. Internal audit may focus more on the risks in the process for undertaking the modelling and test the controls that are in place to mitigate those risks.

- **External Model Audit**

External audit will involve the use of a professional external audit function. An external audit is likely to be a comprehensive test of how the model functions, with the aim of signing off the final set of results as being accurate.

Model Accuracy and Reliability in Use - Detailed Process Guidance

Input Validation

4.35 Similar to input validation for model in development, but will be focused on ensuring the input data and parameters remain valid.

Testing of Model Runs

4.36 Developer testing is a catch-all terms for a wide range of processes and tests that can be undertaken to ensure that the model is reliable. This element of the QA regime, perhaps interacting with the independent review, is likely to prove the most time and resource heavy part of the QA process. There are also likely to be a number of different processes that can be applied and the examples given in the framework are not intended to be exhaustive – it will be for individual policy partners and model owners to identify the processes that are appropriate to their modelling.

Communication of Model Limitations and Uncertainty

4.37 Similar to the processes used for model development, but as the model is likely to have been used in the past, it might be possible to glean more information about uncertainty by comparing past results with actual outcomes. The use of benchmarking against past outcomes and reviewing past model performance should help explore these questions in more depth.

Independent Review

- 4.38** Again, the main types of independent review will be the same for a model in use as for a model in development. However, if a model has been extensively tested during development it may be that stringent reviews only take place when significant changes are made to the model. Where that is not the case, independent review is likely to focus on the interpretation of results, understanding why they have changed from one run to the next and ensuring that the approach is suitable for the question being asked.
- 4.39** If the model forms part of a review or repeated publication, consideration should be given as to the appropriate amount of review that should be undertaken.

Model Governance and Transparency - Detailed Process Guidance

Governance

- 4.40** Governance is perhaps the most important element of the QA framework because it is the process by which risk is managed and the appropriate QA regime signed off. Appointing a Senior Model Owner for each model helps to drive leadership and should form part of the governance framework. The appropriate level for the SMO will be determined by the impact of the model, but for business critical models, it is unlikely that the SMO will be below Grade 5.
- 4.41** Thinking about the three stage lifecycle of model development, governance should play an important role in all three:

Model Specification

- Provide the interface between the requirements of those commissioning the model and the specification and results that the model is designed to produce.
- Establishes the appropriate QA regime using the model impact matrix and the degree of model complexity.
- Identifies resources and timescales for delivery.

Model Build

- Monitor progress towards timescales
- Provide a forum through which risks to delivery can be raised and mitigated.
- Consider whether the QA regime remains appropriate.

Model Testing

- Provide challenge to final outputs.
- Establish that the QA regime was followed.
- Communicate risk and uncertainty to decision-makers.
- Sign-off the completion of the modelling.

4.42 When considering governance around the ongoing use of a model, it can follow a similar process but the first stage will be to ensure that the modelling approach still meets requirements and is fit for purpose. The model build stage will relate to the running and use of the model and the model testing stage will relate to signing of the final results.

4.43 It is the case that when analytical models are performing an integral role in the development and delivery of policy, the analytical governance structures should be embedded within that of the overarching policy.

4.44 **Annex B** contains a checklist for analytical model governance.

Transparency

4.45 Transparency is a powerful tool for improving the quality of analytical work, firstly because it encourages those producing something for public consumption to undertake more rigorous checks but secondly, because it opens up the approach and results to a wide range of external experts and can spark academic challenge, debate and research.

4.46 There are a number of mechanisms for exposing analytical modelling and its results to a wider audience. The modelling can form part of a full publication in its own right, such as the as the DfT Aviation Forecasts and the Road Traffic Forecasts publications. These publications may include more detail on model methods and assumptions in comparison to the release of results as part of a policy document. Analysts can actively seek to engage with interested parties through the use of seminars. There may also be occasions where analytical modelling forms the basis for a planning inquiry or legal challenge, both of which generate high levels of external scrutiny.

4.47 Where appropriate, DfT would expect to defer towards making analytical modelling as transparent as possible. There are however, limits to the degree to which transparency is appropriate. Commercial sensitivity is one such limit, but the model owner may also wish to consider the political environment and whether decision-making should entirely be seen through the prism of analysis. At the least, the appropriate degree of transparency should be considered within the governance structure.

Annex A: Glossary of Terms

A.1 This Annex is intended as a reference guide for the terms used in this document. It begins with a definition of a model but then sets out some of the terms around that basic definition.

Definition of a model

A.2 In the broadest sense, a model is a mechanism for analysing or investigating some aspect of the real world. It is usually a quantitative method, system or approach which applies statistical, economic, financial, or mathematical theories, techniques and assumptions to process input data into quantitative estimates. There are typically three parts to a model:

- inputs – in the form of data and assumptions;
- a processing component – often through calculations; and
- outputs.

Other Terms

A.3 **Model Parameters:** structural feature of the model that defines the relationship between one variable and another. For example a parameter may be an elasticity that relates the growth in GDP per capita to the growth in car ownership

A.4 **Model Assumption:** similar to a model parameter but perhaps harder to generate through a quantitative test for example a behavioural response to a new form of transport.

A.5 **Accuracy:** ensuring that a model or input calculates as intended and that those calculations are free from error.

A.6 **Reliability:** as ensuring that the modelling approach or input is reasonable and conforms with accepted practice and that the outputs intuitively make sense and fit to real-world data.

A.7 **Model Owner:** The lead, working level analyst responsible for the development or use of a model.

A.8 **Policy Partner:** the official responsible for commissioning the analysis and using it within a decision-making process.

A.9 **SMO:** The Senior Model Owner is the lead official with ultimate accountability for the production and use of model outputs. Should provide leadership with regards to QA through the governance structure.

Annex B: Example Model Governance Checklist

B.1 The checklist below, identifies some key questions to consider around the governance of analytical models. The list is not intended to be exhaustive and can be amended in light of the local environment.

Model Specification / Commission

- A suitable SMO for the model is identified.
- Mutual sign-off of the purpose of the model has been achieved and recorded.
- The DfT QA guidelines have been applied and an appropriate QA *strategy* is in place, where the QA strategy should record:
 - Any exceptions or weaknesses in the Model Control Environment.
 - The degree to which inputs and assumptions have been signed off.
 - What developer or user testing are going to be applied and what the residual risk is.
 - What independent review is going to be performed and why this is appropriate.
 - Record any trade-offs between time, quality and cost that have been made and that the resultant risks are owned at the appropriate level.

Model Build

- The model build or run is monitored against agreed timelines.
- Issues and risk are escalated through an agreed process and appropriate mitigation is considered.
- Results, including intermediate results, are released through an agreed process, allowing for sufficient time for high level challenge.

Model Test / Results

The SMO should establish that the final model or model results are:

- Fit for the purpose intended.
- That the risks and limitations in the modelling are properly understood.
- That the QA regime has been properly applied.
- The SMO should record final sign-off of the model build, or results and this should be agreed by the policy partner.

Annex C: Model Impact Matrix

The model impact matrix is used to decide the appropriate level of QA to be applied to the analysis: the greater the impact of a model, the more QA (and more time and resource for QA) is required. This forms part of the early discussion between analyst and policy leads before analysis starts.

Table C.1 Model Impact Matrix

| Description | Low (1) | Medium (2) | High (3) |
|---|--|---|---|
| Role of analysis in the final decision | Model or analysis is one of many factors determining the decision and is not critical. | Model or analysis is one of the most important factors in the decision-making process | The model output is the most important factor in the decision-making process |
| Size of the financial allocation to which the model relates | Modelling is used to allocate DfT or government funding below £50million | Modelling is used to allocate funding >£50million <£250million | Model is used to allocate funding >£250million |
| Size of the wider welfare impacts | Gross welfare impact (positive or negative) <£50 million | Gross welfare impacts (positive or negative) >£50million <£250million | £20billion in wider welfare benefits or costs |
| Legal Consequences of an error | Inappropriate use of modelling is very unlikely to leave the department open to legal challenge. | Inappropriate use of modelling may leave the department open to challenge but either impact of challenge or risk of challenge is low. | Inappropriate use of modelling is almost certain to leave the department open to legal challenge. |
| Level of reputational risk | Model is only used internally | Model output is likely to inform a public statement but not explicitly stated | Model output is quoted publicly by Ministers |
| Model is essential to business plan activities | Model has no relevance to business plan indicators or activities | Model forms part of the reporting for business plan indicator | Model is used to determine success against a business plan objective |
| Number of times model is used | One-off single use | Used for one or two projects annually | Used frequently, for multiple projects |

Annex D: Model Complexity Matrix

The model complexity matrix is used to decide the appropriate level of QA to be applied to the analysis: the more complex a model is, the more QA (and more time and resource for QA) is required. This forms part of the early discussion between analyst and policy leads before analysis starts.

Table D.1 Model Impact Matrix

| Description | Low (1) | Medium (2) | High (3) |
|---------------------------|---|--|---|
| Form of the Model | Simple spreadsheet, perhaps using one or two sheets with limited amount of data transformation | Complicated spreadsheet, perhaps over several worksheets with multiple sources of data and calculations. | Model is an application running through code. |
| Innovation | Model follows well-established techniques with a strong evidence-base | Model follows well established techniques but may involve the generation of new evidence, i.e. econometric modelling or some other forms of innovation | Model is innovative and either applies techniques in a new way or involves developing a new approach. |
| Interactions / Iterations | Modelling requires no iteration or interaction with other models | Modelling requires a limited amount of iterations and interactions with other models but not too complex | Model involves a high degree of iteration and interactions between models, for example supply and demand interactions in transport modelling. |
| Scale / Visibility | Easy to gain oversight of the whole modelling process and limited amount of key information to check and report | Modelling is challenging to oversee, relying on a number of key assumptions and results to gain confidence in the output. | Modelling is very challenging to oversee, relying oversight across a range of models and model interactions. |

