Peer Review of "Peer review of NAPALM" of 28 October 2010 and the Department for Transport Response

1. Introduction

This is a review[1] of the procedures surrounding the commissioning, interpretation and response to an expert peer review of the National Air Passenger Allocation Model (NAPALM[2]). It does not address technical detail in depth, but reviews the underlying procedures and the quality of the expert Peer Review and the Department's Response.

NAPALM, as explained in the expert Peer Review, is a key component of the Department's suite of aviation forecasting and appraisal models. It includes an airport choice model that allocates passenger demand to airports initially without airport capacity constraints; a model that converts passenger flows into air traffic movements (ATMs); and a "capacity restraint" process that generates "shadow costs", based on runway and terminal constraints, with an iterative cycle that derives a total passenger flow allocated to airports allowing for runway and terminal capacity constraints.

The relationship of these processes to other elements of the Department's aviation forecasting processes is shown diagrammatically on page 2 of the expert Peer Review.

2. The expert Peer Review of NAPALM[3]

The Department's commissioning of this peer review of NAPALM did not arise from any doubts about the competence or conscientiousness of either the consultants, Scott Wilson, who maintained and operated the model, or of Departmental officials responsible for this work. It arose rather from an increasing awareness that the complexity of the model had developed over time, to an extent that no more than perhaps one person can have an intimate feel for of all of the modelling itself (and no one can combine a close feel for the modelling and all of the data and the economic rationale underlying all of the associated algebraic inputs). This had already led to some problems where anomalous outputs had proved to be symptoms of flaws that were identified only with difficulty.

The Peer Review therefore addressed two sets of questions. One is "whether the model can be described as fit for purpose in the light of [the time limitation of the current Departmental exercise]". The other is what suggestions might be made for

---

1 This is one of four reviews of procedures produced for publication with the Department's aviation forecasts in July 2011. The other three are of the econometric estimation of price and income elasticities, a paper on the future impacts of market maturity and market liberalisation and a proposed new section on aviation to include in the Department's web based transport appraisal guidance (WebTAG).

2 The Peer Review uses this acronym, while the Department generally uses NAPAlM, with a lower case 'l'. The Department might usefully consider moving away from NAPAlM to the less confusing NAPALM or, perhaps better, simply NAPAM.

3 John Bates Services (2010) "Peer Review of NAPALM", prepared for International Networks Analysis and Support Division, Department for Transport, 28 October, 2010.

"somewhat more radical departures which could be followed in a less constrained timescale".

The expert Peer Review is a well balanced document, of very high technical quality, that covers the ground as comprehensively as could have been hoped for in the time available. It also stresses that the modelling consultants had been extremely cooperative and constructive in answering questions and providing data.

The longest section of the Review is on the (unconstrained) passenger allocation model. This is followed by a short discussion of the conversion from passengers to air traffic movements[4] and a final more substantial section on capacity constraints and shadow costs. Many constructive observations are made throughout the text, several of them being collected in a final conclusions and recommendations section. Some of the specific observations are covered in section 3 below, in the context of the Department's Response to the Review.

In addition to producing his report, the expert peer reviewer also attended two meetings of the Department's aviation Technical Working Group, which included external experts from the CAA and elsewhere, to discuss the findings and recommendations.

3. The Response to the expert Peer Review of NAPALM

The Response to the expert Peer Review is a DfT document drawing heavily on input from their modelling consultants. It addresses twenty one specific comments or suggestions from the Peer Review. The number of issues addressed, within which subsidiary issues are in several cases embedded, reflects the complexity of the modelling. In each case the Response provides some explanatory text followed by a "Summary Response".

I understand that, although during the preparation of the Peer Review there was clearly good, open exchange between the reviewer and the modellers, the written Response itself was not discussed with the expert peer reviewer. However the responses are well informed, comprehensive, well balanced and generally persuasive. I comment here on only a few technical issues and then on more general, high level issues which could usefully be more fully addressed.

One area of concern to the expert peer reviewer is the handling of surface access. This is addressed on pages 9 and 10 of the Peer Review and at some length in the Response, in its paragraphs 2.5 to 2.12. The Response notes (paragraph 2.5) that the Peer Review's main reservation is that the surface access scaling factor (lambda), taken as -0.1 for all modes, is not based on empirical data, and that the Review proposes the use of nested choice models. The expert Peer Review nonetheless

---

[4] A trivial but revealing illustration of how large models develop over time a world of their own is the Review's observation that the so called "Larame" graphs relating numbers of passenger to aircraft size are described in the model documentation as "a well established procedure", implying some external pedigree. In fact they are a perfectly reasonable procedure, but appear to have evolved wholly within the current model and its immediate predecessors (PAM and SPAM) in the Civil Aviation Authority.

judges the *general* methodology as acceptable for its current purposes, but suggests that the procedures could be challenged on several further grounds including the high value of the choice scaling parameter (lambda) at -0.1, the absence of its variation between market segments, the contestability of the valuations of time and the absence of variation of lambda across surface access modes.

The Response addresses these issues reasonably fully and frankly and accepts at least in spirit the Review's various observations and recommendations.[5] It accepts that there is a case for change and declares a longer term aim to introduce a nested hierarchical *domestic* model and the facility to use mode specific constants. It also says that "in the meantime, [the Department] will avoid using the model to perform more detailed assessments of High Speed Rail". The Response suggests that HSR "would have only a very small effect on the overall airport allocation". But it must be hoped that the foreseeable temptations to use the model to give answers to questions which it is at present not really able to handle will be firmly resisted.

Also of substance, but less clear cut, is the concern expressed in the Peer Review about an equal airport capacity constraint shadow price being applied to all passenger segments. The issue is explained clearly and concisely in paragraph 3.12 of the Response and the subsequent discussion is extensive. However the conclusion, in the text of paragraph 3.17 and the subsequent Summary Response, is diffident. The Department maintains that "the current 'neutral' assumption [while stylised] provides a reasonable indication" of how airlines respond to airport capacity restraints. Moreover, as the issue is deeply embedded in the software, the testing of alternative assumptions would be very costly. Indeed I understand that the Department currently feels that should be done only if they were confident, in advance, that it would add significant value to the model outputs as the basis for policy development. The Response does however, in its Summary Response on these issues, undertake to "consider in the next round of software development methods for providing greater flexibility in the application of shadow costs per passenger", although there does not appear to be any mechanism that could effectively ensure that this will happen.

On the question of how airlines actually respond it would seem surprising if this were not amenable to direct research, or to informed judgement by experts closely familiar with the commercial workings of the airline industry. Perhaps that should be investigated.

Of much less substance is the handling of flight frequency. The model's current handling of flight frequency, presumably by accident of history, is based on the procedures used for urban transport, where travellers typically arrive at random at, say, a bus stop and wait for the next bus or tram. As the expert Peer Review notes, and as is accepted in the Response, this framework (even though it is modified in NAPALM) is much less appropriate for long distance transport, and air travel in particular. However the Peer Review establishes that the model in its current form

---

[5] Although, reasonably in my view, the Response does not address the valuation of time. The Department is I believe well aware of the issues and generally makes best use of the available evidence. It must be hoped that in due course more evidence will emerge of on the relative valuations of surface access, airport waiting and flight times (and on the valuation of time uncertainty as distinct from time duration: uncertainty of a few minutes may for example be typically almost costless, whereas the value of time duration may be very close to linear). But time valuation is a transport wide issue and not a priority for the development of NAPALM.

happens to produce outputs very similar to those of a more plausible structure. There is therefore no serious problem here, and the Departmental Response promises, enigmatically but probably sufficiently, the introduction of "a simpler level of service variable".

Less well defined technical issues may arise from the use of the shadow prices generated in the context of airport capacity constraints within cost benefit analysis – as distinct from their use to allocate passengers and air traffic movements. There are aspects of the modelling here that are intuitively quite challenging. The Response discusses the distribution of shadow costs across passenger segments, as noted above. It also briefly addresses (paragraph 3.18) the Peer Review's concern about the potential implications of alternative assumptions to the current convention of treating the shadow cost per passenger as if it is in 'surface access' units. The conclusion here is again that it is an issue to review in the longer term. In itself this is fair. But as elsewhere would be more reassuring were there some mechanism to ensure that this will indeed be picked up rather than left in the long grass.

At a higher level, the expert Peer Review has made even clearer than before how immensely complex is the model and how much it is an inheritance from a long history of incremental development. The Peer Review's "strong plea for more documentation" is addressed in the Response's paragraph 4.1. However the Peer Review concludes that plea with the comment that "These remarks should also be taken to refer to the program code, which I [i.e. the expert peer reviewer] suspect would benefit from a major overhaul before the details become impenetrable". The possible need for "root and branch 're-engineering'" is accepted in the Response (paragraph 3.19) in the context of the algorithm for achieving capacity restraint (with which most of the code is associated) and I understand this applies more generally to all the code.

More generally, the Response categorises activities to handle many of the suggestions in the Peer Review as issues for the medium or longer term. In itself this is perfectly reasonable, but, again, the Response does not indicate any general mechanism to drive such activities.[6] Given that there will nearly always be short term pressures on the modelling there is every likelihood that, without firm commitment to a development plan, with top of the office endorsement, many longer term proposals will fall by the wayside. This is of special concern as the medium to long term suggestions tend to be those that might have most impact on the model outputs. I understand that the Department proposes during 2011 to ensure the updated model is fully documented and to start the process of bringing the bulk of the modelling in-house, and to pick up some of the medium term tasks identified in expert Peer Review. It seems unlikely however that top of the office commitment could be obtained to any more significant programme of development beyond this period

[6] The most that is offered is an explanation that updating of documentation will be included in future contract specifications for model maintenance and development.

4. Conclusion

The Department for Transport deserves to be congratulated for commissioning an independent expert peer review of NAPALM. This has added very considerably to wider understanding of the model's complexity and of ways in which it might usefully be developed in the short and the longer term.

The expert Peer Review itself is well balanced document, of very high technical quality, that covers the ground as comprehensively as could have been hoped for in the time available. It also stresses that the consultants who designed and implement the model, with whom the reviewer dealt for information on the modelling, had been extremely cooperative and constructive in answering questions and providing data.

The Response to the Peer Review is also an impressive document. It is well structured and comprehensive, with responses that are well informed, well balanced and generally persuasive.

Although there was very close exchange between the expert peer reviewer and the modellers during the preparation of the Peer Review the peer reviewer appears not to have been invited to comment on the Department's Response.

The Response notes that the model's use for "detailed assessment of High Speed Rail should only proceed after the types of changes recommended by the Peer Reviewer are implemented". One must hope that this is not subsequently overlooked.

The Response gives a little less emphasis than the expert Peer Review to the case for a major overhaul of the program code, but I understand the case is accepted. More serious is the evident danger that work such as this, that is marked for attention in the longer term, will in the event fall by the wayside as time goes by. I understand that there is a programme in place for development work over the next few months, but that there are no further firm commitments. Without some longer term top of the office commitment one has to fear that practical responses to many of the longer term recommendations may be postponed indefinitely.

Michael Spackman
Special Consultant, NERA Economic Consulting
29 July 2011