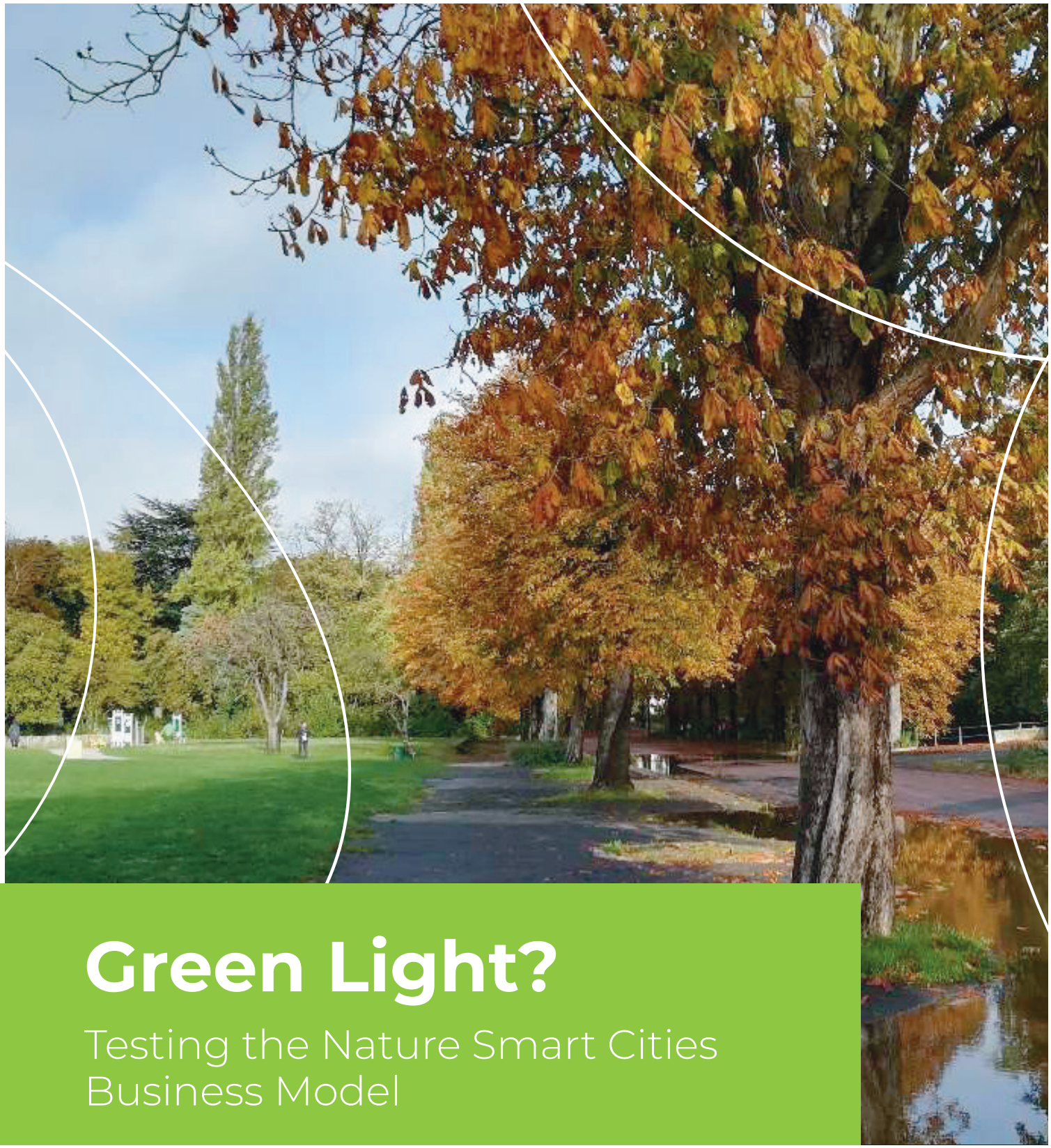

Photo: Phil Back

Green Light?

Testing the Nature Smart Cities Business Model

A report for the Nature Smart Cities Partnership

November 2021

Interreg 2 Seas Mers Zeeën
European Regional Development Fund

NATURE SMART CITIES
ACROSS THE 2 SEAS

Authorship

This report has been prepared for the Nature Smart Cities partnership by Dr Phil Back. Phil is a highly experienced research consultant specialising in local government work, and has considerable experience in the field of testing and observing user interactions with processes and systems. He has worked with local authorities across England and Wales and in a wide range of different settings and services. His first degree, in modern history, was completed at the University of Stirling in 1974, and he has since completed an Advanced Diploma in Historic Environment at the University of Cambridge and a PhD in Landscape History at the University of Sheffield, with a thesis on the origins of Scotland's country park provision.

The Nature Smart Cities project

The Nature Smart Cities (NSCiti2S) project brings together 11 partners across seven cities in France, Belgium, the Netherlands and the UK, funded by the EU through its Interreg 2 Seas programme. The project aims to strengthen and develop LA capacity to provide GI, particularly in smaller, less well-resourced municipalities. It will do this by providing a business model with a robust step-by-step methodology, to determine the financial value of the outputs of a GI project, enabling a more accurate calculation of GI's cost and benefit in comparison with more traditional approaches.

To find out more about NSCiti2S visit www.naturesmartcities.eu

Nature Smart Cities has received funding from the European Union's Interreg 2 Seas Programme under grant agreement No. 2S05-048.

Executive Summary and recommendations

1. This report summarises the results of a series of eight 'Demonstrator Tests' carried out in the autumn of 2021 to prove the effective working of the Nature Smart Cities (NSCiti2S) Business Model and to identify areas where modification might be needed. Green Infrastructure (GI) projects for testing were recruited using an open call and extensive publicity in newsletters and other information mechanisms.

2. The tests were carried out in the UK, the Netherlands and Belgium, and covered a range of green infrastructure project types and scales. Five tests were carried out face-to-face, and three using MS Teams; they lasted from 1.3 to 3.5 hours. Participant authorities have been anonymised in this report, and the detail of their projects and their feedback comments has been removed to protect confidentiality.

3. Detailed notes were taken at each test, and these form the basis of the findings in this report. In addition, each participant was invited to score the test against key criteria of usability, comprehensiveness and credibility/clarity, and our consultant also evaluated each test on this basis.

4. The tests covered most of the Ecosystem Services in the model, but two additional tests were carried out to cover those not explored sufficiently in the contracted work. The feedback from these is included in the report.

5. Although every test raised queries and uncertainties to at least some extent, reaction to the model was positive. Participants commented on its usefulness, felt that it had proved at least its potential, and wanted to engage further with it as it is improved. There were also users who responded positively to model-generated ideas that might improve their projects.

6. The model was assessed by participants at a score of 6.8 out of ten for **usability**. This score can be improved through these measures:

    - Better preparation, with the user forewarned about the data they will need based on the ES they are likely to select, and with GIS support provided beforehand to answer some of these questions, especially those relating to areas and populations.

    - A detailed, step-by-step approach in the Guidance Document, to allow the user to quickly find the answer to questions such as 'what do I put in this cell' and 'where do I go next'.

- Help and guidance on interpreting the results, and whether the changes the results show are significant or not.
- The model may need to be made available in languages other than English, and this may be particularly important for smaller municipalities where English capabilities (and confidence) are more limited.

7. The model was assessed at a score of 7.5 out of ten for **comprehensiveness**. This score can be improved through these measures:
    - The Guidance Document needs to be clear about dimensions of projects that are outside its scope, such as construction materials for buildings and non-green measures such as reducing speed limits.
    - A wider range of landscape types might be added to the available selection. In particular, rivers and canals are a significant omission that will arise frequently in landscapes for improvement.
    - Assurance in the Guidance Document that the list of ES is comprehensive.
    - Removal of the option to add additional ES, since this is entirely user defined and there is therefore no control over accuracy or calculation integrity.
8. The model was assessed at a score of 6.1 for **credibility and clarity**, a score that was distorted by two tests that failed to produce effective results. This score can be improved through these measures:
    - The model has been found not to work at the scale of a city-wide project. Any invitational material linked to the Business Model needs to make it clear that the model is intended for smaller-scale projects.
    - To test the model's effectiveness with mid-range projects, a dry run at the level of a district of a city, or a ward, should take place.
    - The future of the model, and updating of aspects such as costs, needs to be considered and responded to.
    - The provenance of the data used in calculations needs to be more easily referenced, perhaps in the Guidance Document, while the additional worksheets that contain this information may need to be hidden to avoid unhelpful experimentation with the underlying data.

- However, this aspect of the guidance needs to be accessible to non-academic readers, so care will be needed in the use of language and in verbatim citations.
- A database of Business Model users would help to assure new users in particular of authenticity and possible comparability between similar projects. The Green Cities Network envisaged by NSCiti2S may provide a vehicle for this.
- Sheet S5 C may need to be extended to allow input of other costs not currently recognised.

9. A total of 309 comments were made as feedback on the model; these have been organised by worksheet and significance to functionality of the model. They include comments from participants and observations made by our consultant during the tests. Many are repeated or cover similar ground. Of these comments, 103 have been categorised as 'must do', amendments that must be made if the model is to work effectively, and a further 56 are categorised as 'should do', changes that will improve the model. The remainder are categorised as less significant in terms of operating the model, but may include some interesting ideas.

10. Around three quarters of the 'must do' comments focus on nine areas of work:

- Typology, and the need to add some landscape types to the model;
- Cultural ES and measurement, widely regarded as unsatisfactory in its current approach;
- Metrics, with greater clarity of definition;
- Subjectivity, and the problems of non-scientific assessment;
- Interpretation of results for a non-expert user;
- Calculations, especially as regards costs;
- The use of the Shannon - Weaver Index in biodiversity assessment;
- Scaling up, and the limits of the model in terms of project size;
- Programming and layout errors.

11. The most significant of these are the Cultural Ecosystem services, subjectivity and results interpretation, and questions of scale. The 'should do' comments also focus on these areas and together make up 80% of the higher priority comments.

12. A large proportion of the comments arise from using S0, Project Description, and S2, setting ES parameters. There are fewer comments on the Factsheet but the comments on earlier worksheets often have Factsheet implications.

13. It is recommended that attention be focussed on the two lists of more significant comments, and particularly on those classified as 'must do'; and within these, on those comments that touch on the eight key areas of work.

14. Our consultant met with the Lead Partner and the Development team to go through the more significant comments and to agree an approach to the less easily resolved issues, in particular the cultural ES.

15. The draft Guidance Document was not ready in time to be used in the tests (and would have delayed them significantly) but several pointers have been fed back on how the guidance might be presented. Participants need very specific guidance about completing the worksheets, but also need to be able to access material establishing the provenance of the calculations made in the model. The Guidance Document therefore needs to be in two separate parts, one which provides hands-on immediate guidance as to completion, at worksheet cell level, and a second which provides technical explanation and background, as well as a bibliography of reference material used in constructing the model.

16. It is recommended that the Lead Partner and Academic Partners give further thought to

    - sustaining the Business Model beyond the lifetime of Nature Smart Cities, and keeping the supporting data and academic referencing up to date;
    - how to promote the availability of the model, and how to ensure that users always have the most up to date version
    - how to protect themselves against reputational or financial loss consequent on the use or misuse of the model;
    - whether the Guidance Document should be made available in other NSciti2S languages (it has been acknowledged that translating the model itself is both impractical and excessively expensive).

Table of Contents

1 Introduction

1.1 Background

Nature Smart Cities (NSCiti2S) is a partnership between several local authorities in the UK, Belgium, France, and the Netherlands, and three universities in the UK and Belgium. Among the project deliverables is the production of a robust business model, accessible and useable by local government officers and decision-makers in smaller cities and towns, that will reliably demonstrate the value, and the added value, of green infrastructure (GI) projects. The model is informed not only by academic research and knowledge, but also by exploring the state of current practice in the partner authorities (through a comprehensive interview programme) [1] and by examining the expectations of those authorities as regards the finished product.

The model applies techniques of Multi-Criteria Analysis to a diverse range of potential GI projects, and its output will allow a comparative assessment of different scenarios, including comparison between proposed nature-based options and 'grey' alternatives. It is intended to allow decision-makers in smaller cities to make more informed, and more confident, decisions that will take full account of the benefits to be gained through a green approach, and to justify such a decision should it come under challenge.

The model is being developed by the Universities of Antwerp and Ghent, and tested through direct application to the GI projects being developed, in a separate work package, by the partner authorities within Nature Smart Cities. It has been refined and fine-tuned through this process, and is expected to consist of three elements:

- A **spreadsheet,** in MS Excel form, that will process the input data and generate the results, together with a front-end to ease user access;
- A **Guidance Manual** providing a step-by-step methodology to support use of the model;
- An **output document (the 'Factsheet')** that will set out the results in a format to ease interpretation and comparison, and which will provide evidence to inform the decisions to be made.

Although testing within the partnership should point the way towards a working model, the project specification requires that this be proven further through a programme of 'demonstrator testing', assessing the model's value in a wider range of GI scenarios, and facilitate its use within authorities that have no vested interest and no prior knowledge or expectations of it.

[1] This was reported as 'Getting More Green', and is downloadable free of charge from the Nature Smart Cities website.

As Horber has noted, the process of refining a model is greatly enhanced by confronting it with empirical data.[2] This report sets out the results of that programme, the methodology used to secure those results, and recommendations for reshaping the model in the light of these findings.

.

1.2 Research objectives

The demonstrator testing is intended to add these facets to the Business Model development process:

- An **independent critical review** of the mode, exploring its capabilities and limitations, and making recommendations aimed at strengthening the final product;
- An **independent application** of the model to local authorities that are not pilot municipalities of NSCiti2S (and which therefore have little or no prior knowledge of the model) but which fit the target audience of the model, namely cities of less than 550,000 population located within the NSCiti2S programme area.

Within these broad objectives, the following specific outputs were identified

- A **report on the demonstrator testing** itself, with recommendations for refinement, and anonymised material for publication and wider dissemination;
- **Content for a presentation** to the project's capacity building programme, reflecting on the feedback secured;
- **Content for the guidance manual** in the form of changes to its structure, wording and content, and perhaps the inclusion of case studies developed from the testing process;
- Some **business case factsheets** generated from the Demonstrator Test authorities' use of the model.

A proposal based on these requirements, and including a detailed methodology and costing, was accepted by the Project Lead Partner, and Phil Back was appointed in April 2021. The NSCiti2S project timetable required him to complete the task by the end of October 2021, to allow the model development team time to make any modifications needed before their own deadline.

1.3 Methodology

[2] Eugene Horber, *Refining your models (unige.ch), U. of Geneva and FORS, 2021

1.3.1 Familiarisation

The investigation began by seeking familiarity with the way the model works, the data gathering and input requirements, and the content of the guidance manual. This included viewing presentations by Masters students at the universities and asking questions about their work, which has been used to inform aspects of the business model; this helped to flesh out an understanding of the underlying approach.

COVID-19 measures meant that a planned meeting with the Development Team could not take place, but our consultant was able to work with a preliminary version of the Business Model and with subsequent versions as development continued through the summer of 2021, and to interact regularly with the development team to discuss his findings.

After an initial exploration of the model using an invented scenario, it was decided to use two real-life examples that might demand more of the model and which would improve familiarisation, in particular requiring the answering of questions from users. Two 'dry runs' therefore took place in July 2021 with real-life projects being worked on by professional staff at an NSCiti2S partner, using a face-to-face approach and taking notes of issues that emerged. One of these projects was a new country park on the city fringe, the other a new nature reserve being planned for the city.

These two dry runs were very effective in terms of familiarisation, but also highlighted that the model was not yet ready to be tested with external municipalities. A list of possible issues and challenges was drawn up (some of which highlighted weaknesses in understanding, others identifying problems within the model), and sent to the developers. One major issue was an inability to progress beyond a particular point in the process, as the software development had yet to be completed beyond this point; in particular, Factsheets could not be generated, and several other errors became apparent. These were resolved over time using an iterative process in which issues were identified, using dry run data, and relayed to the programmers. The development team also conducted their own internal dry runs.

Two further dry runs were carried out in August 2021 using the same scenarios, and a third dry run in September 2021 using a projected project in a Belgian authority aiming partly to verify that the requested changes had been made, partly to test those worksheets that had so far been inaccessible, and partly to cement an understanding of the model's workings.

These dry runs indicated that it was now possible to work through the model to the final stage, and that many of the issues raised during this internal testing process had been resolved. Although there were still problems in the spreadsheet, many of these were of a cosmetic

nature and would not inhibit formal testing with external bodies; any others that emerged during testing could be attributed to the nature of the project as a work in progress.

By this time arrangements were already in hand to carry out the Demonstrator Tests, precluding further development time without prejudicing deadlines. After discussion with NSCiti2S academic partners and the Lead Partner, it was decided to proceed with the model as it stood in September 2021, and to use the results to inform the further development and polishing of the model.

There were therefore still outstanding issues moving to the test phase. These were already notified to the development team, though, and are not replicated in the report unless they were also picked up during the tests (as many of them in fact were).

1.3.3. Guidance Document

The Guidance Document was supplied in draft form in July 2021, and read and annotated for future attention. However, the team's focus at this point was on improving the functionality of the model itself, and no updated copy was supplied until the day before the first test was due; a skim-read suggested, unsurprisingly, that although the guidance had been considerably developed from the early draft, it had not kept pace with reform of the model itself and might therefore confuse a user. There was also no time to study it closely before launching the test phase, nor to supply it to participants in reasonable time. Moreover, MS Teams was to be used for some of the tests and it would not be possible to observe its use.

Because of these factors, and because of the time requirement that emerged as the tests were conducted, it was decided to leave the guidance to one side during the tests themselves, but to use the feedback from the tests not only to update the model but also to refine the guidance, by reflecting the questions posed as to how to use each worksheet in the spreadsheet. The test results have therefore been used to suggest refinements to the guidance as well.

1.3.2 The call for participation

An open call was issued in June 2021 to all eligible municipalities in the 2seas/mers/zeeën region, using specific and generic e-mail addresses, online newsletters, the project's own website and contacts networks, and other informal networks. These informed relevant personnel of the opportunity to participate; follow-up information was provided on request in the relevant language.

Selection criteria were developed in case of a need to select from a larger than expected pool of applicants, but these were used to select project ideas rather than participants. The original brief was to identify four potential participants, one in each partner country, but in the end, it was decided to invite five interested municipalities to take part, three in the UK and two in the Netherlands. No French or Belgian authority expressed more than a passing interest.

However, two of the project partners also received permission to use surplus funding to develop additional pilot GI projects, and it was decided that the Demonstrator Testing should also be applied to these if possible, as well as a further NSCiti2S partner which had to this point had no contact with the development team. This made eight test sites in total.

Each potential participant was sent some information about the programme, indicating confidentiality and other protocols to be applied, and including a consent form which included acceptance of a disclaimer as to the accuracy of the software, given that this was a test rather than a genuine application.

1.3.3 Preparation for testing

The test process was designed as a combination of observation, questioning/challenge and note-taking to record the users' visible response to what they encountered on-screen and exploring the user's perception of the experience. An observation record, an outline script, and a scoring matrix were all used to support assessment.

Each municipality received some information about preparing for the test by assembling some basic data that the model might require, based on an understanding of the kind of project they would be presenting. Some municipalities took the preparation very seriously, but others less so.

1.3.4 Test authorities

Eight tests were undertaken: five with authorities outside the NSCiti2S partnership, and three with NSCiti2S partners. Two were tests on new pilot schemes added to NSCiti2S late in the project timescale, and one was an NSCiti2S city not covered by earlier work with partners. The identities of the test authorities are withheld, to honour guarantees of anonymity and confidentiality.

Most authorities sent at least two participants to the test; only one was single-handed, while one had no less than five officers present. The tests took place using a variety of equipment

including Windows and Apple laptops of different vintages. The intention had been to conduct the process entirely in person, but COVID-19 protocols inhibited this and, of the eight tests, only five could be undertaken in this way; the others were conducted using MS Teams.

In addition to the eight contracted tests, a further test was designed to explore the treatment of ecosystem services (ES) not covered adequately by the contracted tests; this is referred to as 'Leysers' and is entirely fictitious. There was also a demonstration of the model to a group of officers at a UK Council, using a real-life scenario from the dry runs. Both generated feedback, which has been included.

1.3.5 Analysis

Analysis includes both general examination, looking at aspects such as coverage of ecosystem services, and more specific, looking at the test projects in more detail and at participants' reactions as they worked through the material. The analysis thus moves from the general to the more specific, but includes some overall conclusions and recommendations, as well as some other observations that may or may not be helpful.

Exhaustive notes were made during each test, recording problems that arose but also a general sense of how well the participants were grasping the essence of the spreadsheet and its requirements. In practice, this tended to vary as different elements of the spreadsheet were encountered, and the most straightforward way to analyse the tests is to explore responses on a worksheet-by-worksheet basis; an analysis along these lines is included. For each worksheet, at the record includes

- The general reaction of participants to what was being asked of them
- The faults, suggestions and comments they made on each worksheet as the test progressed, categorised into four groups:
    1. Faults requiring corrective attention ('must do', or Level 4)
    2. Enhancements that might improve function or coverage ('should do', Level 3)
    3. Other possible changes for consideration ('could do', Level 2)
    4. More esoteric suggestions beyond the scope of the model ('won't do', Level 1)

Each of the test groups was scored on the key characteristics identified in earlier research: useability, comprehensiveness and credibility (including clarity), using balanced, weighted sets of criteria agreed with the Lead Partner and set out as Appendix 1. Each participant was also invited to score their experience, and to comment on the score they gave, against each of the key criteria, as a simplex score out of ten.

2 The Business Model

The model takes the form of an Excel spreadsheet organised into a series of worksheets. Some of these require user-generated data to be input, while others allow user data to be entered to supersede the data already contained in the model. The data used to perform the calculations in the model is set out in a series of further worksheets which are available for users to examine but which are not easily modifiable. This section describes the model as it was at the time of carrying out the Demonstrator Testing, and before any significant modifications were made.

The worksheets are as follows:

Introduction: a text-only page that provides a basic introduction to the model and describes the worksheets that follow. This worksheet duplicates material in the Guidance Document and may therefore be considered superfluous.

S0 Project Description: this page invites the user to describe their project in terms of its location, the number of people expected to benefit, and its surface area. There is also an opportunity to provide a brief description of the project and to upload a picture (although this was only rarely successful).

The worksheet also asks the user to define their infrastructure as a series of two or more scenarios. One of these is a Baseline scenario, which would normally (but not always) describe the existing landscape before the green infrastructure project is developed. Further scenarios can be used to define the landscape after the green infrastructure project is in place, and also to define alternatives to the envisaged project. Scenarios can thus explore the impact of a GI initiative, the different outcomes of two alternative propositions, or even the impact of a loss of green components such as through development, to allow exploration of mitigation strategies.

Within each scenario, the user must define their landscape in terms of its components. Several groups of related landscape elements (e.g. trees and shrubs, water, and so on) are available, and within these different types can be selected. For each type, a quantity must be entered; this will normally be the area of landscape described by the type, but for some elements such as trees a numerical quantity is required.

This worksheet is fundamental to the successful operation of the model. The more accurate the data entered here, the more useful the model is. Preparation, in terms of landscape knowledge and measurement, is absolutely essential, and several users commented to this effect. GIS input at this point is especially helpful.

S1 Selection: In this worksheet, the user defines the ES they wish to explore through the model. Although it would be possible to develop results for all 15 listed ES, most projects will have a more limited focus and users may be able to identify ES that are more fundamental to the rationale for the project. The guidance recommends a maximum of six, but this is driven by space limitations in some of the graphics and in practice seven or even eight selections seem to be quite manageable. The more ES selected, though, the heavier the workload for the user in answering questions. It may be noted that the ES cannot be varied between scenarios; all scenarios must use the same ES framework.

S2 Parameters selection: This worksheet looks for input data specific to each ES selected in the previous step. Thus for water retention, the worksheet seeks data on precipitation in the locality, and the average depth of any drainage being incorporated, for instance. Cultural ES are addressed through a series of statements, which must be confirmed or denied individually as either Correct or Incorrect for each scenario. There was a good deal of criticism of the cultural ES approach, and this may well be changed. Using this worksheet generated several comments and challenges for the developers to consider.

Tab A: Water retention: This is the first results sheet, and uses the precipitation data, and a set of 'Retention Co-efficients' to assess the landscapes described in S0 in terms of their capacity to hold water. The sheet aggregates landscape types from S0 into groups that share similar retentive capacities, but this can be modified if the user wishes.

Tab B: Biodiversity: This worksheet uses the same aggregated landscape data to calculate the Shannon-Weaver biodiversity index for structural variation. A spider chart is also generated to indicate the landscape diversity thus calculated. Again, the user can modify the landscape components here to see what effect this has on structural diversity.

Tab C: Biodiversity: In this worksheet, the user explores the landscape in each scenario in a more granular way. A fairly lengthy series of landscape elements must be selected or rejected and from these a calculation assesses the suitability of the habitat for a range of different species of bird, butterfly, bee and amphibian. This enables comparison of the scenarios for their habitat suitability, but also allows the user to explore modifications to the landscape that might promote habitat for particular species that are deemed desirable. A chart is produced for each scenario that assesses the suitability of the habitat, while scores for individual species show those most likely to colonise the landscape.

S3: Qualification: This worksheet invites the user to deploy a score to assess the extent to which each of their chosen ES will change as a result of implementing each of their new scenarios. The premise is that the Baseline scenario will be scored at zero (though this can be changed) while the new scenarios will show change driven by implementation. A spider

diagram that enables comparison between scenarios, and between new and baseline scenarios, is generated. This section was heavily criticised for being excessively subjective, but users found the next worksheet helpful in completing this one.

S4: Quantification: This worksheet takes those selected ES that are capable of measurement and reports the results of the calculations to show how each of these ES changes from one scenario to another. For example, the baseline may itself capture some carbon, and this is reported, as is the carbon captured in each of the new scenarios, to show the extent to which these contribute to the carbon-related ES. The measures used are those most pertinent to the ES in question and are not comparable between ES. Users indicated that they needed help in interpreting these results and particularly the extent to which the results change between scenarios.

S5: Monetisation (Costs): This worksheet is referred to in this report as S5 C. For each landscape type, it provides data gleaned from reliable sources to indicate a cost range for construction (i.e. a capital cost) and maintenance (i.e. a revenue cost). The user can decide to choose a minimum cost, a maximum cost, or an average cost depending on how they wish to present financial information about their project. They can also override the suggested costs by providing more specific financial data if they have this available. Construction costs will normally only be provided for the new landscape elements (the costs of the existing landscape have already been incurred) but maintenance costs can be compared across all scenarios.

S5 Monetisation (Benefits): This worksheet is referred to in this report as S5 B. For each monetisable ES selected in S1, the worksheet calculates a monetary benefit that attaches to the results. This is unlikely to be a cash benefit, but may nevertheless be a realisable outcome in the wider economy, or for a related public service. Thus, for example, if the health ES has been chosen, the value of additional years of good health will be realisable in a reduced or delayed demand for services related to morbidity and mortality. A discount rate can also be applied to allow calculation of the present value of future savings or cost reductions.

Factsheet: This worksheet summarises the results emerging from other worksheets, and allows these to be set side by side to enable comparison between different scenarios (two at a time). It collects information about the project from S0, highlights the ES selected in S1, repeats the spider diagram generated in S3, sets out the quantifiable results from Tab A and S4, and the biodiversity charts from Tab C. It then lists the statements and responses for the selected cultural ES from S2, and presents graphs showing the capital costs for each of the two scenarios chosen, the maintenance costs (both of these collected from S5 C), and the monetary benefits predicted from S5 B. Finally, the Factsheet provides a box in which the

user can add any text commentary that seems appropriate. A facility exists to produce the Factsheet as a printable PDF.

Other worksheets: As the model stands at present, nine other worksheets are reachable beyond the Factsheet. These contain look-up and calculation parameters used in the calculations, and also references to sources used to produce these. They might be of interest to a user seeking assurances over method or provenance, and some can be modified by a skilled Excel user, but the risks associated with modification are high. These sheets were not revealed to the users (though some were aware of them) and might be hidden or at least password-protected to avoid accidental, catastrophic amendment, while providing methodological or provenance-related assurance through the Guidance Document as an alternative reference point.

3 The Tests and Assessment

3.1 The tests and coverage

As the introduction indicates, the tests were spread across three different countries in the 2seas/mers/zeeën region; unfortunately it was not possible find a test partner in France. The main negative implication of this is that the tests do not prove a need for the Business Model to be available in French, although it seems likely that this would be at least desirable, especially perhaps in smaller authorities.

Despite this it was still possible to conduct the tests across a range of projects, as this chart indicates:

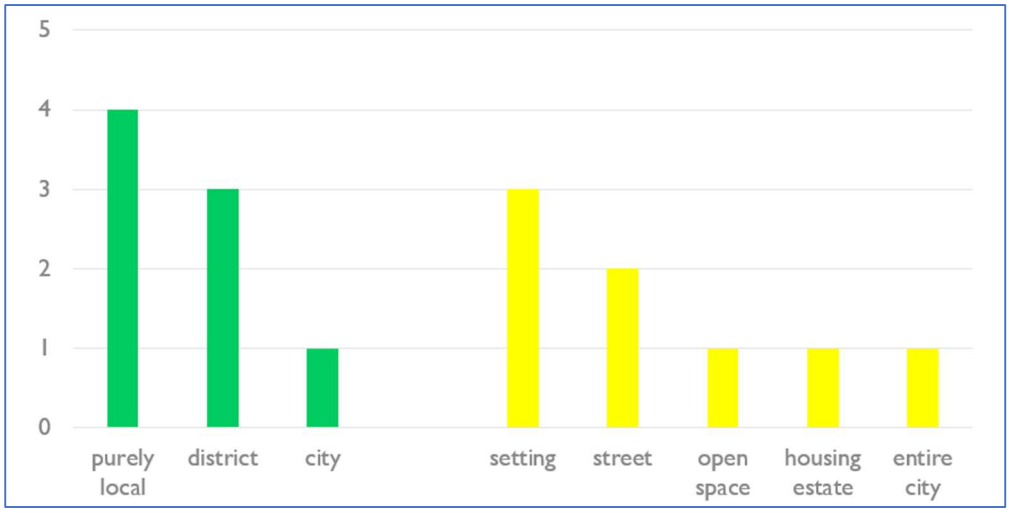

Fig. 3.1 Settings of Demonstrator Tests

Of the eight tests, half were purely local in nature: a street or a local park, for instance. Three covered a larger area such as a housing estate or other district-type setting, while one covered a whole city. Three of the tests concerned the setting for a building – a town hall, a school, an office property – while two were linked to streetscene. Only one test related to an open space, and one to a housing estate. A total of 24 participants took part, with a further 4 added as part of one of the additional tests.

One of the additional two tests was purely local, and one might be considered district-wide; both would be thought of as open space projects.

It was also important that the tests covered the widest possible range of ES, since each ES makes different and distinct demands on the Business Model. This chart sets out the ES covered in the eight contracted tests:

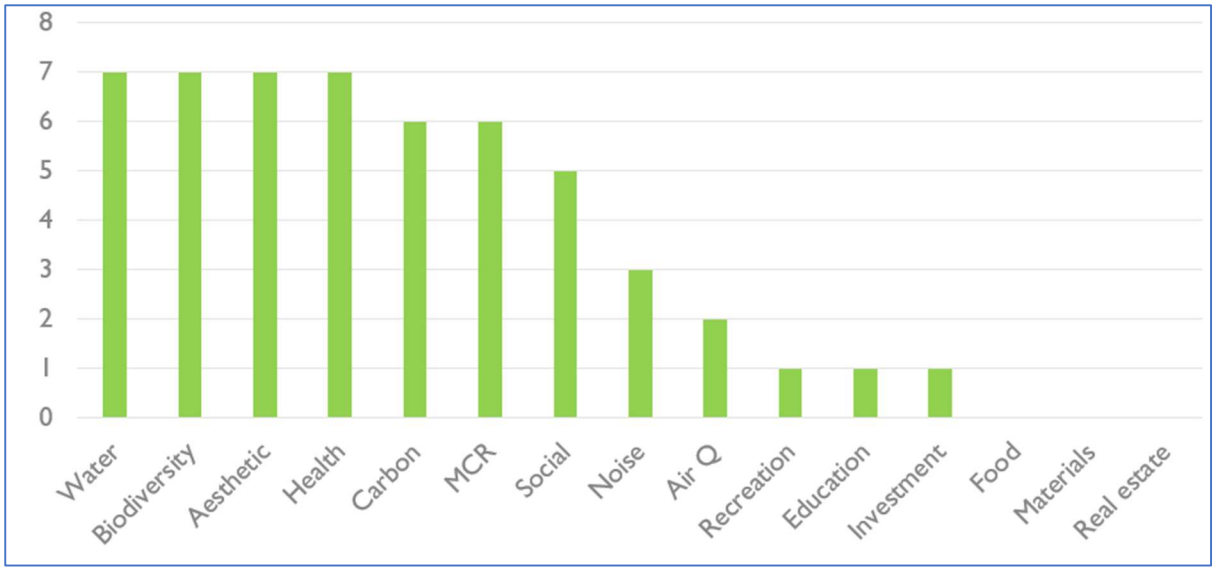

Fig. 3.2 Ecosystem Services covered in contracted tests

Of the eight tests, seven explored water management, biodiversity, aesthetic appreciation or health as key ES the participants were hoping to deliver through their projects. Carbon capture, micro-climate regulation (MCR) and social cohesion were also prominent on participants' agendas. Other ES offered in the Business Model were less significant overall, although noise reduction was a major outcome sought in one test.

The prominence of aesthetic appreciation is interesting, because this was identified in earlier research[3] as the most important priority that decision-makers were looking to address through providing green infrastructure. Since this was identified by practitioners as an important aspect of seven of our eight tests, it clearly still resonates. Health was also a prominent goal attached to GI in that earlier study, carried out before the global pandemic, and it is unsurprisingly still a major aim of GI post-COVID.

Food, materials and real estate prices were not addressed at all in the eight contracted tests, but needed to be tested, which is why additional tests were added. Both the two additional tests included food, and one looked briefly at materials and real estate prices, with the aim being to see how the model handled data in these areas.

[3] Phil Back and Alex Collins, *Getting More Green: Smaller Municipalities' Approaches to Delivering Green Infrastructure* (London: Imperial College London, 2020).

3.2 General feedback

Although no test was problem-free, and in spite of an extensive range of comments and constructive criticism, the overall impressions of participants were positive, as these verbatim comments indicate.

Comments like these are extremely positive, and suggest that the tool will be both welcomed and valuable in making the case for green infrastructure, but also in looking at alternative scenarios and evaluating their different impacts. The third comment here relates to a conversation where parameters were changed to allow the user to evaluate different biodiversity outcomes, and they could see immediately the value of the tool not only for assessment but also in the planning stages and for thinking about what components they should include in their landscape.

These comments indicate a role for the tool beyond its primary purpose of assessing the project and comparing different approaches. The questions the tool asks the user give them the opportunity to consider additional potential dimensions to their ideas, such as the incorporation of a natural playground, or the use of permeable surfacing, which they might not have thought about, but which could add value to what they are planning.

Comments like these illustrate that, for all the problems that still need resolution in the Business Model, the tool is communicating effectively with its target audience and is showing its potential very clearly. The idea of the tool, and the mechanics it uses to do its work, are established as credible and do not need much further justification, while its capacity to turn anecdotal evidence into something more scientific and empirical is also clear and recognised.

Even the least successful of the tests, where data had to be invented and the results could not be evaluated, provoked an enthusiastic response from the participants and a desire to repeat the exercise using real-life data. So too did a test that raised several doubts and questions, but nevertheless left the participants wanting to see how these might be addressed in a future version. These comments suggest the model is not far away from where it needs to be.

3.3 Scoring the tool

The earlier research identified three key criteria that any tool would need to address strongly if it were to succeed and be recognised as of value: useability, comprehensiveness, and credibility. Each participant was invited to assess the tool against these criteria, and our consultant developed a scoring matrix allowing him to carry out a similar exercise using his appraisal of the constituent components of each of these criteria. A clarity measure to the credibility score as this is an important part of credibility: if the results are not communicated effectively, they will not be seen as credible.

Each participant was invited to give a score out of ten, and their scores were then averaged to produce a user score. The consultant's scores were allocated to a list of weighted sub-

criteria to produce a similar score out of ten. The sub- criteria used, and their weights, can be found in Appendix 1; each of the three principal criteria has the same overall weighting.

These are the scores for usability:

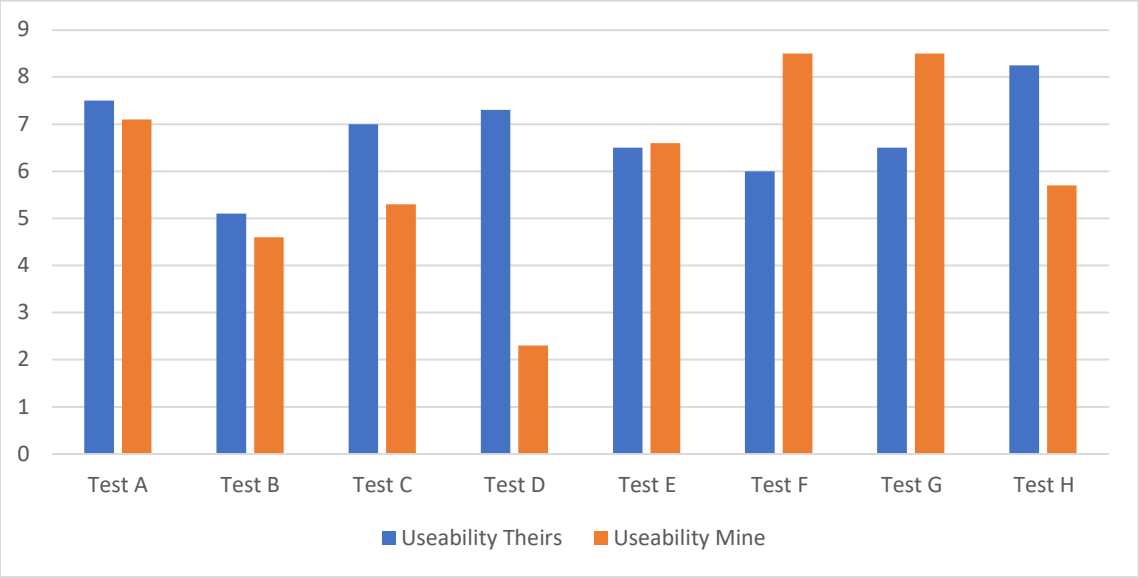

Fig. 3.3 Usability scores

Useability is essentially the user-friendliness of the model, and the ease with which a user adapts to its mechanics and requirements. It encompasses the need for preparation, but also the need for specialist knowledge in order to answer questions, and the help and support the user needs in order to enter data or understand the results they are given.

The scores here are generally high, which is an encouraging result for the model. They reflect the fact that most participants had very little difficulty in 'tuning in' to the model and quickly acclimatised to how to populate the worksheets and how they could modify pre-populated data. Both expert practitioners and non-experts handled the mechanics of the model well. Only one participant struggled with the entry process, and this may well be attributable to an older version of Excel, an older laptop and operating system (or both), and a greater degree of scepticism about the complexity of the model. On the other hand, a junior member of one team was able to handle the demands of the model very easily, thanks to some familiarity with Excel. In two cases, an immediate negative reaction to the appearance of the model was quickly overcome.

All scores bar one are above the midpoint of 5. The lower scores arise because of problems and limitations that are enumerated in the detailed comments. Scores in Tests F and G are marked down by the users because of the problems they encountered in specific areas of the

model – in Test F in relation to S3 and the spider diagram, and in Test G due to a data loss in Tab A and unbelievable costs and benefits in S5. The consultant's perspective, however, was that users in these tests were allowing these issues to obscure the wider useability of the model, which was not the root cause of their issues.

In contrast, the consultant's score for Test D was very low because the municipality had not done their preparation and had no access to the data they needed to populate sheet S0. As a result they were forced to use imaginary data and to generate meaningless results. Preparation was part of the consultant's usability assessment, so they got a low score here. However, this problem did not seem to qualify their own largely positive assessment of the model, and their wish to repeat the exercise.

The average score across all users is 6.8, the consultant's average score across all tests is 6.1 but would be 6.6 without the Test D outlier.

Specifically to improve usability, the following actions are recommended:

- Better preparation, with the user forewarned about the data they will need based on the ES they are likely to select, and with GIS support provided beforehand to answer some of these questions, especially those relating to areas and populations.
- A low-level approach in the Guidance Document, to allow the user to quickly find the answer to questions such as 'what do I put in this cell' and 'where do I go next'.
- Help and guidance on interpreting the results, and whether the changes the results show are significant or not.
- The model may need to be made available in languages other than English, and this may be particularly important for smaller municipalities where English capabilities (and confidence) are more limited.

These are the scores for comprehensiveness:

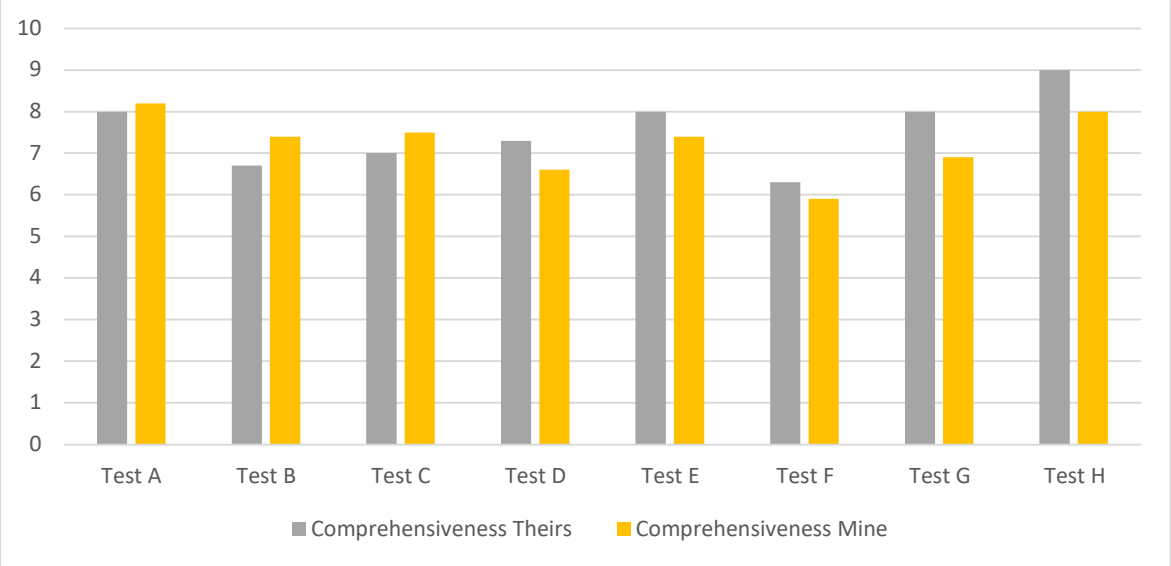

Fig. 3.4 Comprehensiveness scores

Comprehensiveness addresses the extent to which the model deals with the landscape being presented, and the way in which it covers the range of ES being sought through the project. A high score on comprehensiveness will also embrace the extent to which the model seeks the information the user wishes to provide, does not overlook important aspects of ES, and handles the notion of comparability across scenarios.

The scores here are high, both from the users and from myself. They suggest that there is very little wrong with the model in terms of its coverage of the ES that matter to users, although this does not necessarily mean that the results are satisfactory. One user stated that the model had provided tangible and valuable results that had not been available, and not available with this provenance, in the past.

No obvious gaps emerge here, although the testing did not test all ES with the same rigour; real estate prices and attraction of investment would bear further exploration. Also, scores were reduced a little where landscape types proved hard to classify correctly in S0. The Tests B and C scores reflect disappointment that additional non-GI interventions in their projects are not recognised by the model (since these are not GI, the model's approach here is reasonable); the consultant's scores for both are higher because he placed less emphasis on this. One user declined to score this aspect because they needed expert input to know whether all ES aspects had in fact been covered.

The lowest scores come from Test F, where participants would have liked an increased range of ES, and a more visible demonstration of the multifunctionality of GI. However nobody sought to add additional ES, which can be done but at huge cost in terms of providing

calculations and results manually. There was also an extensive discussion about cultural ES in particular, which seemed to highlight a concern over this, and over the subjective nature of sheet S3.

A further factor in two of the scores is data loss from Tab A. This is addressed in the list of comments but denied two groups of participants the chance to see one of the most useful (and least contentious) results sheets on offer for their projects.

The average score across all users is 7.5, the consultant's average score across all tests is 7.2.

To improve comprehensiveness, these actions are recommended:

- The Guidance Document needs to be clear about dimensions of projects that are outside its scope, such as construction materials for buildings and non-green measures such as reducing speed limits.
- A wider range of landscape types might be added to the available selection. In particular, rivers and canals are a significant omission that will arise frequently in landscapes for improvement.
- Assurance in the Guidance Document that the list of ES is comprehensive.
- Removal of the option to add additional ES, since this is entirely user defined and there is therefore no control over accuracy or calculation integrity.

The scores for credibility and clarity are shown here:

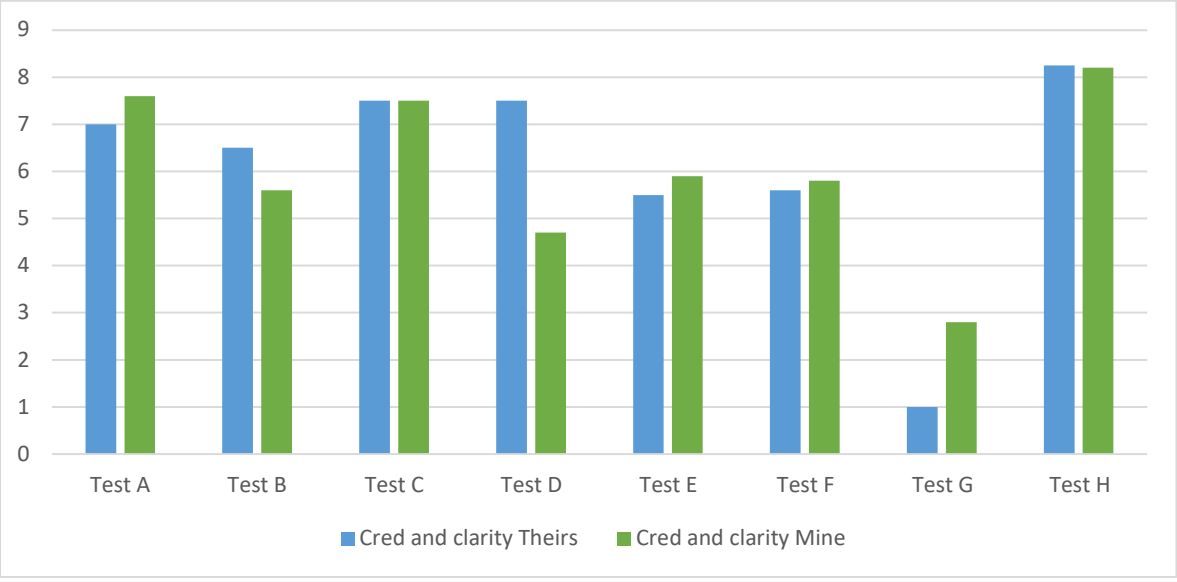

Fig. 3.5 Credibility and clarity scores

Credibility is fundamental to the business model, because if the results are not believable then the work that has gone into producing them is wasted. But the results are only part of the story; the results must also be presented well, using the right language and terminology and set out in a way that inspires confidence, with clear and compelling visuals and numbers that can be explained or interpreted by the person preparing the results.

Here the scores are much more mixed, but only in one case does the user's score fall below 5.0. This is Test G, where the model did not handle the special demands of a city-wide project and generated stratospherically high financial outcomes in particular. These results have overshadowed more successful aspects of the test, and the consultant's higher score reflects this while acknowledging that the test here was in fact a failure. However, although the results in Test D could not be assessed for credibility (hence the consultant's low score) the users clearly had confidence in the process in spite of this.

In contrast, one user found the results entirely credible despite losing the water retention result to a data failure identified in the comments. Another found that the cost results approximated fairly closely to the costs they had themselves agreed with their contractor when setting up the job. However, other costs for removal of existing infrastructure were not present, nor was there an obvious way to add these in.

One user gave the model an enthusiastic maximum score of 10 for this criterion, but this was clearly overstating their experience given the other feedback they gave, and their score was counterbalanced by that of a colleague who recognised this. Other users, while challenging some results, found some others to be entirely believable and useful.

The flexibility offered in costing using custom values was especially welcomed, and at least two users noted that costs were changing rapidly in a period of supply problems (not unique to the UK) and that amending or updating costs would be an essential part of using the model into the future

Several users asked questions about provenance. While not (necessarily) doubting the accuracy of the calculations, they nevertheless wanted to know where the knowledge that underpins these had come from. One user had located some information in this respect on the additional worksheets, but this was not encouraged; another would want to explore some of the calculations and rationales with local experts (although the premise of the model is that this expertise does not exist, the model needs to be fireproof in this respect). Users would also welcome some assurance as to which other municipalities are using the model, as a basis for providing reassurance to their decision-makers as to authenticity.

The average score across all users is 6.1, the consultant's average score across all tests is 6.0. If Test G is excluded (as perhaps it should be), then these scores rise to 6.8 and 6.5 respectively.

To improve credibility and clarity, these actions are recommended:

- The model has been found not to work at the scale of a city-wide project. Any invitational material linked to the Business Model needs to make it clear that the model is intended for smaller-scale projects.
- To test the model's effectiveness with mid-range projects, a dry run at the level of a district of a city, or a ward, should be set up.
- The future of the model, and updating of aspects such as costs, needs to be considered and responded to.
- The provenance of the data used in calculations needs to be more easily referenced, perhaps in the Guidance Document, while the additional worksheets that contain this information may need to be hidden to avoid unhelpful experimentation with the underlying data.
- However, this aspect of the guidance needs to be accessible to non-academic readers, so care will be needed in the use of language and in verbatim citations.
- A database of Business Model users would help to assure new users in particular of authenticity and possible comparability between similar projects. The Green Cities Network envisaged by NSCiti2S may provide a vehicle for this.
- Sheet S5 C may need to be extended to allow input of other costs not currently recognised.

3.4 Feedback on the Business Model

As the test took place, copious notes were taken on the questions being asked and the issues and uncertainties that the users were encountering. These have been collated into sets of comments, grouped by the significance attached to them and using the categorisation already described (*S. 1.3.5*).

A total of 309 comments has been collected, but these are by no means all at the same level of significance, as the chart below indicates:

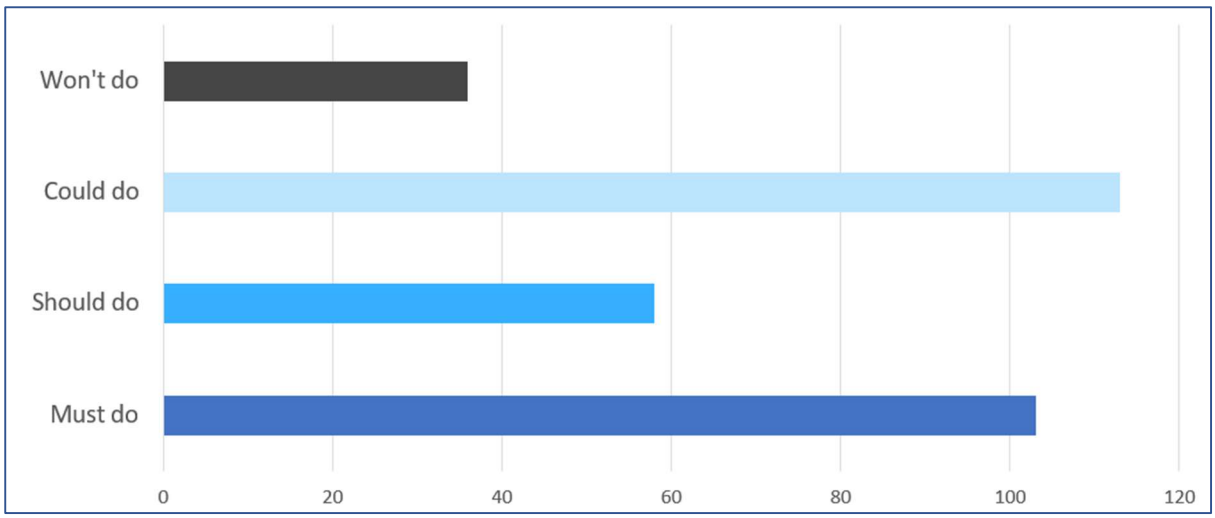

Fig. 3.6 Categorisation of comments (N=309)

A third of the comments are classified as 'must do', putting them at Level 4, and meaning that they affect the successful function of the Business Model. Not addressing these will cause the Business Model to fail, or to generate inaccurate or misleading results, at the very least affecting user confidence and undermining credibility, and with the potential to take users towards an unjustifiable decision. However, although the number of these comments seems large, several emerge more than once through the tests, and there are overlaps between comments, so the workload for the development team is not so great as it might seem from the numbers. Also, although some of the comments focus on areas that are complex and require careful consideration, many are simple to address, with relatively easy corrections and adjustments.

Just under 60 comments are categorised as 'should do', or Level 3. Attending to these comments will enhance the model, and improve it in different ways, without being fundamental to successful operation.

There are over 100 comments that can be classified as 'could do' (Level 2), suggestions that are not outrageous in any way, and which might improve the model, but which might be seen as less important or urgent than those in the higher categories, and therefore less pressing when considering the workload and the time available to make adjustments. And there are a further 36 comments that, while made in good faith, are more speculative in nature and are unlikely to be carried through to reality. These are classified as 'won't do' and are at Level 1 in the comment hierarchy.

The comments have also been classified according to the worksheet to which they relate:

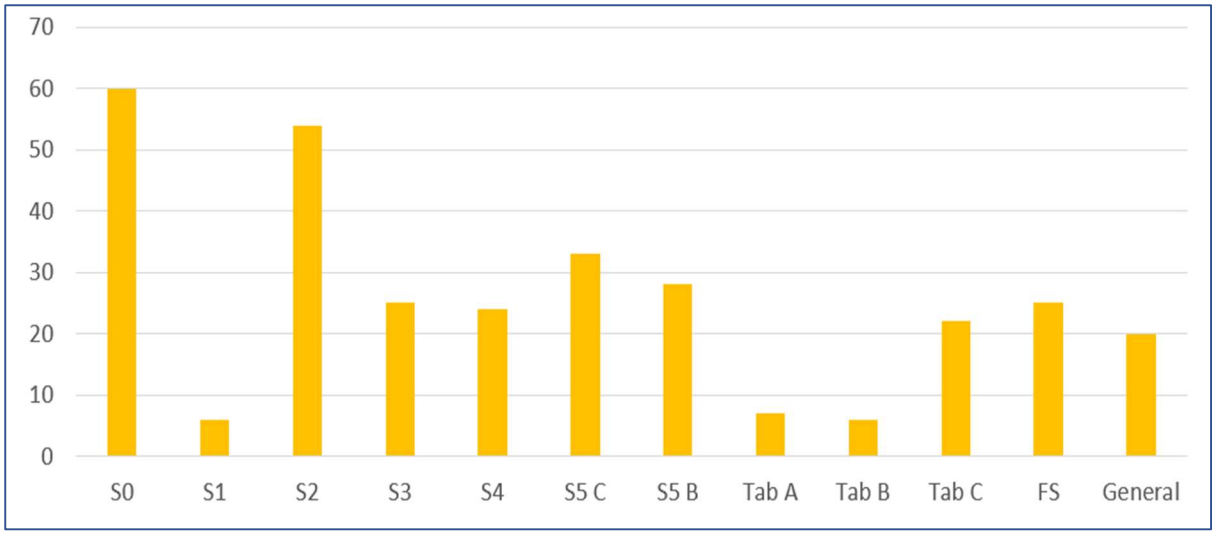

Fig 3.7 Comments classified by worksheet

The comments are most heavily concentrated on the sheets where user input is greatest. Sheet S0, where the scenarios and their landscapes are defined, attracts the most comment, closely followed by sheet S2 where the user sets the parameters for the ES they have selected. These two worksheets dominate the chart, with the next most prominent worksheets being the two financial ones, S5 C and S5 B.

There are far fewer comments on S1, where ES are selected, and on Tabs A (water retention) and B (Shannon-Weaver index). Tab A is largely regarded as satisfactory, but Tab B was not visited very much in these tests and will need further discussion with the developers. And although the number of comments on the Factsheet is limited, many of the observations made on other worksheets have implications for what carries through to the Factsheet.

Looking at this in more detail highlights ten areas for particular attention by the developers, as this chart of the most significant (i.e. 'must do', or Level 4) comments shows:

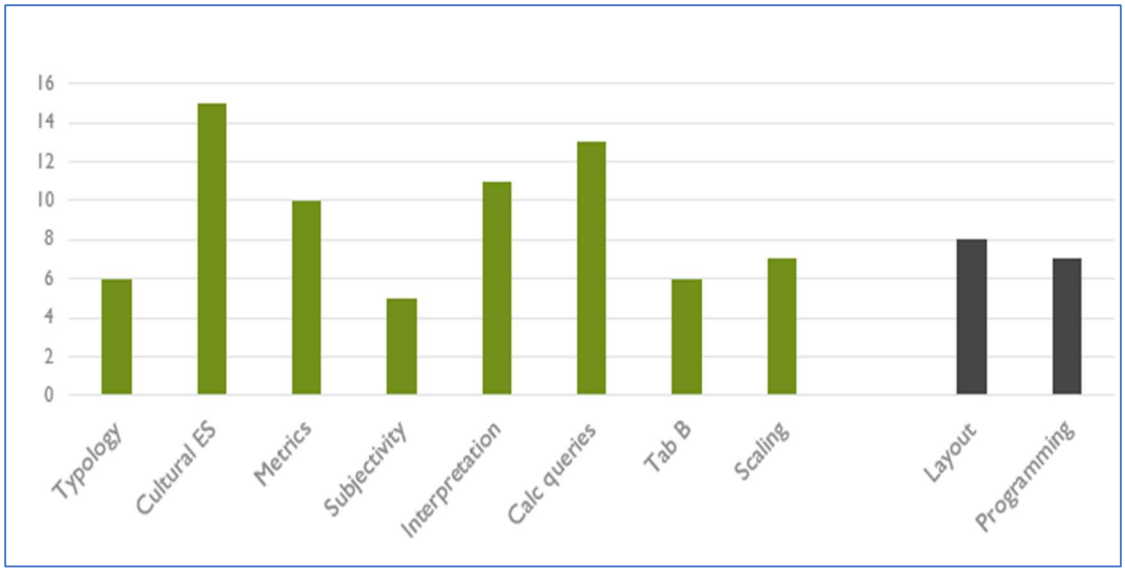

Fig. 3.8 Level 4 comments by subject

Nearly three quarters of the 'must do' comments are focussed on eight areas of work. Adding in the programming and layout-related comments takes this to 85% of these comments, around five out of six of the total. And adding in the 'could do', Level 3 comments results in this chart:

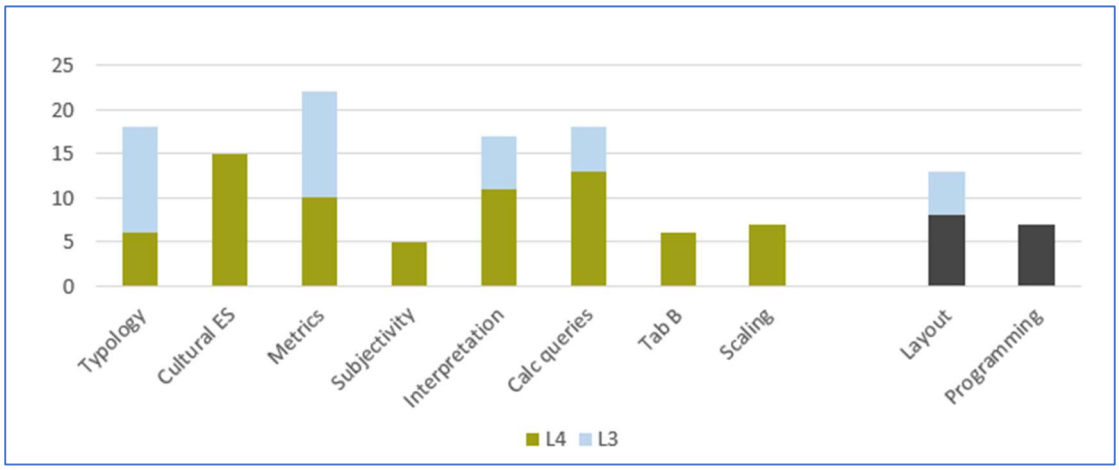

Fig 3.9 Level 4 and 3 comments by subject

There are Level 3 comments in five of these groups, so addressing both sets of comments would address 161 of the 208 comments in these two levels, around 80% of both lists.

The issues being raised can be summarised thus:

- ***S0 importance and typology***

The model is built around the definitions of scenarios and landscape elements identified in this worksheet, and two critical issues have been identified here. One is the absolute, critical importance of preparation and the need for the user to understand what the model needs in order to fly. The other is that some landscape elements have been overlooked or omitted, including some important and common ones like watercourses, and this needs to be addressed.

- ***S2 Cultural ES measurement***

Users express widespread dissatisfaction with the way the model handles cultural ES. In terms of pure numbers, this is the most important of the Level 4 comments, and the focus is consistently on the same issues. The present list of statements is seen as tedious, the selection of 'Correct' or 'Incorrect' responses to each is too absolute and too binary in nature – not allowing the user to show development beyond a second scenario – and some of the statements are reversed in polarity, confusing the user and leading to errors. The statements are also not as comprehensive as they could be in defining the characteristics of the ES in question; some are drawn from academic sources, but others are admittedly not. Users also dislike the way these are managed into the Factsheet and do not think this will impress a decision-maker. They also think that there is more potential for calculation here than has been allowed for thus far.

This is by far the most far-reaching issue identified through the feedback, not least because the approach taken here has implications for sheets S3 and S4 and touches on the issues of subjectivity and interpretation, and has massive implications also for the Factsheet as it currently exists. A lengthy but constructive discussion of the problems and possible solutions has taken place with the development team, and it has been agreed that:

- We will review the lists of statements and revise them where necessary to ensure they cover the ES adequately;
- We will ensure that all Cultural ES statements are of the same polarity;
- We will explore the possibility of using a graduated scale to assess each statement for each scenario;
- We will look at whether it is possible to program sheet S2 so that a user can choose a limited number of statements they especially want to promote to their decision-maker on the factsheet (the easier solution considered);
- We will look at whether it is possible to programme sheet S2 so that a user can allocate a score to each statement under each scenario, and so generate a

global score for each cultural ES selected, that can be used to populate S3 and S4.

- ***Numbers and metrics***

There are issues with some metrics that are not defined, including the very important total area in sheet S0, which has to be in square metres to verify subsequent data entry but is not labelled in this way. Radial measurements need to be clarified as to whether the radius is drawn from the centre of the site, or from its boundary, and language also needs to be clarified for the user to know whether a figure required is houses or buildings, and whether those who enter and leave a space but do not live there (school children, for example) should be included or excluded from some calculations. Although metrics represent a substantial proportion of the higher-level comments, this is a disparate group of comments and many of the issues here are very easily resolved through clearer and definition.

- ***S3 subjectivity***

This worksheet is regarded as excessively subjective, and this is an almost universal view among participants. They feel it undermines the science behind the other worksheets, and is too easily manipulated – and that this is too transparent to anyone challenging the assessment. Some participants love the spider diagram, others are less moved by it, but few are happy with the way it is created or changed. They feel they are 'making things up' and are unhappy about this. Although the numbers are lower here, this can be assessed as as the second most important issue in feedback because of its implications for confidence in the model generally, and for populating the other results worksheets, particularly the Factsheet.

- ***S4 interpretation***

This worksheet, on the other hand, is welcomed. It contains hard results from hidden, but undoubtedly scientific, calculations, and is clearly making a good impression. However, there is widespread concern over interpreting the results, and assessing whether a change in the result from one scenario to the next is significant or not. It might be useful, nevertheless, to introduce S4 ahead of S3 so as to allow S3 to be populated with hard data rather than subjective interpretation.

- ***S5 C and S5 B calculations***

There are several concerns about the costs deployed in worksheet S5 C, to the extent that some users challenge the amounts being suggested (although this can be overcome easily using custom values, it is an element in credibility). Users also want a facility to add in

additional costs not currently considered by the assessment, such as the removal of existing infrastructure, or the need for heavy plant. Maintenance costs in particular are challenged (although these can also be overridden), and the amounts flowing through to the Factsheet for maintenance are alarming. In sheet S5 B, some of the ES are deemed not monetizable, and some users think this should be challenged.

- ***Making use of Tab B***

The Shannon-Weaver Index is not widely understood. The user does not therefore know how to interpret the result or what to do with it, and although the worksheet says it is exported to the Factsheet, this actually doesn't happen.

- ***Scaling up***

Test G has highlighted that the model cannot cope with city-scale projects. What is not currently known is the safe working limit of the model in terms of scale, and some further exploratory work is needed in this area to inform the Guidance Document.

- ***Layout***

Layout issues do affect the operation of the model, but are generally easily dealt with. Some figures in S5 B are not aligned correctly, for instance. This does not affect the results, but is a highly visible error. Similarly, layout issues in the Factsheet charts may affect interpretation.

- ***Programming issues***

There are a small number if programming issues that need to be addressed. These result in errors in calculation, or even catastrophic failure of the model with the need to restart. These errors arise in several worksheets, but most affect the Factsheet in some way; the one in Tab A was the most prominent in the tests, but others are also vitally important.

3.5 Feedback and observations on the Factsheet

Many of the issues with the Factsheet are reflective of problems already identified through feedback on other worksheets. The list of correct/incorrect statements, pulled through from sheet S2 to present results for Cultural ES, is already noted as an issue that needs rectification within that worksheet, and is generally noted as something that, if included in the Factsheet, will not be read by a decision-maker and will tend to detract from an otherwise pithy summary of the results. Similarly, the astronomic costs presented in the Factsheet for maintenance are drawn through unchanged from worksheet S5 C where an annual maintenance figure is

replaced by a cost calculated over a much longer period. Resolving problems on other worksheets will help to improve the accuracy and content of the Factsheet as well.

The Factsheet has some elements missing, though. Food results are calculated but not reported in the Factsheet as it stands, and nor are the numerical figures for the health ES; both are calculated in S4 and could easily be presented here. Post-COVID, health is perhaps an especially serious omission.

There was some discussion among participants about showing the full list of ES on the Factsheet. At the moment, all ES are listed but those not selected on sheet S1 are 'greyed out'. Some liked the idea of showing a decision-maker that other ES had been available and considered, but were not chosen; others were concerned that their inclusion might prompt a decision-maker to identify additional ES they thought were important, setting unwanted hares running and delaying a decision.

Concerns arise about a visual impression given by the financial charts. These use different x-axes, necessarily because of the ranges of figures involved, but the result is potentially visually misleading, suggesting that small amounts in one chart are similar to much larger amounts in another chart. This could give a careless reader the impression that (for instance) annual monetary benefits outweigh maintenance costs, when in fact the figures are very different. Repositioning or recolouring the charts might offer a solution here.

The Factsheet also provides a space in which the author can offer some additional text. At the moment this is labelled 'Conclusion' but might be better left as free text where additional commentary can be added – for example, to add in a comment on non-GI interventions that add weight to the argument for the project, such as reducing the speed limit to further reduce noise, or using sustainable building methods to further reduce indoor temperatures. It might even be possible to suggest some additional content for this box in the Guidance Document.

Reviewing the Factsheet content has also prompted consideration about what a decision-maker would be looking for in this document. These focal points are suggested:

- What and where is this project?
- Why are we doing it?
- What will it achieve: main ES aim(s) and secondary results? How do options compare?
- Who will benefit, and how many of them are there?
- Is there public support for this approach?

- What will it cost (capital cost and annual revenue costs)?
- Is it cheaper than other options? And if not, is it better value?
- Can I trust these results?
- How does it deliver against the municipality's wider corporate objectives e.g. net zero, climate change goals, sustainability etc.?

This is not necessarily an order of service for the Factsheet, but could act as a useful checklist to ensure the Factsheet is answering the questions a decision-maker will have in mind as they consider a proposal. Some of the questions can be answered with results, but others will need a commentary from the Factsheet author, and it might be prudent to allow space for this and to suggest these questions as prompts in the Guidance Document.

4 Guidance

The Guidance Document was not ready in time to be included in the testing process, and in any event would have added significantly to the time required to carry out the tests. Nevertheless there is evidently a need for guidance to help the user navigate the Business Model, to populate those worksheets that require user data and to interpret the results that they produce. The maximum possible help was offered during the tests, and some participants made it clear that they would not have managed without this support.

There were also a couple of tests where the advance preparation was done by our consultant, and several others where participants recognised the fundamental importance of this. There were also two tests that had no preparation, and although this was overcome in one case, the other was a disaster as far as meaningful results are concerned. The Guidance Document needs to emphasise the importance of doing this groundwork, and might usefully provide a checklist of information the user is likely to need. Much of this is readily available, but may require the help of a GIS officer who needs to be forewarned, or may require an internet search where useful links can be provided. All of this will not only improve the results, but also strengthen the user's confidence that they know what they are doing and are getting it right.

At the moment, the draft Guidance Document provides help in two key areas: the academic rationale behind the Business Model content and calculations, and the practical help that the user needs to make the Business Model work. The material is organised by worksheet, so that the guidance explains the purpose of each sheet, includes some academic reference material, and then moves on to guide the user on the process of completing the sheet. Feedback from the tests suggests strongly that both these components are needed, but there is a significant practical difficulty with the current approach.

There is no doubt that some material describing the purpose of each worksheet, and especially the academic provenance behind it and its inner workings, are essential. Several users asked for this, and one went looking for it in the worksheets that lie beyond the Factsheet. However, it seems unlikely that this material would be referred to on a regular basis as the Business Model is being used, so linking it to the more practical material involved in data entry and interpretation makes this more difficult to find and generally less accessible to the user. Participants describing the kind of guidance they need identified a more cell-by-cell approach, with the guidance advising them what data is needed and which cell(s) to put it in, at a very simple level. The theoretical material is needed, but is much less useful when completing the model and can be separated out, to allow the cell-by-cell material to followed through as a series of uninterrupted steps. This has led to the conclusion that the Guidance Document either needs to be in two parts – one focussed on the practical, day-to-day use of the model

and the other providing the theoretical underpinning – or even perhaps two separate documents, each with a different focus. The project calls for a step-by-step methodology, and the model provides this, but the user also needs **step-by-step guidance** to allow them to navigate successfully.

This is not to diminish the **importance of provenance** in supporting the Business Model. Several participants noted the importance of this, not least in dealing with challenge from their more expert colleagues in other disciplines, or in reassuring a decision-maker, and some wanted the provenance to appear in the Factsheet. This is probably unnecessary, but the theoretical material that explains how the model is built should include a bibliography of the academic material used in its development. Even if the works themselves are unfamiliar to practitioners, such a list will help to reassure colleagues and decision-makers that the model is built on sound foundations.

The guidance can also play an important role in **interpreting the results**. This will be especially important in understanding sheet S4, which presents the results for quantifiable ES, where users need to be able to decide whether the differences between the results for different scenarios are significant or relatively minor. Even where the metric used is familiar – weight of food produce, for instance, or ambient temperature – the user will need help to understand whether a small difference is significant or not. Where the metric is less familiar – mass of carbon sequestered, or DALY (a health measure that links to morbidity) – help will be needed to understand the metric as well as the change in the metric from one scenario to another. Help will clearly be needed in understanding Tab B, the Shannon-Weaver index, and the Guidance Document can support this and also provide links to additional explanatory material online if the user needs it.

As an alternative to the cell-by-cell guidance, the possibility of a series of **short video tutorials** might be considered, taking one worksheet at a time and helping the user to complete their version by seeing how this is done and what is involved, for themselves. This approach could also be used to help the user understand the results that they see. One of the dry run examples would lend itself to this approach, as it includes multiple scenarios but is straightforward enough, and familiar enough as a concept, to be useful to a less skilled practitioner. YouTube is full of this kind of material which will be highly familiar to many users, and this would be a relatively inexpensive and cost-effective addition to a printed manual.

Thinking about the downloadable hard copy, **design** can play a very helpful role here. None of the development team is an expert in design, and it would be worth exploring the possibilities of using design techniques to make a manual easy to read and use.

The possibility of **translation** should also be explored; having guidance material in a familiar language may help a user cope with an English-language spreadsheet, so avoiding the inevitable complexities of traying to translate the model itself (something which would be incredibly difficult and costly to do, extremely error-prone, and would require extensive retesting in unfamiliar languages). In this context, it should be noted that some of the screenshots in the draft guidance are Apple-based and that the screens look different in MS Windows, which presents a challenge; the developers use Apple IT, but most local authorities use Windows as their operating system. It's also the case that some of the screenshots are in Dutch, because the underlying operating system used by the development team is itself in Dutch; as a result, some of the Excel-generated pop-ups that appear (for instance when a user tries to enter a value in a protected cell) are in Dutch, and these are the examples that appear in places in the guidance.

5 Other Observations

5.1 Help moving forward

While the internal testing and Demonstrator Testing have highlighted issues and allowed corrections and adjustments to be made to the Business Model, it is also the case that each new test throws up challenges and questions that may flummox anyone using it. Some of these can be anticipated, and addressed in the Guidance Document, but it seems likely that there will be unexpected or bizarre results or other issues that require support. It's also clear that some users are more confident in Excel than others are, and these users may struggle without a resource to turn to beyond the Guidance Document.

At present there is no provision for ongoing support for using the Business Model after it is signed off with Nature Smart Cities, leaving future users to their fate. A user forum would be a possible low-cost solution here but would depend on regular attention from stronger users.

5.2 Ongoing value of the Business Model

A key issue yet to be addressed in the development of the BM is that of post-project sustainability. Data in the BM may need to be updated periodically to reflect advances in knowledge surrounding GI, and financial data in the BM will need to be updated to reflect inflation and/or technological developments in both grey and green approaches. At present no provision has been made for this, either in the work programme of partners or in financial support for this work. If no provision is secured, the data in the software will become less and less valid, and the credibility of the BM will begin to be eroded. The partnership may wish to give this some thought with a view to perhaps securing some ongoing funding for maintenance. A subscription model is unlikely to work and may not be permitted by Interreg.

5.3 Version control

At present the mechanism to acquire the Business Model is to download it in Excel format from the NSCiti2S website. However, having once downloaded it, there is no guarantee that a returning user will check to see if there is a more recent version that supersedes the one they already have. An alternative approach would be to set up the model to be accessed and stored online using a user account, which would mean that data was securely stored but also that only the most up to date version of the model was available. This could also help with ongoing licensing (see below).

5.4 Promotion

A phenomenal amount of work has gone into the Business Model, in underpinning research, software design and development, programming, testing and revision. More work will take place after receipt of this report. But it is noted that, aside from those already 'plugged in' to the Nature Smart Cities network, or drawn in through the capacity building programme, there will be very limited promotion of the Business Model, especially after NSCiti2S itself draws to a close. It would be a huge anti-climax, and a loss of significant skill, expertise and effort, if the model were simply to become yet another largely unused green infrastructure tool. Capacity building will certainly help here but other avenues of publicity might also prove fruitful. It is therefore strongly suggested that means of making it more widely known should be explored and discussed by the partnership.

5.5 Conditions of use

Although the BM is intended for free use by municipalities, the developers and the partnership still need to be protected against any consequences of an error or omission in the software, inadvertent or deliberate changes to the underlying programming, or the misuse of the software beyond its intended purpose. Even though the Guidance Document makes it clear that the model should be regarded as advisory only, it is possible that a municipality relying on the results produced in the software might make an investment decision based on this, which turns out to be flawed and causes financial and/or reputational damage. They may then seek to hold the developers/partners legally responsible for this. At the very least a disclaimer is needed to transfer responsibility to the user – in much the same way as other software does – requiring the user to accept the terms and conditions before they can use the software. A license agreement which a user has to accept to gain access to the model would be an alternative, and more legally secure, solution, and is now normalised in gaining user access to proprietary software.

Licence models are available online, but the language will depend on the expected nature of use, the freedom to be allowed to modify the underlying code or import it into new models, and the country or countries whose legal system is to be used as the basis for any legal action. It is recommended that specialist professional advice is sought.

5.6 Language

Language has emerged as a challenge for the Business Model. At present it is in English, as is the prototype Guidance Document. But the NSciti2S project covers a geographical area in which two other languages are spoken, and informal feedback has encouraged the

development of at least a Dutch version, which would be understood in Belgian Flanders as well as in the Netherlands. Had a test been conducted in France, a French version would probably have been requested as well, and this would also extend the currency of the Business Model into the French-speaking areas of Belgium (although these are outside the 2 seas/mers/zeeën area). Informal feedback from elsewhere also suggests that, while English is widely understood in the Netherlands (and perhaps this is also true in Flanders), those who work for smaller local authorities are less likely to be fluent or confident in using English, and this may discourage them from trying out the Business Model for themselves. These are, of course, the very municipalities that the model is attempting to support.

Appendix 1: Scoring criteria and weights

Scoring group	Component	Weight	Score	Result
Useability	Need for specialist knowledge	3		
	Availability of necessary input data	2		
	Work needed to generate input data	1		
	User confidence in working with the tool	3		
	Need for support and direction	1		
Comprehensiveness	Coverage of the range of landscape and other types in the project	1		
	Coverage of the ES sought from the project	1		
	Generation of the information of greatest importance to the user	3		
	Asking the right questions and not overlooking important dimensions	3		
	Handling comparatives successfully	2		
Credibility (and clarity)	Generating believable results	3		
	Using the right language and terminology	1		
	Accessibility of results to decision-maker	2		
	Visuals that answer questions clearly and unambiguously	1		
	User confidence in explaining/interpreting results to non-expert	3		

Interreg
2 Seas Mers Zeeën
European Regional Development Fund

NATURE SMART CITIES
ACROSS THE 2 SEAS

www.naturesmartcities.eu