A planning-support tool for spatial suitability assessment of green urban stormwater infrastructure

Martijn Kuller a, b, *, Peter M. Bach a, c, d, Simon Roberts e, Dale Browne e, Ana Deletic b

a Monash Infrastructure Research Institute, Department of Civil Engineering, Monash University, Clayton, 3800, VIC, Australia
b School of Civil and Environmental Engineering, University of New South Wales Sydney, 2052, Australia
c Swiss Federal Institute of Aquatic Science & Technology (Eawag), Überlandstrasse 133, 8600 Dübendorf, Switzerland
d Institute of Environmental Engineering, ETH Zürich, 8093 Zürich, Switzerland
e E2DesignLab, Carlow House, 289 Flinders Lane, Melbourne, VIC 3000, Australia

HIGHLIGHTS

• Current WSUD planning is ad-hoc, causing sub-optimal functioning of systems and uncapitalized potential benefits;
• We developed a spatial suitability model (SSANTO) to inform placement of WSUD, assessing opportunities and needs;
• Besides biophysical, socio-economic and planning aspects, SSANTO brings knowledge from ecosystem services into WSUD planning;
• SSANTO facilitates sophisticated Geo-Information Science-based Multi-Criteria Decision Analysis;
• Testing of SSANTO shows its usefulness to robust WSUD planning/implementation;
• SSANTO enables deep stakeholder engagement with urban planning focussing on opportunities and needs of WSUD.

GRAPHICAL ABSTRACT

ABSTRACT

Distributed green stormwater management infrastructure is increasingly applied worldwide to counter the negative impacts of urbanisation and climate change, while providing a range of benefits related to ecosystem services. They are known as Water Sensitive Urban Design (WSUD) in Australia, Nature Based Solutions (NBS) in Europe, Low Impact Development (LID) in the USA, and Sponge City systems in China. Urban planning for WSUD has been ad-hoc, lacking strategy and resulting in sub-optimal outcomes. The purpose of this study is to help improve strategic WSUD planning and placement through the development of a Planning Support System. This paper presents the development of Spatial Suitability Analysis Tool (SSANTO), a rapid GIS-based Multi-Criteria Decision Analysis tool using a flexible mix of techniques to map suitability for WSUD assets across urban areas. SSANTO applies a novel WSUD suitability framework, which conceptually spatial suitability for WSUD implementation from two perspectives: ‘Needs’ and ‘Opportunities’ for WSUD. It combines biophysical as well as socio-economic, planning and governance criteria (‘Opportunities’) with criteria relating to ecosystem services (‘Needs’). Testing SSANTO through comparing its results to work done by a WSUD consultancy successfully verified its algorithms and demonstrated its capability to reflect and potentially enhance the outcomes of planning processes. Manual GIS based suitability analysis is time and resource intensive. Through its rapid suitability analysis, SSANTO facilitates iterative spatial analysis for exploration of scenarios and stakeholder engagement.
1. Introduction

Distributed green stormwater control measures are increasingly applied in cities around the world. They are primarily designed to protect surface water quality, mitigate flood risk and increase resilience of urban drainage systems in response to the challenges posed by urbanisation and climate change. These nature-based stormwater control measures are commonly referred to as Water Sensitive Urban Design (WSUD) in Australia and this paper. WSUD also provides benefits including amenity and recreational values, mitigation of urban heat, an alternative source of water provision and habitat for increased biodiversity (e.g. Ashley et al., 2017).

Strategic planning approaches are critical for the spatial allocation of WSUD to suit their physical and social urban context, while optimising the benefits derived (Thévenot, 2008). However, current planning of WSUD is often the result of opportunistic and ad-hoc decision-making processes, which is reflected in its current spatial distribution (Kuller et al., 2018a). These may result in less than optimal outcomes for both infrastructure operation and service delivery. Consideration of multiple criteria is essential to respond to both the multi-faceted nature of WSUD and the urban environment it is integrated into. The application of planning support systems (PSS) may significantly improve WSUD planning outcomes through their capacity to combine, analyse and present diverse spatial information in a format that is meaningful to stakeholders (Geertman and Stillwell, 2012; Klosterman, 1997). They can be used to promote collaborative planning and strategic decision-making. A plethora of PSS, models and tools have become available to the WSUD planner over the past two decades, such as SUSTAIN-EPA (Lee et al., 2012), SUDSLOC (Ellis and Viavattene, 2014) and BeST (Digman et al., 2015). However, recent reviews of tools and models for WSUD planning still conclude the current models and tools are insufficient (Kuller et al., 2017; Lerer et al., 2015). These reviews indicate tools need to be (i) more spatially explicit, (ii) broader in scope (in terms of technologies), (iii) more comprehensive (in terms of assessment criteria) and (iv) more rigorous. Furthermore, a recent study of causes behind a lack of uptake for PSS in urban planning – referred to as the ‘implementation gap’ (Vonk, 2006) – reveals these are also present in WSUD planning (Kuller et al., 2018b). Some of the most important barriers to the adoption of such tools include their (i) lack of user-friendliness, (ii) required time and effort to run, (iii) complexity, (iv) lack of industry standards and (v) inability to produce relevant outputs. Perhaps the most significant cause of these shortcomings is a lack of engagement between PSS developers and the planning practice.

Engagement with WSUD planning practice suggests a specific need for spatially explicit tools. Kuller et al. (2017) specifically review recent tools that integrate Multi-Criteria Decision Analysis (MCDA) with spatially explicit algorithms using Geographical Information Systems (GIS), which have the potential to benefit planning (Massam, 1988). GIS-MCDA offers capabilities to integrate the complexities emerging from both technical and social perspectives, such as the integration of social, environmental and economic factors as well as consideration of non-monetary values (Ferretti and Montibeller, 2016). Complex, multi-disciplinary, multi-stakeholder and group decision-making processes are facilitated through techniques offered by GIS-MCDA (Boroushaki and Malczewski, 2010; Jankowski and Nyerges, 2001). The assessments and tools currently available have at least one and usually more of three limitations, including their:

(i) insufficient or incomplete number of assessment criteria,
(ii) insufficient methodological sophistication (e.g. spatial explicitness, consideration of preferences, combination rules) and
(iii) lack of automation and reproducibility. In particular, integrated tools that link social and technical factors are lacking in literature and practice, as discussed by Prudencio and Null (2018).

In response to these limitations, this study aims to develop a WSUD planning support system that automates user(s)-driven spatial suitability assessment for the planning of green stormwater control systems in urban environments. We specifically focus on the following objectives:

(a) Operationalising a novel and comprehensive WSUD suitability framework proposed by Kuller et al. (2017), which identifies two sides of suitability: WSUD needs a place and a place needs WSUD. The latter is generally overlooked in current WSUD PSS. Operationalisation is achieved by coupling the indicators related to the criteria to spatial datasets;
(b) Advancing spatial WSUD suitability mapping by applying a mix of GIS-MCDA techniques on the above criteria, in a flexible and replicable manner;
(c) Developing a spatial software supporting strategic decision-making by integrating the above in alignment with practitioner insights (Kuller et al., 2018b);
(d) Testing the PSS on an existing real-world case study in Melbourne, where suitability analysis was performed previously by industry stakeholders.

By utilising a novel suitability framework, for the first time, ecosystem services and the community needs for green infrastructure, as well as suitability related to the biophysical, social and urban context underlying WSUD performance are systematically incorporated into a WSUD PSS. The strength of the presented tool lies in its capability to automate advanced spatial MCDA techniques, which normally requires considerable time and human resources. In doing so, the tool can generate easily interpreted output, thereby facilitating deeper collaboration between stakeholders.

2. Development of SSANTO

The Spatial Suitability ANalysis TOol (SSANTO), which automates spatial suitability assessment for the planning of green stormwater control systems in urban environments, was developed to meet four objectives:

(a) Presents an easy-to-use interface that enables use by experts as well as non-experts and practitioners
(b) Performs quick but rigorous assessment of the complete array of relevant factors in a spatially explicit way
(c) Combines opportunity assessment with the spatial assessment of needs, using principles from ecosystem services
(d) Produces ready-to-use intuitive outputs that can be interpreted by experts as well as lay people.

SSANTO can perform individual suitability analyses for seven of the most commonly implemented WSUD infrastructure types: (1) Bioretention & rain gardens, (2) Infiltration systems, (3) Green roofs, (4) Ponds & lakes, (5) Swales, (6) Rainwater tanks, and (7) Constructed wetlands. The tool’s flexible architecture allows for easy extension to include additional infrastructure types. SSANTO was built as a Python add-in to the spatial software ArcMap by Esri, connected to several python toolboxes and coded in Python version 2.7. SSANTO’s analysis is raster-based, using a customisable cell size of 20 × 20 m.
Outputs include individual criterion suitability maps, an opportunity map, needs map and overall suitability map. Furthermore, summary statistics for the study area, as well as detailed background information on local suitability results can be interactively queried on the output maps (for more details, see Appendices A and B).

SSANTO does not consider system size of WSUD, catchment hydrology, treatment trains or quantitative and qualitative performance. As such, it is meant to be used in conjunction with other models such as MUSIC (eWater, 2011) for design, UrbanBEATS (Bach et al., 2013) for options generation or even an agent-based model for green infrastructure uptake by Castonguay et al. (2018), which explicitly requires some form of suitability maps as inputs.

2.1. Methodological background

SSANTO operationalises the WSUD suitability framework developed by Kuller et al. (2017), as presented in Fig. 1. To 'measure' suitability, this framework starts from the notion of 'two sides of suitability':

- 'Opportunities' (referred to as 'WSUD needs a place'); and
- 'Needs' (referred to as 'A place needs WSUD').

Opportunities describes favourability of locations for the implementation of green stormwater infrastructure based on the biophysical, socio-economic as well as planning & governance context.

Needs describes locations based on their need for the benefits derived from WSUD related to provision (e.g. irrigation water harvesting), regulation (e.g. water quality), cultural values and ecological habitat. Each category contains several suitability factors (Fig. 1).

SSANTO allows the user to build suitability maps for both sides of suitability using a four-step procedure adapted from Malczewski and Rinner (2015): (1) compiling geodatabase, (2) masking, (3) value scaling, and (4) combination rules (Fig. 2). Below, each step is introduced, while their implementation in SSANTO is detailed in Appendix A.

2.1.1. Compiling geodatabase

All relevant spatial datasets are compiled into a geodatabase. Spatial data corresponding to each measurable indicator from the suitability factors in the framework are sought. Spatial MCDA tools like SSANTO are typically data-hungry (Lee Jr, 1973) and their application may suffer from a lack of readily available data (Ferretti and Montibeller, 2016). In certain cases, alternative datasets can be used to measure an indicator (e.g. Rhea et al., 2014). For example, landfill and petrol station sites can act as a proxy to the indicator 'soil contamination'. Proxies should be applied with caution, as their relation to the actual indicator may not always be straightforward (Marttunen et al., 2018).

A pragmatic approach was used to compile SSANTO's geodatabase, selecting priority indicators from existing datasets available for our case-study (Table 1).

2.1.2. Masking

Masking is the process of removing all areas where at least one aspect of the urban context constrains infrastructure implementation. Two types of masking are possible, depending on the type of data: (1) Boolean masking and (2) masking using a threshold value. The former is used for discrete data (e.g. features) while the latter is used for continuous data (e.g. slope).

2.1.3. Value scaling

Before combining and comparing diverse criteria, we transform raw spatial datasets to a common suitability scale (Malczewski and Rinner, 2015) through value scaling. This process answers questions of the following kind: What does a slope of 5.5% mean for the suitability of a location for the implementation of rain gardens? Value scaling is an essential step in GIS-MCDA, and a mathematical representation of human judgement and knowledge in the form of 'value functions' (Keeney, 1992). Value functions describe the relationship between raw data values and suitability values, thus representing various datasets in comparable units. The shape of a value function is unique to its corresponding criterion and WSUD system type. For simplicity, linear value functions are commonly applied (Malczewski and Rinner, 2015). However, Stewart (1996) showed the shape of value functions.

---

Fig. 1. Suitability framework (simplified) adapted from Kuller et al. (2017). Both sides of suitability comprise of categories (in rectangles with icons) and their factors.
matter, and that using a linear form is often an over-simplification of reality. As an alternative, SSANTO applies value functions of a piecewise linear form, which can account for non-linearity (Malczewski and Rinner, 2015; Pereira and Duckstein, 1993). Thus, suitability maps for each criterion are produced using a common suitability scale from 0 (least suitable) to 100 (most suitable).

### 2.1.4. Combination rules

This involves combining all criteria. As not all aspects carry the same importance for final suitability, criteria need to be weighted. Weights can be elicited from stakeholders, or calculated. Three weighting methods are currently adopted in SSANTO. **Hierarchical weighting** is commonly applied for decision problems that can be divided into sub-objectives, such as the suitability framework used by SSANTO. Weight definition can be the source of biases, which are discussed by Marttunen et al. (2018). According to them, hierarchical weighting suffers most from splitting and asymmetry biases as well as higher variance, where user weights are affected by the structure of objectives and sub-objectives (i.e. criteria) in a branch of a hierarchy. **Non-hierarchical weighting** may suffer from range insensitivity and equalising biases. The former occurs when the range of possible criterion values is insufficiently reflected in weights while the latter occurs because of the user’s tendency to avoid assigning very high or low weights. **Entropy-based weights** are calculated weights which reflect the information-density of the data (Shannon and Weaver, 1963). Lower entropy means higher variation in the data, more discriminative power and therefore higher criterion weights. Fully or partly (i.e. combined with another weighting method, as optional in SSANTO – Hwang and Yoon, 1981b) entropy-based weights mitigate biases associated with user-defined weighting and can be useful when combined with other weighting methods (Nijkamp and van Delft, 1977). It is important to note that entropy-based weighting results in output maps with more emphasis on ‘relative suitability’, i.e. suitability of a location is relative to the suitability of all other locations in the studied area. User-defined weights, on the other hand, result output with emphasis on ‘absolute suitability’, i.e. the suitability of a location based on a global standard rather than compared to other locations in the studied area. Appendix H presents a discussion on this phenomenon.

According to Ferretti and Montibeller (2016), the inherent subjectivity of user-defined criteria weights in GIS-MCDA necessitates participatory processes with stakeholders. They argue the rating method (Malczewski and Rinner, 2015) adopted by SSANTO responds to the observed need for simple weight elicitation protocols. This is contrasted by the widely used ‘pairwise comparison’, which has been criticised for inflated spread of weights and inconsistencies (Lienert et al., 2016; Malczewski and Rinner, 2015) and may require users to redo the entire process. The simplicity of the rating method comes at the cost of certain biases (Eisenführ et al., 2010). Therefore, the flexible architecture of SSANTO allows the addition of other weighting methods in the future, to give users more options.

### Table 1

Summary of datasets used in SSANTO for both sides of suitability (Fig. 1) with corresponding indicator and category from suitability framework. Complete table of datasets used in SSANTO, including data sources and further explanation can be found in Appendix C.

| Category               | Indicator              | Dataset                  | Data type   | Proxy |
|------------------------|------------------------|--------------------------|-------------|-------|
| Opportunities          | Biophysical            | Slope                    | Digital Elevation Model (DEM) | Raster |
|                        |                        | Roof areas               | Building footprints | Features |
| Socio-economic         |                        | Environmental awareness  | Election results | Aggregated |
|                        |                        | Sense of community       | Volunteering   | Aggregated |
| Planning & governance  | Utility infrastructure  |                          | Easements     | Features |
|                        | Cultural Heritage      |                          | Cultural Heritage Sites | Features |
|                        | Land ownership         |                          | Cadastre      | Features |
| Needs                  | Provisioning           | Proximity to water demand| POI (Points of Interest) | Features |
|                        | Regulating             | Heat vulnerability       | Heat vulnerability | Aggregated |
|                        |                        | Connected impervious area| Directly connected imperviousness (DCI) | Aggregated |
|                        |                        | Floods                   | Flood extents  | Raster |
| Cultural               | Visibility             |                          | Land-use, POI | Features |
After weighting, a model is applied to combine all criteria and create a suitability map. The weighting method and combination rules are closely related to each other. Malczewski and Rinner (2015) describe a range of methods varying in complexity, ranging from simple linear additive models to complex and non-linear ideal point and outranking methods. One of the most widely applied methods is weighted linear combination (WLC). As an intuitive method to decision-makers, this method was chosen for its simplicity. WLC assumes linearity (constant marginal values) and additivity (mutual preferential independence). Although the assumptions behind WLC are not easily applied in spatial decision problems, it was found to perform almost as well as far more complicated, non-linear methods such as reference-point methods (Hwang and Yoon, 1981a) and can be easily implemented in GIS using map algebra (Tomlin, 1990). The equations from SSANTO’s combination rules algorithm are presented and explained in Appendix A.

SSANTO’s architecture is flexible and allows for iterative analyses to respond to some of the key challenges of GIS-MCDA as identified by Ferretti and Montibeller (2016) and discussed in this section. Because of the nature of GIS-MCDA, a certain level of uncertainty and bias is unavoidable. Clear reporting on the effect of model choices to the user is therefore critical. This is achieved through clear user guidance and default settings that minimise unwanted biases.

2.2. Testing SSANTO

This section describes the case-study location, E2D’s methods and outputs, SSANTO’s setup for testing and the method for comparing E2D’s outputs with SSANTO’s outputs. The notion that quantifying uncertainty in socio-technical systems and the models that represent their behaviour is an unresolved academic challenge (Van Asselt and Rotmans, 2002) is valid to date. The modelled phenomena in these models (e.g. suitability) often are not physically (objectively) measurable metrics. Suitability reflects a stakeholder’s preference and expertise, and estimates depend on the applied definition of suitability. Such estimations are scarce due to high investments required to generate them.

Guidance from literature is limited, as systematic reporting on testing and validation is largely absent for socio-technical models. Studies that do include testing and validation generally adopt one or a combination of three approaches from a wider research field. Firstly, comparing model outputs to those of established models, as applied in e.g. flood modelling (e.g. Guidolin et al., 2016; Jamali et al., 2018). Secondly, comparing model outputs to historical data, as applied in e.g. scenario modelling (e.g. de Haan et al., 2016; Ozturk and Batuk, 2011). Thirdly, qualitative evaluation of model outputs using expert and stakeholder opinions elicited through structured workshops, as applied in social sciences. Established comparable models as well as historical data were unavailable for testing of SSANTO. Following a scarce number of examples from literature (e.g. Bach et al., 2015), we adopted a combination of the first and third approach for testing by comparing SSANTO’s outputs to recent and high-quality consultancy case-study results, which were available for a municipality in Melbourne.

Testing of SSANTO aimed at (i) ensuring mathematical correctness of the algorithms (verification) and (ii) ensuring practical correctness of the model to produce similar outcomes to in-depth manual suitability studies (validation). To warrant the meaningfulness of the latter test, value scales were developed independent of the weight assignment. SSANTO’s outputs were compared with the outcomes of a suitability mapping and prioritisation study carried out by a consultancy firm called E2Designlab, hereafter referred to as E2D. We acknowledge that the testing adopted for SSANTO is far from perfect. However, considering the highly contested nature of socio-technical model testing and validation, the methodology followed in this paper is considered the best available.

2.2.1. Case study description

The City of Darebin is a local government area in the suburbs directly northeast of the centre of Melbourne, Australia. Its population is just under 150,000 in an almost fully built-up area of 53 km² (ABS, 2016). Darebin is situated in the Lower Yarra Catchment with Merri Creek forming its western border and Darebin Creek forming its eastern border (Fig. 3). Natural waterways in Darebin have been degraded due to urban development and communities are facing occasional flood events.

2.2.2. Validation data

Melbourne-based WSUD consultancy E2D performed a spatial prioritisation study for Darebin in 2017 (Roberts et al., 2017), for two cases: (1) street-scale systems such as small rain gardens and tree pits, and (2) precinct-scale WSUD options such as wetlands, large rain gardens and stormwater harvesting schemes that treat moderate to large catchments (from 10 ha to over 100 ha). In the first case, values from four spatial datasets selected by E2D (all part of ‘Needs’) were assigned a suitability score between 1 (least suitable) and 3 (most suitable). E2D then overlaid the maps by summing up the suitability scores, to generate a final vector-based suitability map with scores between 4 and 12. This created a coarse suitability map without mask, intended as a high-level aid to decision-makers for streetscape system planning. For the purpose of testing, this map was first normalised to represent scores from 0 to 100, and subsequently rasterised.

In the second case, 68 priority sites (available as a polygon shapefile) were manually identified as opportunities to retrofit precinct-scale systems (hereafter called: ‘E2D priority sites’). E2D priority sites were identified through GIS screening, masking and a manual prioritisation exercise, using E2D’s experience with planning for green stormwater infrastructure combined with spatial data from 12 criteria and additional local knowledge. More detail about the selection process can be found in Appendix E as well as Roberts et al. (2017). It should be noted that E2D priority sites were created using detailed information about the urban context as well as tacit knowledge of the urban planner and political motivations, all of which are not always available or reflected in the form of spatial datasets.

2.2.3. Tool setup

A total of seven model runs were performed to validate and test SSANTO, as presented in Table 2. One model run analyses suitability for street-scale systems (rain gardens in SSANTO), using identical weighting and value scales as E2D (S-Case-Expert). Six model runs analyse suitability for precinct-scale systems (wetlands in SSANTO), using two sets of criteria (‘case-limited’ and ‘full’) and three different weighting methods (‘equal’, ‘expert’ and ‘entropy’). The case-limited criterion set is a selection of criteria as used in the study by E2D, while the full criterion set includes all relevant criteria as defined by Kuller et al. (2017) for which data was available. A complete overview of these criteria is presented in Table 3. Weighting method ‘equal’ refers to a model run where all criteria have equal weights, while for ‘expert’, weights were provided by E2D. The ‘entropy’ weights are calculated from the variability of data within a criterion, as explained in Appendix A.

More information about the criteria from Table 3 and all other criteria used by SSANTO for other WSUD types, their data sources, as well as value scales applied to each criterion and WSUD type can be found in Appendices C and D respectively. Expert value scales for model run S-Case-Expert as well as expert weights for precinct-scale systems can be found in Appendix F and Table 3 respectively.

2.2.4. Comparing SSANTO’s outputs to suitability results generated by E2D

For model run S-Case-Expert, the normalised suitability map from SSANTO’s setup for testing and the method for comparing SSANTO’s outputs, SSANTO’s combination rules algorithm are presented and explained in Appendix A.
SSANTO’s suitability map is prevented by the incompatibility of output types. To assess whether the outputs of SSANTO reflect E2D suitability site selection, SSANTO’s calculated suitability at E2D priority sites was compared with the average calculated suitability for Darebin. Statistical significance of the differences in mean suitability between Darebin average and E2D priority sites was tested using the Mann-Whitney U test (Mann and Whitney, 1947), as the normality assumption for a student t-test could not be fulfilled for any of the cases. The suitability maps for the three weighting methods were assessed in greater detail for model runs using full criteria. Finally, the output for ‘Opportunities’ and ‘Needs’ were compared for model run P-Full-Expert by comparing their suitability values on a cell-by-cell basis.

3. Results & discussion

Setting up and running a suitability analysis in SSANTO can be completed within an hour by a trained user. This can be considered rapid compared to manual GIS-MCDA exercises which can take days or even weeks to complete. Total runtime of the different algorithms for a simulation is under 30 min and depends on the user preferences and parameters chosen.

3.1. Street-scale systems

Verification resulted in a 96.7% match for street-scale systems (identical cell values between the output generated by SSANTO compared with E2D), 2.65% of area had a deviation of 1, and only 0.61% deviated more than 1 (Fig. 4). While the deviation of 1 can be attributed to rounding errors, deviations above 1 are likely due to misalignment of cells resulting from the transformation of vector data to raster data. The close match between the modelled output and the output from E2D is promising, as it demonstrates that SSANTO can reproduce the technical steps and thinking required to undertake this analysis in an algorithmic and automated way.

3.2. Precinct-scale systems

Results of the model run using full criteria and expert weights as provided by E2D are shown in Fig. 5. In total, 72% of the case-study area is masked out. SSANTO’s suitability scores are relatively higher (p < .000) at E2D priority sites with a mean of 69, compared to 64 in the remainder of the unmasked study area (Fig. 5c). This difference is greater for ‘Opportunities’ (62 vs. 55, p = .000) than for ‘Needs’ (79 vs. 78, p < .000). This suggests that the ‘Opportunities’ side of suitability was more important in the selection of E2D priority sites than the ‘Needs’ side.

Fig. 6 presents the performance of all six model runs, through a full comparison of the suitability distributions. Model runs using full criteria result in a higher positive difference in mean between E2D priority sites and Darebin average (i.e. better fit) than model runs using limited criteria (mirroring E2D’s analysis). This suggests that E2D and the City of Darebin implicitly used more information than the indicators they reported to inform their E2D priority site selection, including some of the information that SSANTO uses. Thus, some of the tacit knowledge used for E2D priority site selection (for which the geographic data may not have been obtained at the time) was potentially captured by SSANTO’s full criteria set, supporting the choice to use a comprehensive criteria set.

For the case-limited set of criteria, expert and entropy weighting perform better than equal weighting. For model runs using the full criteria set, entropy weighting results in a significantly better fit than the other weighting sets, while expert weighting results in a slightly better fit than equal weighting.
The full comparison of key suitability summary statistics between E2D priority sites and Darebin average for ‘Needs’, ‘Opportunities’ and combined suitability for all scenarios is presented in the table of Appendix G. From this table we observe that ‘Needs’ fitted E2D priority sites better than ‘Opportunities’, suggesting those criteria from the case-limited set were more closely considered by E2D. This order is reversed in tool runs with a full set of criteria, where the fit for ‘Needs’ is lower than that for ‘Opportunities’.

The outperformance of expert weighting by entropy weighting for the full criteria set is notable, considering experts weights were elicited from the same people who selected the E2D priority sites. These results highlight the complexity of accurate weight elicitation as it suggests we underestimated the importance of some and overestimated the importance of other criteria from the planning process. It furthermore points to the potential strength of entropy weighting, as P-Full-Entropy was found to be the best performing model run. Part of the explanation may lay in the fact that the decision to implement precinct-scale systems (predominantly wetlands)

had been taken before E2D undertook the spatial analysis, and therefore relative suitability took priority over absolute suitability (see Appendix H for further discussion). For a full comparison between expert- and entropy weights, refer to Table 3.

Fig. 7 compares the results for expert (b) with equal (a) and entropy weighting (c). While suitability values vary between the three outputs for certain locations, other locations appear more stable (Fig. A6 of Appendix G). The spread of suitability values across Darebin is highest for entropy weighting and lowest for equal weighting, where 97.7% of suitability values lie between 50 and 70 (see also: Fig. 6). The small number of criteria that dominate the suitability result for entropy weights (Table 3) could explain the high spread compared to expert and equal weighting, where suitability results are evened out by a greater number of influential criteria.

Finally, it is notable that in most locations, the score for ‘Needs’ is higher than that for ‘Opportunities’ (Fig. A7 of Appendix G). This could indicate that either (i) there is a bias towards ‘Needs’ in value scales or (ii) green infrastructure is needed in more locations than it can be implemented.

### 4. Implications to practice

SSANTO brings several critical improvements and innovations to the existing suite of GIS-MCDA studies. SSANTO is one of few spatial suitability analysis tools for WSUD planning (Kuller et al., 2017). It is the first tool to adopt a comprehensive set of criteria from both the needs and opportunity sides of suitability, considering the multiple benefits derived from WSUD by incorporating ecosystem services into its underlying framework. It combines this with a selection of advanced GIS-MCDA techniques to accurately represent suitability, where previous studies tend to focus only on comprehensiveness of criteria (e.g. Hirschfeld et al., 2005; Viavattene et al., 2008) or sophistication of

---

**Table 2**

Overview of model runs performed for testing.

| WSUD type | Criteria set | Weighting method | Model run |
|-----------|--------------|------------------|-----------|
| Street-scale | Case-limited | Expert* | S-Case-Expert |
| Precinct-scale | Case-limited | Equal | P-Case-Equal |
|            | Entropy      | Expert       | P-Case-Expert |
|            |             | Entropy      | P-Case-Entropy |
| Full       | Equal       | P-Full-Equal |
|            | Entropy      | P-Full-Entropy |

* Both weighting and value scales used by E2D were also applied for this model run. The value scales can be found in Appendix F.

**Table 3**

Criteria used for testing and their respective precinct-scale expert and entropy weights.

| Category               | Criterion                        | System type* | Mask (threshold) | Precinct-scale weights |
|------------------------|----------------------------------|--------------|------------------|------------------------|
| WSUD needs a place     | Biophysical                      | Slope        | ✓ S, P           | Expert 41 Entropy 30  |
| (opportunities)        | Building footprints              | ✓ S, P       | –                |
|                        | Pre-human wetland structureb     | ✓ S, P       | –                |
|                        | (Distance to) landfill sites     | – S, P       | 30 28            |
|                        | Sports fields                    | ✓ S, P       | –                |
| Socio-economic         | Education level                 | S, P         | 10 7             |
| Planning & governance  | Sense of community               | S, P         | 20 7             |
|                        | Environmental awareness          | S, P         | 20 15            |
|                        | Distance to drainage infrastructure | S, P       | 90 1             |
|                        | Utility infrastructure           | S, P         | 70 19            |
|                        | Land value                       | S, P         | 20 5             |
|                        | Street width/type                | S ✓          | –                |
|                        | Lot Size                         | P –          | 50 100           |
|                        | Land ownership                   | S, P         | 70 30            |
|                        | Cultural Heritage                | S, P         | 30 36            |
|                        | Geological Heritage              | S, P         | 20 14            |
|                        | Natural Heritage                 | S, P         | 30 39            |
|                        | Distance to airports             | ✓ P (1000 m) | 60 3             |
| A place needs WSUD     | Provisioning                      | Irrigation demand | S, P | 80 12 |
| (Needs)                | Regulating                       | Effective imperviousc | S, P | 100 – |
|                        | Total impervious                 | ✓ S, P       | 80 25            |
|                        | Current WSUD                     | ✓ S, P       | 90 0             |
|                        | Heat vulnerability               | S, P         | 60 11            |
|                        | Flood risk                       | S, P         | 30 5             |
|                        | Visibility                       | S, P         | 60 100           |
|                        | Social cohesion                  | S, P         | 40 37            |
|                        | Green cover                      | S, P         | 60 20            |
|                        | Recreation                       | P –          | 40 4             |

* S: street-scale systems such as tree pits and rain gardens, P: precinct-scale systems such as constructed wetlands and large bioretention. In bold are the criteria used by E2D and applied in the ‘case-limited’ model runs.

b Pre-human wetland structure was excluded from the analyses as no overlay features are present for the case study location.

c Effective imperviousness was excluded from the evaluation for the Darebin case study, as this fully developed area has homogeneously high rates of effective imperviousness, far above the threshold for which WSUD could improve water quality.
methodology (e.g. Ozturk and Batuk, 2011; Rahman et al., 2012). Furthermore, SSANTO is the only GIS-MCDA tool for WSUD planning that has been tested against industry practice. Finally, where many GIS-based MCDA studies are either single efforts, frameworks for analysis or tool development (e.g. Chow et al., 2014), SSANTO is one of very few fully functional automated software tools enabling repeated suitability assessment.

Although SSANTO is currently in the research application phase, it presents a promising opportunity to practice for solving prominent planning challenges. Thus, SSANTO has the potential to become valuable to urban planners if several challenges are overcome, which are discussed in the remainder of this section. Furthermore, we make some suggestions for proposed further development of SSANTO in the future.

4.1. Operation of SSANTO

As is inherent to many types of models and tools, the availability, quality and format of input data are fundamental to SSANTO’s operation. This principle is often referred to as “garbage in, garbage out” (Eysenck, 1978). The most important data-related issues include:

- data quantity: many criteria and related datasets are required, including biophysical, socio-economic and planning-related data (Alexander, 1989);
- data quality, accuracy and collection date (how recent the data are);
- fuzziness of the relation between data/indicator and decision criterion (Chen et al., 2011; Malczewski, 1999);
- high variety of data formats and resolutions related to the different types of data (Openshaw, 1983).

The nature of data inputs, i.e. socio-economic and urban form data, makes SSANTO specifically applicable to infill development modelling. Sufficient biophysical data in combination with detailed statutory planning data can warrant SSANTO useful for greenfield developments as well. In cases where master plans are detailed enough, certain ‘virtual’ datasets can serve as criteria in the planning & governance, provision, regulating, cultural and habitat categories. Limitations related to data are described in greater detail in the Appendix H.

Different weighting methods are associated with different advantages, biases and limitations, as described in Section 2.1. Varying data formats were found to have an impact on entropy weights, potentially compromising their validity. Further discussion on the power of entropy-based weights to remove biases associated with user-based weighting (Boroushaki, 2017) as well as the limitations related to data formats can be found in Appendix H. Clear communication of uncertainties relating to data and weighting is essential to enable appropriate, mindful interpretation and application of the outputs (Walker et al., 2003). Because user-defined weights reflect preferences and expertise, suitability is a human concept, which is not objectively measurable, as mentioned in Section 2.2.

The tested performance (difference between suitability of E2D priority sites vs. Darebin average) is adversely affected in some cases by E2D’s access to ‘inside information’ for the selection of E2D priority sites, which is not reflected in input data. A notable example is the location of a priority site at a golf course that would have been considered unsuitable, but was soon to be decommissioned. Such discrepancies highlight the role of tools and models as supporting (rather than replacing) planning, in conjunction with the human decision-making
processes. Nevertheless, SSANTO’s ability to capture some of the tacit knowledge behind prioritisation by E2D using its full criteria set is encouraging.

4.2. Validation challenges

One of the greatest and unsolved scientific challenges within the field of socio-technical modelling is the lack of understanding on how to validate such models (Van Asselt and Rotmans, 2002). Unlike ‘traditional’ urban water models, such as hydrological models, these models do not simulate or predict measurable phenomena, but rather represent and explore human judgement and preferences. Therefore, models such as SSANTO cannot be easily validated against measured field data. Broadly speaking, two contrasting philosophies exist in the academic debate: (1) models cannot be trusted until they are validated, including socio-technical models and (2) focus of socio-technical model testing should be on usefulness and useability rather than validation. To date, validation does not play any role of significance in GIS-MCDA models, as evident from the lack of discussion in highly cited textbooks (e.g. Malczewski and Rinner, 2015) and research articles (e.g. Chang et al., 2008; Rahman et al., 2012). The absence of rigorous and widely accepted solutions to this debate can critically hamper wider acceptance and application of socio-technical models. Thus, modelling communities urgently need to engage more with this question to find credible yet workable solutions. Since making any conclusive statements in this debate falls outside the scope of this article, the highest achievable was to seriously attempt model testing and validation of SSANTO.

Given that suitability cannot be measured, caution is required when concluding that model run P-Full-Entropy is better at modelling suitability. More accurately, it ‘performs better’ in mirroring the outcome of human decision-making processes. In line with earlier statements, just as important as validity is a tool’s usefulness: the ability to reduce the cost while increasing the speed and rigour of decision-making processes and enhance the quality of decision outcomes. Usefulness also depends on a tool’s user experiences, such as user-friendliness, flexibility and relevance to planning problems (Lee Jr, 1973). All of these issues

Fig. 5. Output maps for model run P-Full-Expert. a) suitability map for the ‘Opportunities’ side of the suitability framework, b) suitability map for the ‘Needs’ side of the suitability framework, c) suitability map of ‘Opportunities’ and ‘Needs’ combined, overlaid by E2D priority sites and optional sites (stars).
were found crucial in WSUD planning practice (Kuller et al., 2018b). Therefore, further research should test SSANTO’s performance against these aspects in greater detail (see Section 4.3).

Value scales were developed using sources from diverse and international sources to increase SSANTO’s validity in a diverse set of international contexts. Despite our best efforts, these sources are still limited to information and expertise available to us. Therefore, further research will focus on implementing the option of user-defined value scales (see Section 4.3) to ensure wider applicability.

4.3. How to use SSANTO in practice

SSANTO’s design is intended to foster stakeholder collaboration through its rapid generation of preference-based suitability outcomes. It enables users to compare and discuss individual model runs to gain deeper understanding of the underlying context and preferences of suitability for WSUD. Thus, SSANTO can support more robust decisions. It is, however, not meant to replace human judgement (Reed and Kasprzyk, 2009), and should never be used in isolation. Rather, it should be considered a support tool in a wider decision-making context, which also includes other tools and models as well as human decision-making. For example, SSANTO does not consider catchment hydrology, and therefore high suitability of a location does not warrant appropriateness for WSUD’s place in the treatment train. Such considerations can be accounted for by coupling SSANTO to planning simulators which incorporate flow paths, such as UrbanBEATS (Bach et al., 2013). The place of SSANTO in such iterative decision-making processes is depicted in Fig. 8. Preliminary discussions with planning professionals suggest that SSANTO would provide a valuable addition to their work. SSANTO was developed in response to the urgent need of WSUD planning practitioners for spatial decision analysis (Kuller et al., 2018b).

Fig. 6. Comparison of distribution of suitability between Darebin average (blue) and E2D priority sites (green) for all precinct-scale scenarios. μD: mean suitability for Darebin, μP: mean suitability at priority sites. p: significance level of difference in means resulting from Mann-Whitney U test. Greater positive difference between μD and μP (i.e. green bars appear to the right of blue bars) indicates better performance.

Fig. 7. Comparing SSANTO results for different weighting methods for large systems using full criteria. a) Equal weights, b) expert weights, and c) entropy weights.
4.4. Proposed further development

Future work will focus on qualitatively testing and validating SSANTO through workshops with practitioners and application in decision-making processes. Quantitative validation will include comparisons with multiple suitability analyses previously undertaken in Melbourne and in cities outside Australia, and in-depth sensitivity analysis of weights, value scales and methodologies, to gain deeper understanding of the identified biases and uncertainties (Delgado and Sendra, 2004). Also, future work includes the addition of more advanced functionality for weight elicitation such as SWING (Edwards and von Winterfeldt, 1986), which could improve user experience as well as weight consistency, and has previously been implemented in urban water management research (Scholten et al., 2015). Future work will also ideally add the option for users to define their own value scales, as these can be (to a certain extent) preference dependent and may change with progressive insights in WSUD placement. Finally, enhanced user-friendliness as well as coupling and integration with other models could be achieved by migration of SSANTO to open-source-, standalone- and online platforms.

5. Conclusions

This paper presents a methodology and associated software tool called SSANTO, developed to rapidly assess spatial suitability for the planning and implementation of green and distributed urban stormwater infrastructure, using GIS-MCDA techniques. SSANTO allows suitability mapping for seven different system types including rain gardens and constructed wetlands. The tool allows for diverse criteria sets to be included such as biophysical, environmental, socio-economic and planning related data. For the first time, the two sides of suitability for a broad range of system types are incorporated, applying principles from ecosystem services. The tool’s architecture was designed to facilitate the application of various weighting methods, combination techniques and parameter settings, tailored to the user’s preferences. Analyses of ‘Opportunities’ and ‘Needs’ are presented, using an intuitive and spatially explicit way through colour-coded maps.

Running SSANTO with similar inputs to those used by a WSUD consultancy demonstrated its ability to produce comparable outcomes to manual suitability mapping, confirming the internal validity of the algorithms used. Further testing demonstrated that SSANTO can reflect human decision-making processes by successfully calculating relatively higher suitability values for selected priority locations from a decision-process by a WSUD consultancy.

It was found that using the most comprehensive set of criteria, SSANTO was more successful to reflect the outcomes of a human decision-making process than using only the selection of criteria used for that human decision-making process. This suggests that SSANTO can capture some of the tacit knowledge that planning practitioners use for WSUD placement for our single case study. Furthermore, entropy weighting performed better than expert weighting and equal weighting. However, caution should be used as entropy weighting is associated with certain methodological limitations relating to the impact of different data types on entropy weights. Throughout our study area, ‘Needs’ for green infrastructure is consistently higher than the ‘Opportunities’ for them.

The development of SSANTO aimed for a simple and user-friendly interface and workflow. The aim is to enable experts in sustainable urban water management and planning, as well as lay people, to undertake thorough spatial analysis without the need to invest high amounts of time and resources, associated with manual processes. SSANTO’s rapid suitability analysis aimed at facilitating the assessment of multiple scenarios, increasing our understanding of the interaction between urban planning decisions and our urban context. By comparing the outcomes of iterative application by multiple stakeholders, SSANTO should promote discussion, collaborative planning and a deeper understanding of the variation of stakeholder preferences and their impacts on decision-making. In doing so, SSANTO has the potential to bridge the

![Fig. 8. WSUD planning and decision-making cycle with SSANTO.](image-url)
gap between perceived need for planning support and low utilisation levels of models and tools and improve the outcomes of planning for sustainable urban water management.

**Declaration of Competing Interest**

None.

**Acknowledgements**

The authors express their gratitude towards the City of Darebin for providing us with crucial data and Melbourne Water for funding the Darebin prioritisation study that we used for tool validation and testing. This work was supported by the Australia-Indonesia Centre under the project code RCR-BrownMON: Urban Water Cluster and fund code SRP16 52057764. Funding sources did not have any involvement in the research. Author contribution: The design and scope of the research was done by Martijn Kuller, Peter M. Bach and Ana Deletic, while the research and development of the presented tool was performed by Martijn Kuller with the support of Peter M. Bach. Testing of the tool was conducted by Martijn Kuller with the support and data input from Simon Roberts and Dale Browne. The article was written by Martijn Kuller and reviewed by all authors. All authors have approved the final article.

**Appendices. Supplementary data**

Supplementary data to this article can be found online at [https://doi.org/10.1016/j.scitotenv.2019.06.051](https://doi.org/10.1016/j.scitotenv.2019.06.051).

**References**

ABS. 2016. 2016 Census. Australian Bureau of Statistics, Canberra.

Alexander, E.R. 1989. Sensitivity analysis in complex decision models. J. Am. Plan. Assoc. 55 (3), 323–333.

Ashley, R., Ceronisius, B., Digman, C., Horton, B., Baczkin, T., Smith, B.,...Baylis, A., 2017. Demonstrating and monetizing the multiple benefits from using SuDS. J. Sustain. Water Environ. Build. 4 (2), 05017008.

Bach, P.M., McCarthy, D.T., Ulrich, C., Sitzenfrei, R., Kleidorfer, M., Rauch, W., Deletic, A., 2013. A planning algorithm for quantifying decentralised water management opportunities in urban environments. Water Sci. Technol. 68 (8), 1857–1865.

Bach, P., McCarthy, D., Deletic, A., 2015. Exploring green urban water infrastructure as part of the urban form: a critical tool of best planning practice. Environ. Model Softw. 64, 265–282 Retrieved from [http://www.sciencedirect.com/science/article/pii/S1364815216301623](http://www.sciencedirect.com/science/article/pii/S1364815216301623).

Ben-Dor, É., Kuller, M., Bach, P.M., Ramirez-Lovering, D., Deletic, A., 2017. Framing water sensitive urban design as part of the urban form: a critical tool of best planning practice. Environ. Model Softw. 96, 168–187 Retrieved from [http://www.sciencedirect.com/science/article/pii/S136481521630147X](http://www.sciencedirect.com/science/article/pii/S136481521630147X).

Lee, J.B., Farrelly, D., Deletic, A., Bach, P.M., 2018B. Building effective planning support system for green urban water infrastructure—practitioners’ perceptions. Environ. Sci. Pol. 99, 153–162 Retrieved from [http://www.sciencedirect.com/science/article/pii/S0925857418305486](http://www.sciencedirect.com/science/article/pii/S0925857418305486).

Linnert, J., Duygan, M., Zhang, J., 2016. Preference stability over time with multiple elicitation methods to support wastewater infrastructure decision-making. Eur. J. Oper. Res. 257, 746–760 Retrieved from [http://www.sciencedirect.com/science/article/pii/S0377221716301382](http://www.sciencedirect.com/science/article/pii/S0377221716301382).

Kuller, M., Farevell, D., Deletic, A., Bach, P.M., 2018B. Building effective planning support systems for green urban water infrastructure—practitioners’ perceptions. Environ. Sci. Pol. 99, 153–162 Retrieved from [http://www.sciencedirect.com/science/article/pii/S0925857418305486](http://www.sciencedirect.com/science/article/pii/S0925857418305486).
Rhea, L., Shuster, W., Shaffer, J., Losco, R., 2014. Data proxies for assessment of urban soil suitability to support green infrastructure. J. Soil Water Conserv. 69 (3), 254–265.

Roberts, S., Browne, D., Lloyd, S., 2017. Priority Stormwater Projects for a Water Sensitive Darebin (Retrieved from Unpublished).

Scholten, L., Schuwirth, N., Reichert, P., Lienert, J., 2015. Tackling uncertainty in multi-criteria decision analysis – an application to water supply infrastructure planning. Eur. J. Oper. Res. 242 (1), 243–260 Retrieved from: http://www.sciencedirect.com/science/article/pii/S0377221714007838 https://doi.org/10.1016/j.ejor.2014.09.044.

Shannon, C.E., Weaver, W., 1963. The Mathematical Theory of Communication. University of Illinois Press, Urbana, IL, USA.

Stewart, T.J., 1996. Robustness of Additive Value Function Methods in MCDM. J. Multi-Criteria Decis. Anal. 5 (4), 301–309.

Thévenot, D.R., 2008. Daywater: An Adaptive Decision Support System for Urban Stormwater Management. IWA publishing.

Tomlin, C.D., 1990. Geographic Information Systems and Cartographic Modeling. Prentice Hall.

Van Asselt, M.B.A., Rotmans, J., 2002. Uncertainty in integrated assessment modelling. Clim. Chang. 54 (1-2), 75–105.

Viavattene, C., Scholes, L., Revitt, D.M., Ellis, J.B., 2008. A GIS based decision support system for the implementation of stormwater best management practices. Paper Presented at the 11th International Conference on Urban Drainage, Edinburgh, Scotland, UK.

Vonk, G.A., 2006. Improving Planning Support: The Use of Planning Support Systems for Spatial Planning. KNAG/Netherlands Geographical Studies.

Walker, W.E., Harremoës, P., Rotmans, J., van der Sluijs, J.P., van Asselt, M.B.A., Janssen, P., Krayer von Krauss, M.P., 2003. Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. Integr. Assess. 4 (1), 5–17.