DEVELOPING AN OPEN SOURCE PLUGIN FOR SPATIAL MULTI-CRITERIA DECISION ANALYSIS: ORDERED WEIGHTED AVERAGING IN QGIS

By
Gregory Huang
H.BSc. Bioinformatics and Computational Biology, University of Toronto, 2019

A Major Research Paper
presented to Ryerson University
in partial fulfilment of the
requirements for the degree of
Master of Spatial Analysis (MSA)
in the Program of
Spatial Analysis

Toronto, Ontario, Canada, 2020
©Gregory Huang, 2020
AUTHOR’S DECLARATION

I hereby declare that I am the sole author of this MRP.

This is a true copy of the MRP, including any required final revisions.

I authorize Ryerson University to lend this MRP to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this MRP by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my MRP may be made electronically available to the public.
Developing an Open Source Plugin for Spatial Multi-Criteria Decision Analysis: Ordered Weighted Averaging in QGIS.

By Gregory Huang, Master of Spatial Analysis, Ryerson University, 2020

ABSTRACT

Multi-Criteria Decision Analysis (MCDA) is a technique commonly used in GIS for decision support. In the QGIS environment, few pre-existing plugins exist to help calculate vector-based MCDA techniques such as Ordered Weighted Averaging (OWA). This research paper creates a set of two new Python 3-based plugins: The “Field Standardizer” plugin for quick standardization of vector data, and the “WLC/OWA Tool” plugin for an intuitive calculator to carry out both Weighted Linear Combination (WLC) and OWA to expand MCDA functionalities in QGIS. The plugins are then illustrated in an investigation of socio-economic status of neighbourhoods in the City of Toronto. The investigation also analyzes the mapped results of different OWA weight options tested, such as MAX, MIN, WLC, and other customized weights.

Keywords: Multi-Criteria Decision Analysis (MCDA), Ordered Weighted Averaging (OWA), Standardization, QGIS, Python Plugin
ACKNOWLEDGEMENTS

I would like to thank Dr. Claus Rinner for his continued guidance throughout my entire time in the MSA program at Ryerson, as well as his willingness to discuss research topics at length even before my admittance to the program. I would also like to thank Dr. Evan Cleave and Dr. Andrew Millward for being in the examination committee for this MRP, who offered valuable insights and helped me improve this paper. Finally, I would like to thank my family for supporting me and helping me along the way get to where I am today.
# TABLE OF CONTENTS

AUTHOR’S DECLARATION .................................................................................................................. ii

ABSTRACT ......................................................................................................................................... iii

ACKNOWLEDGEMENTS .................................................................................................................... iv

LIST OF TABLES .................................................................................................................................. vi

LIST OF FIGURES ............................................................................................................................. vii

1 INTRODUCTION .............................................................................................................................. 1

1.1 Background ................................................................................................................................... 1

1.2 Research Questions and Objectives ............................................................................................ 2

1.3 Structure of the Paper .................................................................................................................. 2

2 LITERATURE REVIEW ..................................................................................................................... 3

2.1 MCDA in GIS ............................................................................................................................... 3

2.2 Weighted Linear Combination (WLC) Technique ....................................................................... 4

2.3 Ordered Weighted Average (OWA) Technique ........................................................................... 6

3 METHODOLOGY ............................................................................................................................ 11

3.1 Plugin Development ..................................................................................................................... 11

3.1.1 “Field Standardizer” Plugin ................................................................................................ 11

3.1.2 “WLC/OWA” Plugin ............................................................................................................. 14

3.2 Case Study Design ....................................................................................................................... 17

3.2.1 Data Overview and Criteria Selection ................................................................................ 17

3.2.2 Data Standardization Scheme ............................................................................................. 18

3.2.3 Criteria Weighting Scheme .................................................................................................. 22

3.2.4 Order Weighting Scheme ...................................................................................................... 23

4 CODE IMPLEMENTATION RESULTS ......................................................................................... 24

4.1 “Field Standardizer” Plugin ..................................................................................................... 24

4.2 “WLC/OWA Tool” Plugin .......................................................................................................... 27

5 CASE STUDY RESULTS AND ANALYSIS .................................................................................... 30

5.1 Criteria Standardization ............................................................................................................. 30

5.2 Comparison Between Different OWA Weighting Schemes .................................................... 34

6 DISCUSSION AND CONCLUSION ............................................................................................. 43

APPENDIX .......................................................................................................................................... 45

REFERENCES ....................................................................................................................................... 46
LIST OF TABLES
Table 1: Pairwise comparison matrix template for WLC weight determination ........................................ 4
Table 2: Score/Scale for comparisons used in Pairwise Comparison Matrix ........................................ 5
Table 3: Summary of standardization methods for each ID assigned.................................................. 13
Table 4: Example calculation of an OWA score at location i................................................................. 15
Table 5: Criteria selected for the study and descriptions .................................................................... 17
Table 6: Pairwise comparison for the five chosen criteria ................................................................. 22
Table 7: Criterion weights calculation for the five chosen criteria ..................................................... 22
Table 8: Sets of order weights from $\alpha = 0$ (pessimistic) to $\alpha = 1$ (optimistic) for five criteria. Source: Wu et al. (2009) ........................................................................................................ 23
**LIST OF FIGURES**

Figure 1: Example calculation of a WLC MCDA technique in GIS in a data set of four locations each having three criteria considered for analysis. ................................................................. 6

Figure 2: Decision Strategy Space diagram with two criteria. ................................................................. 9

Figure 3: Histogram results for the five criteria selected (bin = 50). ............................................................ 19

Figure 4: Normal Q-Q plots for the five criteria selected. .............................................................................. 20

Figure 5: Boxplots for the five criteria selected. ......................................................................................... 21

Figure 6: Initial startup page when “Field Standardizer” plugin is launched. ................................................. 24

Figure 7: Example of a populated “Field Standardizer” plugin ready for execution. ...................................... 26

Figure 8: Initial startup page of the “WLC/OWA Tool” plugin when launched. ........................................... 27

Figure 9: Example of a populated “WLC/OWA Tool” plugin ready to be executed. ......................................... 29

Figure 10a: Standardized scores for average dwelling value in Toronto Census Tracts (CTs) .................. 30

Figure 10b: Standardized scores for average household income criterion in Toronto. ............................. 31

Figure 10c: Standardized scores for percentage of bachelor’s degree criterion in Toronto. ....................... 31

Figure 11a: Standardized scores for unemployment rate criterion in Toronto. ........................................ 32

Figure 11b: Standardized scores for prevalence of poverty criterion in Toronto. ...................................... 32

Figure 12a: Inputs for MIN OWA ($\alpha = 0$). ............................................................................................... 34

Figure 12b: Inputs for OWA with $\alpha = 0.3$ ............................................................................................... 34

Figure 12c: Inputs for WLC ($\alpha = 0.5$) ..................................................................................................... 35

Figure 12d: Inputs for OWA with $\alpha = 0.7$ ............................................................................................... 35

Figure 13: Mapped MIN OWA Scores ($\alpha = 0$). ......................................................................................... 36

Figure 14: Mapped $\alpha = 0.3$ OWA Scores .................................................................................................. 36

Figure 15: Mapped WLC Scores ($\alpha = 0.5$) ............................................................................................. 37

Figure 16: Mapped $\alpha = 0.7$ OWA Scores ................................................................................................. 37

Figure 17: Mapped MAX OWA Scores ($\alpha = 1$) ...................................................................................... 38

Figure 18: Main areas in Toronto discussed in analysis. .............................................................................. 38
1 INTRODUCTION

1.1 Background

Multi-Criteria Decision Analysis (MCDA) is commonly used in Geographic Information Systems (GIS)-based decision support, as it allows the distillation of multiple variables (criteria) into a new singular index for geographical features (Malczewski, 2006). There are multiple methods that can be utilized for MCDA in the realm of GIS, but the more commonly used techniques include Weighted Linear Combination, or WLC, and Ordered Weighted Averaging, or OWA (Malczewski, 1999). The use of MCDA can help with decision support on a multitude of spatial problems, from physical geography topics such as site selection, to social or retail geography such as demographic analyses.

While MCDA is a widely known technique in GIS, there are not a lot of vector-based MCDA tools available in GIS software, especially in the open-source Quantum GIS (QGIS) environment (Malczewski & Rinner, 2015). The issue was further exacerbated with the QGIS platform moving from Python 2 to Python 3, as many plugins that were not updated became unusable. There are no current tools in QGIS that perform the OWA technique, and an existing WLC QGIS tool previously developed by Renacin Matadeen (2019) has room for improvement in areas such as user interface design and insufficient data standardization functionalities. In addition to the lack of existing MCDA tools in QGIS, there is also a lack of standardizing tools, which are essential to any type of numeric aggregation method such as MCDA, as the standardization of data allows the different criteria to be compared numerically. As such, spatial analysts who wish to conduct vector-based MCDA may have to individually standardize and calculate WLC or OWA scores, which not only takes up valuable time but also creates opportunities for errors in calculation (Malczewski & Rinner, 2015).

The goal of this paper is to help address the lack of MCDA tools in the QGIS environment by creating both a tool to standardize data and a tool to calculate the scores of two common MCDA techniques, WLC and OWA, and demonstrate the tools through a case study. With the creation of these tools, carrying out MCDA in QGIS can be greatly simplified for spatial analysts and hopefully reduce sources of error in these analyses.
1.2 Research Questions and Objectives

This research paper documents the creation of two (2) Python-based QGIS plugins to achieve the main objective of creating a user-friendly environment in QGIS for MCDA:

a. A “Field Standardizer” plugin that allows users to easily standardize the data that they wanted to use by offering multiple available standardization methods, and

b. A “WLC/OWA Tool” plugin that offers an intuitive platform for WLC and OWA analyses, where desired criteria can be easily selected, criterion weights and order weights can be easily defined, and pre-set values for quick exploratory analyses are made available.

In addition to the software development, the plugins will be tested for their functionality through a case study regarding socio-economic status of Toronto Census Tracts (CTs), using two research questions:

a. What are some observed differences between the WLC technique and the OWA technique?

b. What are the observed outcomes of using different order weight schemes with OWA?

These two research questions should highlight the conceptual differences between the two MCDA techniques and how these differences translate in a spatial context. With the dissemination of the Python plugins as open-source software and the demonstrations through the case study that answer the research questions, this research project offers QGIS users an accessible and sustainable platform for data standardization and performing WLC/OWA MCDA.

1.3 Structure of the Paper

The paper first provides a literature review of the two MCDA methods that will be implemented in the plugin: The WLC and OWA techniques. Subsequently, the paper explains the methodologies for the creation and design of the two plugins as well as the case study. Then, the paper examines the implementation results of the plugin and analyzes the results from applying the plugins on the case study data. Finally, limitations of the study and a conclusion of the research will be given at the end of the paper.
2 LITERATURE REVIEW

2.1 MCDA in GIS

There are often problems in GIS where the solutions require assessing multiple attributes of geographical locations to come up with some sort of decision or evaluation; while multi-criteria decision analysis (MCDA) techniques were not created initially for use on spatially referenced data in GIS analyses, there has been a steady growth of GIS use cases of MCDA for these kinds of geographical decision support (Carver, 1991; Malczewski, 2006). The applications of various MCDA methods are multi-faceted and encompass a wide range of geographical problems, both by virtue of the ubiquity of data integration problems in the GIS realm and by virtue of the extraordinary range of available customizations that can be achieved with MCDA processes (Malczewski & Liu, 2014). Some examples of different criteria combination techniques include Weighted Linear Combination (WLC) and Ordered Weighted Averaging (OWA). Within each of the principal MCDA methods, there are also varying degrees of customization that can tailor the combination results to each individual analysis (Malczewski, 2006). Even with the relatively simplistic WLC technique, where criteria values were simply assigned weights and combined for the MCDA score, there are a plethora of available weighting schemes for specific needs or goals of the user; in the case of OWA, the order weights add an additional layer of complexity and customizability compared to WLC (Eastman, 1993; Malczewski & Liu, 2014). The applications of these MCDA methods in GIS span fields from physical geography such as site suitability analysis, to social geography such as socio-economic statuses of local neighbourhoods (Hajizadeh et al., 2020; Malczewski, 2006).

With the available MCDA methods, to carry out these analyses in a GIS environment requires a multitude of tools and varying levels of automation. In most GIS software environments, the MCDA calculations can be done in different ways: in the desktop environment, calculations can be done incrementally in table calculations, through pre-existing tools that automate the process, or by using a combination of both, such as user model-building for MCDA calculation modules. In the web-based environment, most of the options are constrained to pre-existing scripts that can be executed (Malczewski & Rinner, 2015). For more simplistic approaches such as WLC that utilize a relatively small set of criteria, perhaps using the incremental Field Calculator approach may be sufficient; however, when more complex MCDA
analyses needed to be carried out, existing solutions may be limiting for the user who desires to perform MCDA on a dataset (Malczewski & Rinner, 2015).

2.2 Weighted Linear Combination (WLC) Technique

As mentioned above, one of the more straightforward MCDA methods used for solving GIS-based MCDA problems is the Weighted Linear Combination (WLC) technique. When using WLC, each of the selected criteria is given a specific “weight”, where the sum of all weights given out in the analysis totals to 1.0; then, at each of the geographic locations, the normalized and standardized values at the location for each of the selected criteria that the analyst wishes to investigate is multiplied by its respective criterion weight (Malczewski, 2006). Finally, the values with the weights applied are summed for each location, becoming the aggregated WLC score. The weights are often subjective to the user conducting the research, and there are multiple published methods to generate those weights (Carver, 1991; Malczewski, 2000; Eldrandaly, 2013). One of the many ways to obtain criterion weights is through the pairwise comparison matrix method, since the method is relatively straightforward and intuitive (Carver, 1991; Saaty, 1980). If the pairwise comparison method is chosen, in order to determine the weights for the criteria, a pairwise comparison matrix that examines the criteria used will be generated based on the template shown in Table 1 (Carver, 1991). The scores given for each comparison should follow the definitions given in Table 2 (Saaty, 1980), or their reciprocals.

Table 1: Pairwise comparison matrix template for WLC weight determination

<table>
<thead>
<tr>
<th></th>
<th>Criterion 1</th>
<th>Criterion 2</th>
<th>…</th>
<th>Criterion n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion 1</td>
<td>1</td>
<td>$s_{12}$</td>
<td>…</td>
<td>$s_{1n}$</td>
</tr>
<tr>
<td>Criterion 2</td>
<td>$\frac{1}{s_{12}}$</td>
<td>1</td>
<td>…</td>
<td>$s_{2n}$</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>Criterion n</td>
<td>$\frac{1}{s_{1n}}$</td>
<td>$\frac{1}{s_{2n}}$</td>
<td>…</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2: Score/Scale for comparisons used in Pairwise Comparison Matrix

<table>
<thead>
<tr>
<th>Strength of Relative Importance</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Equal importance</td>
</tr>
<tr>
<td>2</td>
<td>Equal to moderate importance</td>
</tr>
<tr>
<td>3</td>
<td>Moderate importance</td>
</tr>
<tr>
<td>4</td>
<td>Moderate to strong importance</td>
</tr>
<tr>
<td>5</td>
<td>Strong importance</td>
</tr>
<tr>
<td>6</td>
<td>Strong to very strong importance</td>
</tr>
<tr>
<td>7</td>
<td>Very strong importance</td>
</tr>
<tr>
<td>8</td>
<td>Very to extremely strong importance</td>
</tr>
<tr>
<td>9</td>
<td>Extreme importance</td>
</tr>
</tbody>
</table>

Through the combination of Tables 1 and 2, a weight for each criterion can be calculated. To populate a matrix, a number indicating the strength of relative importance chosen by the analyst will be given to each pairwise comparison of criteria. For example, if Criterion 1 has strong importance for the decision relative to Criterion 2, the matrix at index $s_{12}$ will be given a score of 5. Conversely, since Criterion 1 is relatively more important than Criterion 2 in a pairwise comparison, the inverse relationship (i.e., $s_{21}$, where Criterion 2 is compared to Criterion 1) will therefore also be the reciprocal of 5, receiving a score of 0.2. Note that the indexes in Table 1 such as $s_{12}$ only indicate the location and not the values assigned.

Subsequently, when a criterion is compared against itself in a pairwise comparison matrix, the score should always be 1. After pairwise comparisons for each criterion were completed, the individual criterion weights for that particular analysis can then be inferred through the pairwise comparison matrix.

Additionally, since multiple criteria of potentially vastly different ranges and units may be used to calculate the WLC score, the data used for the analysis need to be standardized to avoid certain criteria from dominating the rest, and the choice of standardization technique can also vary based on both the statistical characteristics of the raw data and the user’s preferences.
An example of a WLC operation workflow on four locations using three criteria can be seen in Figure 1.

![Example calculation of a WLC MCDA technique in GIS in a data set of four locations each having three criteria considered for analysis](image)

The relative simplicity contributes to the widespread use of WLC-based MCDA in GIS, where it is often applied in the context of suitability analysis for land and site selections (Malczewski, 2000). For example, in a suitability analysis by Blachowski (2015), WLC, in combination with the pairwise comparison matrix for criteria weighting, combined a wide range of 15 criteria such as railroads, ecological corridors, urban areas, and gas networks, to assess the accessibility of a coal deposit in Poland.

2.3 Ordered Weighted Average (OWA) Technique

The Ordered Weighted Average (OWA) Technique in GIS can be viewed as a “step up” from the WLC technique, where in addition to the criterion weights given and applied for each of the criteria selected, an additional “order weight” is applied to obtain the final aggregated score at each location. Initially proposed by Yager (1988) for generalized multiple-criteria analysis, Eastman (1997) was first to apply the principles of OWA in a spatial context. Yager (1988) proposed that in multiple-criteria decision making, the aggregation methods will sometimes have
to consider the satisfaction of criteria. In OWA, the measures of criteria satisfaction is generally based on two extreme operators: the “AND” (or “MIN”) operator, and the “OR” (or “MAX”) operator; where for the “AND” operator, the aggregation demands that all criteria should be satisfied, whereas for the “OR” operator the aggregation only seeks the satisfaction of any one criterion used in the analysis. In between the two extremes of those two operators, other order weights can be calculated based on the “ORness” of the analysis, giving rise to a new set of complexities to multiple-criteria analysis (Yager, 1988; Eastman, 1997). In addition to the measures of “ORness”, measures of trade-off and dispersion can also be used as the driver of order weight assignments (Rinner & Malczewski, 2002). With the ability to add order weights based on metrics such as degrees of ORness, the OWA technique is most ideal for situations where there are degrees of uncertainties, which in turn contribute to “risk” in the decision-making process (Ahn & Yager, 2013; Rinner & Voss, 2013). Given OWA’s advantages in risk management for uncertainties, many applications of the OWA technique in GIS were focused on cases related to risk assessment and vulnerabilities. For example, utilizing OWA for flood risks, landslide susceptibility, and using the maximum entropy approach to calculate OWA scores for the contiguous United States’ flood susceptibility in the context of global warming (Tang et al., 2018; Feizizadeh & Blaschke, 2013; Runfola et al., 2017).

There is an abundance of ways in which order weights can be assigned, with multiple metrics that can be used for calculation. Using the example of measurements of ORness for order weight calculations, O’Hagan (1988) and Yager (1988) suggested the maximum entropy approach by utilizing the following equations as a means to arrive at order weights:

\[
\alpha = \sum_{k=1}^{n} \frac{n-k}{n-1} \lambda_k
\]  

(1)

Where \(\alpha\) measures the degree of ORness in the analysis, \(k\) is the number of criteria, and \(\lambda_k\) is the given order weight at the \(k^{th}\) criteria. Also note that the sum of order weights should be 1, same as the criterion weights, shown in Equation (2):

\[
\sum_{k=1}^{n} \lambda_k = 1
\]  

(2)

In Equation (1), \(\alpha\) measures ORness by its similarity to the OR/MAX operator. If \(\alpha\) is 1, this means that the decision-makers are looking for maximum ORness, or the most generous case where any criteria can be satisfied; to solve for \(\alpha = 1\), it means that \(\lambda_k = 1\) at the first order
weight, and the remaining order weights are zero. This can also be interpreted as an “optimistic”
or maximum (MAX) order weight scheme. If \( \alpha \) is 0, this means the opposite case, where the
order weights demand satisfaction of all criteria, and \( \lambda_k = 1 \) at the last order weight, while all
preceding order weights are zero. Conversely, this can be interpreted as a “pessimistic” or
minimum (MIN) approach for order weighting. With the two extremes defined, the middle
ground where \( \alpha = 0.5 \) is in fact the WLC approach, where \( \lambda_k = \frac{1}{n} \) for all of the order weights.
The varying levels of \( \alpha \) other than 0, 0.5, and 1 allow for a wide range of order weights that can
be used given the different levels of ORness determined by the analyst as well as the number
of criteria selected; however, given the unique characteristics of \( \alpha \) at 0, 0.5, and 1 for AND, WLC,
and OR operators, respectively, where the extreme cases of ORness exists, the number of criteria
used in the OWA analysis will not have an effect on the set of order weights given (Malczewski
& Rinner, 2015).

Equation (1) can be used in conjunction with dispersion measures to calculate OWA weights
using the maximum entropy approach (O’Hagan, 1988; Malczewski & Liu, 2014). The
dispersion measure is calculated with Equation (3) below:

\[
\varphi = - \sum_{k=1}^{n} \frac{\lambda_k \ln \lambda_k}{\ln n}
\]  

(3)

Where \( \varphi \) is the measure of dispersion from range 0 to 1. In the dispersion equation, when \( \varphi 
= 0 \), the order weight \( \lambda_k = 1 \) at the last order weight, similar to the pessimistic approach, and at
\( \varphi = 1 \), which is maximum dispersion, the order weight \( \lambda_k = \frac{1}{n} \) for all criteria (Malczewski &
Liu, 2014). Therefore, for the maximum entropy approach, the set of order weights can be
obtained by solving Equation (4) below through non-linear programming (Malczewski & Liu,
2014).

\[
\text{maximize } \phi = - \sum_{k=1}^{n} \frac{\lambda_k \ln \lambda_k}{\ln n}
\]  

(4)

Note that Equation (4) should also adhere to the boundaries set by Equations (1) and (2),
which act as constraints to the maximization of the algorithm (Fullér & Majlender, 2003).

In addition to the ORness of OWA operators where \( \alpha \) examines the similarity to the logical
OR operator, the “trade-off” measure can also be looked at for substitutability. The trade-off
measure indicates how well one criterion can compensate for another criterion with lower values; in other words, the higher the trade-off, the easier it is for one criterion to substitute for another (Jiang & Eastman, 2000; Malczewski & Liu, 2014). The trade-off measure, in conjunction with the ORness measure, can be utilized for the evaluation of the chosen decision strategy (Jiang & Eastman, 2000). In the decision strategy space, an OWA analysis utilizing just two criteria will form a triangular shape, as shown in Figure 2, but will approach a more rectangular shape (with vertices at AND/OR/WLC maintained) as new criteria are added (Rinner & Malczewski, 2002). The evaluation strategy graph indicates the relationship between tradeoff and ORness measures, which can help assist an analyst in selecting an \( \alpha \) number to determine order weights based on the needs of each particular study.

![Figure 2: Decision Strategy Space diagram with two criteria](image)

While the maximum entropy approach to generate the set of order weights is more commonly used, there are also a number of other, more complex algorithms accounting for different user needs (Wu et al., 2009). Some of these methods include some degree of transformation from the maximum entropy approach done by O’Hagan (1988). For example, a polynomial transformation
of the non-linear maximum entropy function done by Fullér & Majlender (2001); Majlender’s (2005) development of maximal Rényi entropy method, which specifies that the maximum entropy approach can be one of its use cases; and Liu and Chen’s (2004) development of a variation of the maximum entropy approach with parametric maximum entropy OWA (PMEOWA). Other examples include novel variations such as the minimal variability method, which focuses on the set of weights (weighting vector) itself in cases where the analyst wishes to obtain minimal variations between the different order weights, as well as using Monte Carlo simulation to obtain order weighting vectors (Fullér & Majlender, 2003; Tang et al., 2018). With the wide range of available algorithms to generate sets of order weights, the OWA technique is relatively robust in terms of tackling a user’s specific decision strategy, and that variability can be useful for multiple GIS applications as well.
3 METHODOLOGY

This section explains the design rationale for both of the implemented plugins, as well as the method choices for the case study. The design for the plugins has the goal of maximizing their possible areas of application, hence the decision to separate the standardizing tool and the WLC/OWA tool. On the other hand, the design for the case study was aimed to clearly demonstrate the effects of assigning different weights for the MCDA calculations, in addition to its goal of demonstrating the utility of the plugins. Note that for this study, the data used and the assumed use cases for the created plugins are based on vector data, not raster data.

3.1 Plugin Development

The plugins were designed to be used in the QGIS 3.0+ environment, where the plugins were coded in Python 3, with QGIS-specific functionalities. The plugins developed in this project were created with QGIS version 3.14 “π”, with frontend design completed in the Qt Designer application that contains QGIS-customized user interface widgets. In order to package the code that carries out the analyses into a plugin, the existing “Plugin Builder” plugin was used to generate the templates for the implementation of the two plugins in this project. The Plugin Builder tool was essential for the initialization of the graphic user interface (GUI) for the plugins, which was the key component of the creation of the plugins. Once the templates were set up using Plugin Builder, the development of the plugins then split into two components: The GUI design component (frontend) and the algorithm implementation component (backend). The frontend design was completed first by utilizing the Qt Designer application. Then, the backend design was implemented, often referencing the objects created in the frontend. The details of the designs for each plugin will be elaborated in the following sub-sections.

3.1.1 “Field Standardizer” Plugin

The “Field Standardizer” plugin was created as a response to answer the lack of a QGIS function or plugin that could quickly and intuitively standardize columns (fields) of a given vector layer, and to support the required standardizing step for WLC/OWA calculations. Note that for the context of this research paper, “standardization” means to re-scale all criteria into a common range. As mentioned earlier, the WLC/OWA calculations are composite scores of
multiple criteria that can have vastly different ranges, and to prevent one criterion from being the
dominant factor in determining WLC or OWA scores, the criteria needed to be re-scaled to a
common range, which gave rise to the necessity of a standardization tool.

The decision to separate the field standardizer plugin was influenced by two major
factors: First, the consideration was that the inclusion of a standardization widget within the
WLC/OWA plugin itself might inundate a prospective user of the tool with too much
information displayed on the GUI, as the standardization protocols and explanations take up
space in the interactive window that is already crowded with various components. Second,
having a standalone field standardizer plugin makes the plugin available to all kinds of other
applications that may require standardization of field data.

With that in mind, the decision was made to implement the standalone plugin, using three
(3) standardization techniques: maximum score, score range, and z-score. Each of these
standardization techniques have two variants implemented as well – a cost and a benefit function.
Since the WLC and OWA MCDA analyses determine a location’s suitability by a final
aggregated score, where the higher the score the better, when a criterion (field) is standardized
from a raw value, if a higher raw value is less favourable for the analysis, a cost standardization
has to be applied, whereas if a higher raw value is more favourable for the analysis, a benefit
standardization will be applied. As a result, a total of six different standardization formulae were
made available for the user of this plugin.

The frontend of the plugin was designed to be as simple as possible, where the user can
intuitively pick the layer that they would like to work on, easily select the fields that they wanted
to standardize, and assign a standardization method to each of the fields selected. It was decided
that each of the six available standardization options were assigned an ID, with the assignments
displayed in the GUI, in order to limit the chances of user error and streamline the process.

In addition to implementing the standardizing formula, the backend of the plugin also
monitors the user actions to check for data compliance. For example, the code will only allow
users to select numeric fields for standardization, and check that the user has entered a valid
standardization ID. If a wrong ID was entered, the process will terminate and throw an error
message, alerting the user that the input that they have entered was not valid. If the correct IDs
were entered, the code will perform the corresponding standardization calculations. The calculation overwrites the original raw values to standardized/rescaled values.

The IDs for each standardization method as well as their formulae are as follows, shown in Table 3.

<table>
<thead>
<tr>
<th>ID</th>
<th>Standardization Method</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Max Score – Benefit</td>
<td>( V'<em>{ij} = \frac{V</em>{ij}}{\text{Max}(V_j)} )</td>
</tr>
<tr>
<td>2</td>
<td>Max Score – Cost</td>
<td>( V'<em>{ij} = 1 - \frac{V</em>{ij}}{\text{Max}(V_j)} )</td>
</tr>
<tr>
<td>3</td>
<td>Score Range – Benefit</td>
<td>( V'<em>{ij} = \frac{V</em>{ij} - \text{Min}(V_j)}{\text{Max}(V_j) - \text{Min}(V_j)} )</td>
</tr>
<tr>
<td>4</td>
<td>Score Range – Cost</td>
<td>( V'<em>{ij} = \frac{\text{Max}(V_j) - V</em>{ij}}{\text{Max}(V_j) - \text{Min}(V_j)} )</td>
</tr>
<tr>
<td>5</td>
<td>Z-score – Benefit</td>
<td>( V'<em>{ij} = \frac{V</em>{ij} - \mu}{\sigma} )</td>
</tr>
<tr>
<td>6</td>
<td>Z-score – Cost</td>
<td>( V'<em>{ij} = -\frac{V</em>{ij} - \mu}{\sigma} )</td>
</tr>
</tbody>
</table>

Where \( V'_{ij} \) is the standardized value of field \( j \) at location \( i \), \( V_{ij} \) is the raw value of field \( j \) at location \( i \), \( \text{Max}(V_j) \) is the maximum raw value in field \( j \) across all locations, \( \text{Min}(V_j) \) is the minimum raw value in field \( j \), \( \mu \) is the raw mean of field \( j \), and \( \sigma \) is the standard deviation of field \( j \). These formulae are implemented accordingly in the code.

The multiple standardization methods are selected for different potential use cases a user of the plugin may encounter. IDs 1 to 4 turn the raw data into values between 0 and 1. IDs 5 and 6 transform the data to have a mean of 0 and a standard deviation of 1 (z-score). Table 3 shows that the “cost” functions reverse the “benefit” functions’ rescaled/standardized scores, where larger raw values in the “cost” functions will be standardized into comparatively smaller values, and smaller raw values will be standardized into comparatively larger values. Implementing the
“cost” functions is crucial for the purpose of calculating WLC or OWA MCDA scores, as the scores are interpreted in a way that a larger value represents better suitability. With the wide range of available options, it is up to the user’s discretion to determine which of the techniques to use, based on the intended use cases for the standardized fields.

3.1.2 “WLC/OWA” Plugin

The “WLC/OWA” plugin utilizes a general design principle similar to that of the “Field Standardizer” plugin, where the goal was to help the user of the plugin intuitively perform either a Weighted Linear Combination (WLC) calculation or an Ordered Weighted Average (OWA) calculation. At this stage, the user should have already standardized the criteria that will be used in the plugin.

The Weighted Linear Combination (WLC) formula is as follows:

$$WLC_i = \sum_{k=1}^{n} v_{ik} w_k$$

(5)

Where $WLC_i$ is the WLC score at location $i$, $v_{ik}$ is the value of location $i$ for criterion $k$ from a set of $n$ criteria, and $w_k$ is the defined weight of criterion $k$. Note that the sum of all criterion weights should equal to 1 ($\sum_{k=1}^{n} w_k = 1$).

The Ordered Weighted Average (OWA) formula is as follows:

$$OWA_i = \sum_{k=1}^{n} z_{ik} o_k$$

(6)

Where $OWA_i$ is the OWA score at location $i$, $z_{ik}$ is the ordered weighted value of location $i$ in criteria $k$ from a set of $n$ criteria, and $o_k$ is the order weight of criteria $k$. To clarify, $z_{ik} = v_{ij} \times w_{ij}$, with the indices $j$ and $k$ rearranged so that $z_{i1} \geq z_{i2} \geq \cdots \geq z_{in}$. Also note that the sum of all order weights should equal to 1 ($\sum_{k=1}^{n} o_k = 1$). These two equations were implemented in the backend of the plugin. Additionally, there are two variations of OWA score calculation in the literature. While Malczewski & Liu (2014) applied the order weights based on the sorted values of the standardized criterion values before the criterion weights were applied, we use the approach by Malczewski & Rinner (2005) who applied criterion weights before order weights. Therefore, the user of this plugin will be made aware of the potential discrepancies in calculation methodology in the README document of the plugin to avoid confusion.
Table 4 shows a hypothetical calculation at location $i$ using the OWA formula used in the plugin. In this example, four criteria are used, and random weights are assigned. The order weight assignments, from maximum weighted criteria value to minimum weighted criteria value, are 0.1, 0.2, 0.3, and 0.4. Note that the same calculation will happen for every location in the specified vector layer.

Table 4: Example calculation of an OWA score at location $i$

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Standardized Value</th>
<th>Criterion Weight</th>
<th>Standardized Value * Criterion Weight</th>
<th>$z_{lk} \cdot o_k$ (sorted weighted criteria values * order weights)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion 1</td>
<td>0.2</td>
<td>0.5</td>
<td>0.10</td>
<td>0.10*0.2 = 0.020</td>
</tr>
<tr>
<td>Criterion 2</td>
<td>0.9</td>
<td>0.2</td>
<td>0.18</td>
<td>0.18*0.1 = 0.018</td>
</tr>
<tr>
<td>Criterion 3</td>
<td>0.3</td>
<td>0.2</td>
<td>0.06</td>
<td>0.06*0.4 = 0.024</td>
</tr>
<tr>
<td>Criterion 4</td>
<td>0.7</td>
<td>0.1</td>
<td>0.07</td>
<td>0.07*0.3 = 0.021</td>
</tr>
</tbody>
</table>

$OWA_i = 0.083$

Upon examining the WLC and OWA formulae (Equations 5 and 6), it can be observed that the OWA formula can also be used to carry out a WLC calculation, if the order weights assigned to each criterion in the OWA formula were the same, as the comparative values of the criteria scores do not impact the order weights applied. While the main purpose of the study was to add the OWA functionality to the QGIS repository of plugins, with this special relationship between OWA and WLC, the added benefit of developing an OWA plugin was that the plugin will be able to perform both MCDA methods. To ensure that the availability of the WLC technique was obvious to the user in an interface geared primarily towards the calculation of an OWA score, a “WLC” button was added to one of the order weight pre-sets so the user could easily convert an OWA calculation to a WLC one.

Given the similarities in the initial steps of user interaction between the “Field Standardizer” plugin and the “WLC/OWA Tool” plugin, the frontend of the WLC/OWA Tool also aimed to create a streamlined experience for the user to select the desired layer and criteria for the analysis. It was also decided that two “pre-set” sections, one for criterion weights and
another for order weights, should be added to the GUI in order for the user to quickly fill out weights and decrease potential chances of user error in specific use cases. Note that in the GUI, for ease of understanding, the user inputs the order weights from maximum to minimum. This means that the largest weighted criteria value gets multiplied by the first order weight entered, the second-largest weighted criteria value gets multiplied by the second order weight, and so on.

In order to calculate the OWA scores, given the necessity to sort the weighted values, the backend implementation for the formula was organized in the following steps:

Step 1: The names of the selected criteria are identified.

Step 2: The code iterates through all locations (features) in the input layer. At each location, the code visits each of the selected criteria. The location’s value at each criterion is multiplied by its designated criterion weight.

Step 3: At the same location, the code sorts the weighted criteria values from maximum to minimum – this is the “ordering” part of OWA – and the sorted list of criterion values is multiplied by the order weights.

Step 4: After the order weights are applied, the calculated values for each criterion at the current location are summed, resulting in the location’s OWA score. This score is added to a newly created field called “OWA Score” at each location. A message tells the user that the calculations have been completed successfully.

In addition to the implementation of the OWA technique, similar to the “Field Standardizer” plugin, additional checks were put in place to make sure the user provided correct input data. This included a check to see if all criterion weights sum to 1, and all order weights also sum to 1. Additionally, only numeric criteria can be selected for analysis. These restrictions were also displayed in the GUI itself to remind the user to check their input. In the event of incorrect criterion or order weights being provided, the plugin will throw an error notifying the user that incorrect weights were given, and the process terminates. If all user inputs were correct and passed the checker, the process runs, and the new OWA scores will be calculated for each location according to the user’s inputs.
3.2 Case Study Design

In order to test the two completed plugins, in addition to conducting tests on random sample data in the plugin development stage to make sure that the calculations for every scenario were correct, a case study with real data was conducted as a proof of concept. The case study helped test the calculations, illustrate the functions of the plugins, and also enabled an analysis of how different criterion and order weights affect outcomes of MCDA in a mapped setting. The case study employs 2016 Census data in a polygon vector layer to examine socio-economic status of Census Tracts in the City of Toronto.

3.2.1 Data Overview and Criteria Selection

The 2016 census data obtained from Statistics Canada contained a multitude of potential indicators for socio-economic status, which included metrics such as a census tract (CT)’s average household income, unemployment rate, and percentage of holders of a bachelor’s degree or higher. For this case study, the scope of the analysis was limited to the boundaries of the City of Toronto, while the data included the entire Toronto Census Metropolitan Area (CMA).

Five criteria were selected for this analysis and are shown in Table 5 with details about the characteristics of the criteria. Note that since the bachelor’s degree criterion was a raw count, it was normalized by each feature’s total population. Normalization of criteria before analysis is critical in order to avoid unwanted bias and skewness.

Table 5: Criteria selected for the study and descriptions

<table>
<thead>
<tr>
<th>Criterion Name</th>
<th>Description</th>
<th>Unit of measurement</th>
<th>Benefit (maximize) or Cost (minimize)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNEMP_RATE</td>
<td>Unemployment Rate</td>
<td>Percentage (%)</td>
<td>Cost (minimize)</td>
</tr>
<tr>
<td>AVG_VAL</td>
<td>Average Value of Dwellings</td>
<td>Dollars ($)</td>
<td>Benefit (maximize)</td>
</tr>
<tr>
<td>AVG_HH_INC</td>
<td>Average Household Income</td>
<td>Dollars ($)</td>
<td>Benefit (maximize)</td>
</tr>
<tr>
<td>POP_PRV</td>
<td>Prevalence of Low Income</td>
<td>Percentage (%)</td>
<td>Cost (minimize)</td>
</tr>
<tr>
<td>BACH_DEG</td>
<td>Population aged 25-64 with bachelor’s degree</td>
<td>Percentage (%)</td>
<td>Benefit (maximize)</td>
</tr>
</tbody>
</table>
These criteria were selected based on their relevance to the goal of analyzing socio-economic status of the city, as well as the opportunity to test the cost/benefit variations in the “Field Standardizer” plugin. Household income and dwelling value are criteria that can potentially provide direct insights on a neighbourhood’s economic status, whereas the other criteria can be indicators of a combination of both social and economic status. Note that from Table 5, a benefit/cost field indicates the standardization scheme used for these criteria – in an aggregation scenario where higher scores indicate potentially higher socio-economic status, lower unemployment rates and low income prevalence is beneficial to the location’s status, therefore during standardization these values should use a cost, or minimizing, function; whereas for the remaining criteria, a benefit function should be used for standardization. After the criteria were selected, the values in each of the criteria needed to be standardized and given appropriate weights. These steps will be elaborated in the following subsections.

3.2.2 Data Standardization Scheme

Given the variety of units of measurement and the need to penalize higher values in some of the criteria through cost functions (in this case, unemployment rates and low-income prevalence), the criteria needed to be normalized and standardized. With normalization already completed, the remaining step is to standardize the data. In the previous section on plugin development, there were three standardization methods at our disposal through the “Field Standardizer” plugin: maximum score, score range, and z-score, each with cost and benefit function variations. To decide which standardization metric to use, a brief analysis of the distribution of each criterion was conducted in R in order to determine the suitability of the data to be standardized, either by rescaling (score range; raw values re-scaled to 0-1) or by calculating its z-scores (mean of criteria set at 0; 1 indicates a standard deviation away from the mean). Usually for z-score standardization, the raw data should have some semblance of a normal distribution for the z-scores to be most effective. To determine if the criteria used in this case study were normally distributed, histograms as well as normal Q-Q plots were generated for analysis, with the results shown in Figures 3 and 4 on the following pages.
Figure 3: Histogram results for the five criteria selected (bin = 50)
Figure 4: Normal Q-Q plots for the five criteria selected
The histograms and normal Q-Q plots show that, perhaps unsurprisingly, the criteria selected did not have normal distributions, and could potentially benefit from log transformations. From the histograms, average dwelling value, average household income, and number of bachelor’s degree holders all skewed quite heavily to the right, and their normal Q-Q plots showed non-linear trends that reinforced the histograms in suggesting non-normality. More importantly, many of the criteria selected had quite a few outliers as well, as shown through boxplots in Figure 5 below.

![Boxplots for the five criteria selected](image)

Figure 5: Boxplots for the five criteria selected

In addition to the histograms and normal Q-Q plots, the Shapiro-Wilk test for normality for each of the five criteria all had p values significantly lower than 0.05, which suggested that the distributions were not normal, as the assumptions for normality were rejected. As a result of these findings, it was decided that z-score standardization was not to be used in this case study, and instead the score range standardization will be applied to all of the criteria for WLC and OWA calculations by using the “Field Standardizer” plugin.
3.2.3 Criteria Weighting Scheme

After the raw values from the five criteria were standardized, the associated weights were chosen. For this case study, a pairwise comparison matrix was used to determine the criterion weights. Tables 6 and 7 below shows the matrices used to generate the criterion weights, based on Tables 1 and 2 in the literature review section.

Table 6: Pairwise comparison for the five chosen criteria

<table>
<thead>
<tr>
<th></th>
<th>UNEMP_RATE</th>
<th>AVG_VAL</th>
<th>AVG_HH_INC</th>
<th>POP_PRV</th>
<th>BACH_DEG</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNEMP_RATE</td>
<td>1</td>
<td>2</td>
<td>0.5</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>AVG_VAL</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>AVG_HH_INC</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>POP_PRV</td>
<td>1</td>
<td>1</td>
<td>0.33</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>BACH_DEG</td>
<td>0.33</td>
<td>1</td>
<td>0.33</td>
<td>0.5</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 7: Criterion weights calculation for the five chosen criteria

<table>
<thead>
<tr>
<th></th>
<th>UNEMP_RATE</th>
<th>AVG_VAL</th>
<th>AVG_HH_INC</th>
<th>POP_PRV</th>
<th>BACH_DEG</th>
<th>WEIGHTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNEMP_RATE</td>
<td>0.21</td>
<td>0.29</td>
<td>0.19</td>
<td>0.15</td>
<td>0.30</td>
<td><strong>0.23</strong></td>
</tr>
<tr>
<td>AVG_VAL</td>
<td>0.10</td>
<td>0.14</td>
<td>0.19</td>
<td>0.15</td>
<td>0.10</td>
<td><strong>0.14</strong></td>
</tr>
<tr>
<td>AVG_HH_INC</td>
<td>0.41</td>
<td>0.29</td>
<td>0.38</td>
<td>0.46</td>
<td>0.30</td>
<td><strong>0.37</strong></td>
</tr>
<tr>
<td>POP_PRV</td>
<td>0.21</td>
<td>0.14</td>
<td>0.12</td>
<td>0.15</td>
<td>0.20</td>
<td><strong>0.16</strong></td>
</tr>
<tr>
<td>BACH_DEG</td>
<td>0.07</td>
<td>0.14</td>
<td>0.12</td>
<td>0.09</td>
<td>0.10</td>
<td><strong>0.10</strong></td>
</tr>
</tbody>
</table>

Sum = 1

Note that the relative importance assigned in Table 6 was not based on domain-specific research, but rather on subjective values for demonstration purposes. After applying the pairwise comparison matrix, it can be seen that average household income and unemployment rate were
given relatively higher weights, whereas the percentage of holders of bachelor’s degrees or above received the lowest criterion weight.

3.2.4 Order Weighting Scheme

In order to compare the effects of different order weights on the selected data, five order weight schemes were included in this analysis: a “pessimistic” scheme for MIN OWA (AND operator; $\alpha = 0$), an equal weighting scheme for WLC ($\alpha = 0.5$), an “optimistic” scheme for MAX OWA (OR operator; $\alpha = 1$), and two schemes for ORness measures of $\alpha = 0.3$ and $\alpha = 0.7$ by utilizing the maximum entropy approach. Table 8 shows the order weights for different $\alpha$ values when the number of criteria used is 5; these are the solutions of the nonlinear Equations (3) and (4) in the literature review section (Wu et al., 2009). The columns that are in bold are the $\alpha$ values that were used for the case study.

Table 8: Sets of order weights from $\alpha = 0$ (pessimistic) to $\alpha = 1$ (optimistic) for five criteria.
Source: Wu et al. (2009)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>0.0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$o_1$</td>
<td>0</td>
<td>0.0050</td>
<td>0.0289</td>
<td><strong>0.0706</strong></td>
<td>0.1277</td>
<td><strong>0.2</strong></td>
<td>0.2884</td>
<td><strong>0.3962</strong></td>
<td>0.5307</td>
<td>0.7105</td>
<td>1</td>
</tr>
<tr>
<td>$o_2$</td>
<td>0</td>
<td>0.0175</td>
<td>0.0599</td>
<td><strong>0.1086</strong></td>
<td>0.1566</td>
<td><strong>0.2</strong></td>
<td>0.2353</td>
<td><strong>0.2574</strong></td>
<td>0.2565</td>
<td>0.2068</td>
<td>0</td>
</tr>
<tr>
<td>$o_3$</td>
<td>0</td>
<td>0.0602</td>
<td>0.1240</td>
<td><strong>0.1672</strong></td>
<td>0.1920</td>
<td><strong>0.2</strong></td>
<td>0.1920</td>
<td><strong>0.1672</strong></td>
<td>0.1240</td>
<td>0.0602</td>
<td>0</td>
</tr>
<tr>
<td>$o_4$</td>
<td>0</td>
<td>0.2068</td>
<td>0.2565</td>
<td><strong>0.2574</strong></td>
<td>0.2353</td>
<td><strong>0.2</strong></td>
<td>0.1566</td>
<td><strong>0.1086</strong></td>
<td>0.0599</td>
<td>0.0175</td>
<td>0</td>
</tr>
<tr>
<td>$o_5$</td>
<td>1</td>
<td>0.7105</td>
<td>0.5307</td>
<td><strong>0.3962</strong></td>
<td>0.2884</td>
<td><strong>0.2</strong></td>
<td>0.1277</td>
<td><strong>0.0706</strong></td>
<td>0.0289</td>
<td>0.0050</td>
<td>0</td>
</tr>
</tbody>
</table>

Note that in Table 8, the weights are mirrored and reversed images of each other (e.g., the order weights for $\alpha = 0.4$ are the weights for $\alpha = 0.6$ but in reverse order) on either side of $\alpha = 0.5$ (WLC). At this point, the criteria were selected and standardized using the Field Standardizer plugin in QGIS, criterion weights were calculated, and order weights were determined. The remaining steps are to input these values into the WLC/OWA Tool plugin in QGIS to calculate their respective OWA scores.
4 CODE IMPLEMENTATION RESULTS

4.1 “Field Standardizer” Plugin

Figure 6 shows the initial landing page of the GUI (frontend) after the “Field Standardizer” plugin is launched by a user.

Figure 6: Initial startup page when “Field Standardizer” plugin is launched
Referencing Figure 6, a typical flow of actions performed by the user of this plugin is as follows:

Step 1: Upon launching the plugin, the user should see a drop-down menu requesting them to select a layer from the current QGIS project to work on.

Step 2: Once a layer is selected, a table underneath the layer selection box will be populated by available fields to be standardized. The user can then select the desired fields.

Step 3: Once the fields are selected, the user can click the “Add selected fields” button to populate a new table underneath with the selected fields, where each field will have an empty space next to it where the user can select which standardization method to apply to that field. The methods are referenced by a whole number from 1-6. The user interface shows a text box titled “Std Code Directory”, indicating which number refers to which standardization method.

Step 4: Once the standardization methods are assigned, the user clicks “OK” to execute the plugin, and the selected fields are updated with their new, standardized values.

Figure 7 shows an example of a filled-out GUI for the “Field Standardizer” plugin, ready for execution. The example shows the input used to standardize the case study data (see Chapter 5).
Figure 7: Example of a populated “Field Standardizer” plugin ready for execution
4.2 “WLC/OWA Tool” Plugin

Figure 8 shows the initial landing page of the GUI (frontend) after the “WLC/OWA Tool” plugin is launched by a user.

Figure 8: Initial startup page of the “WLC/OWA Tool” plugin when launched
Referencing Figure 8, a typical flow of actions performed by the user of this plugin is as follows:

Step 1: Similar to “Field Standardizer”, the user is prompted to select an input layer for this plugin.

Step 2: Once the input layer is selected, the table of available criteria is populated, where the user can then select the criteria of interest from the input layer.

Step 3: When the user clicks “Add selected criteria”, the two tables below are populated by the names of the selected criteria, with empty cells for the “Criteria Weight” and “Order Weight” columns.

Step 4: The user can select preset weights for criteria and order weights under the “Presets” boxes. For Criteria Weight Presets, if “Equal Weights” is selected, the empty cells for criteria weight will be populated by assigning each an equal weight based on the number of criteria selected. If “Manual Entry” is selected, the Criteria Weight cells are cleared for the user to manually input the criterion weights. For the Order Weight Presets, if “MAX” is selected, the first row of the Order Weights table will receive a score of 1, whereas the remaining cells receive 0; if “MIN” is selected, the last row of the Order Weights table will receive a score of 1, whereas the remaining cells receive 0. If “WLC” is chosen, equal weights are assigned to the Order Weights column. If “Manual Entry” is chosen, the order weights table clears and allows user to manually input the order weights.

Step 5: Once the appropriate fields have been filled out, the user then clicks “OK” to execute the plugin. A new field, “OWA Score”, will be added to the layer as a result of running the plugin.

Figure 9 shows an example of a filled-out GUI for the “WLC/OWA Tool” plugin, ready for execution, which has the criterion weights and order weights filled out for the case study to carry out the MAX OWA calculations.
Figure 9: Example of a populated “WLC/OWA Tool” plugin ready to be executed
5 CASE STUDY RESULTS AND ANALYSIS

5.1 Criteria Standardization

First, the standardized values of each criterion were mapped and inspected for trends that may help with understanding some of the OWA results. While the standardization method used is the same (score range) across all criteria, two of the criteria were standardized with the “cost” function and three were standardized with the “benefit” function, so they are presented differently. Figures 10a, 10b, and 10c show the benefit criteria, where Figure 10a represents the standardized average dwelling value, 10b the standardized average household income, and 10c the proportion of people with a bachelor’s degree. Figures 11a and 11b show the cost criteria, which are standardized unemployment rate (Fig. 11a) and standardized prevalence of poverty (Fig. 11b). Note that all maps in this study utilizes the projection system EPSG: 26917, or NAD83 / UTM zone 17N.

Figure 10a: Standardized scores for average dwelling value in Toronto Census Tracts (CTs)
Figure 10b: Standardized scores for average household income criterion in Toronto

Figure 10c: Standardized scores for percentage of bachelor’s degree criterion in Toronto
Figure 11a: Standardized scores for unemployment rate criterion in Toronto

Figure 11b: Standardized scores for prevalence of poverty criterion in Toronto
The standardization of all five criteria was completed in the newly created “Field Standardizer” plugin in QGIS in one run. The inputs for this operation are shown in Figure 7 in the previous chapter. Through this standardization, the raw values in each of the criteria were re-scaled to a range of 0 to 1.

Notice that for the two criteria with cost functions (Figures 11a and b), the lower scores received a darker shade of red, indicating that these CTs have less favourable values for socio-economic status. Since the data were standardized with the cost function, these CTs receiving lower standardized scores actually had higher raw values for that criterion. On the other hand, for criteria with benefit functions (Figures 10a-c), higher standardized scores do mean that the CT does have more favourable values for socio-economic status, and the higher score will be accounted for in the subsequent WLC/OWA calculations.

From the observations made in the benefit criteria, CTs located around midtown Toronto had a clear advantage over most of the other CTs in terms of average household income and average dwelling value, while the High Park area to the west showed some favourable numbers as well. Unsurprisingly, the downtown area did better for the criterion on the percentage of people with bachelor’s degrees. In contrast, the midtown and High Park areas also received relatively higher standardized scores for the two cost criteria, meaning that these CTs also had low poverty prevalence and lower unemployment rates compared to the rest of the city. The combination of the favourable numbers in these criteria will make it likely for the WLC and OWA calculations to conclude that these are the neighbourhoods with better socio-economic status compared to the rest.

The standardized maps also identified some areas that did not perform as well and are likely to receive lower aggregate scores. CTs in Scarborough towards the East end of the city and Etobicoke towards the Northwest end of the city show lower scores across the board from average household income to prevalence of poverty, which makes them likely to be identified as areas of lower socio-economic status compared to the remainder of the city.
5.2 Comparison Between Different OWA Weighting Schemes

After examining the standardization results for each of the criteria, the standardized values were put through the “WLC/OWA Tool” plugin for calculation. The calculations for each scenario were carried out separately. The plugin inputs were shown in Figures 12a-d below. Inputs for the MAX OWA calculation can be seen in Figure 9 in the previous section.

Figure 12a: Inputs for MIN OWA ($\alpha = 0$)

Figure 12b: Inputs for OWA with $\alpha = 0.3$
After the WLC/OWA plugin has been run five times with these different inputs, the calculated scores were mapped. Since different $\alpha$ values for order weights would yield vastly different ranges of final OWA score, which makes it unfeasible to numerically compare the different OWA results, the choropleth maps were classified using the quantiles method with five classes. With the quantiles method, each class will have the same number of features, which supports comparisons between the maps. Figures 13-17 each shows the results of a different OWA calculation, in ascending order of the selected $\alpha$ value, with Figure 15 being WLC.
Figure 13: Mapped MIN OWA Scores ($\alpha = 0$)

Figure 14: Mapped $\alpha = 0.3$ OWA Scores
Figure 15: Mapped WLC Scores ($\alpha = 0.5$)

Figure 16: Mapped $\alpha = 0.7$ OWA Scores
Figure 17: Mapped MAX OWA Scores ($\alpha = 1$)

Figure 18: Main areas in Toronto discussed in analysis
For reference, Figure 18 shows the main areas of interest in the analysis of the mapped OWA scores. The Midtown Area highlighted in green and the High Park/West End Area highlighted in blue serve as generalized focus areas for high-level observations of overall OWA performances, whereas the Downsview Census Tract highlighted in red serves as a localized example of how different $\alpha$ values affect final OWA results.

From observing the coverage of the highest OWA scores, the MAX OWA map from Figure 17 shows a rather scattered distribution of highest scoring CTs instead of defined clusters. Since MAX OWA utilizes an “optimistic” approach, where the largest weighted criteria value gets counted and nothing else in the OWA combination, it appears that some of the CTs identified earlier when examining the individual criteria that suggested potentially lower socio-economic statuses were actually included in some of the areas that received higher OWA scores based on the characteristics of the calculation method. For example, in terms of determining the highest scoring CTs, in a CT where there are four criteria with low weighted criteria values and one criterion with a very high weighted criteria value, that high value can outrank other CTs that may have a much larger number of criteria that have high weighted criteria values. This calculation method likely contributed to the observation that the MAX OWA map showed rather mosaic-like OWA scores for the downtown area proceeding west, where there is a wide range of classes, and some unexpected CTs showing higher OWA scores such as the CTs on the western edge of Toronto in Etobicoke. While the MAX OWA map did show that the midtown and High Park areas had CTs with higher OWA scores, they did not seem to cluster as well as the other maps, which may not be a good identifier for finding broader regions of interest with low or high socio-economic status in Toronto.

The MIN OWA calculation on the other hand shows that the “pessimistic” approach indeed had an effect on the selection of CTs that were deemed to have relatively higher socio-economic status. When compared to the MAX OWA calculation, the CTs deemed to have high socio-economic status by the MIN OWA calculations expanded to a much wider range of CTs in the midtown area, and also included the affluent High Park area which the MAX OWA calculations missed. With the expansion of high scoring CTs in those two neighbourhoods, a more clustered map appeared compared to MAX OWA, identifying a more definitive region of CTs with higher socio-economic status. With the MIN OWA approach only accounting for the
smallest weighted criteria values, it was interesting to see the formation of two expanded clusters; in order for this to happen, it meant that the smallest weighted criteria score in these CTs carry a relatively higher value than the smallest weighted criteria scores in the other remaining CTs, which highlighted the fact that it was likely that these CTs received higher scores in all five of the criteria used in this analysis. In the neighbourhoods that scored lower, it might be just due to a single criterion performing poorly compared to the midtown and High Park neighbourhoods, but because in the MIN OWA calculation it was the only one counted they cannot compete with the CTs that scored high on all five categories. The results demonstrate the extremity of the MIN OWA approach, which is highly sensitive to the lower scoring criteria and, as a result, can be impacted greatly by the assigned weights for the criteria used in the analysis. This can sometimes yield results that are further departures from expectations based on individual observations of each criterion.

The WLC scores when compared to the MAX and MIN OWA scores appear to be right in the middle ground between the two. With the order weights the same, the trade-off between criteria is 1, where any criterion is equally likely to substitute for another. The WLC technique maintains a strong indication that the midtown area, as well as the High Park area in the west, enjoy high socio-economic status, and while the location of CTs with the highest grouping in these two clusters showed some variation to MIN OWA (expansion of cluster in west end; shrinkage of cluster in midtown), it still showed a higher level of clustering compared to MAX OWA. The areas that were expected to have lower scores – Scarborough and Northern Etobicoke – both scored relatively lower with the WLC technique as well, with no major departures from the MIN and MAX OWA methods.

As for the two other OWA scores at $\alpha = 0.3$ and $\alpha = 0.7$ using the maximum entropy approach, the effects of optimistic and pessimistic approaches can also be observed. At $\alpha = 0.3$, where the OWA calculation erred on the more pessimistic side with ORness between 0 (MIN) and 0.5 (WLC), the resulting map resembled the MIN OWA map more, and showed that the higher socio-economic statuses were clustered in the two hot spots identified in MIN OWA – midtown and High Park – perhaps even better than the MIN OWA clustering. At $\alpha = 0.3$, it also appears that there are more distinctions within the CTs that perform well in the downtown area, with more CTs in downtown getting elevated to the highest classification (darkest shade of
green). In comparison, when $\alpha$ was set at 0.7, the resulting map greatly resembled that of $\alpha$ set at 1 (MAX OWA), where the side-by-side comparisons actually do not see the maps differ much at all, with just slightly better clustering compared to MAX OWA. It appears that at lower $\alpha$ values, the sensitivity to criterion weights increases and produces more distinctive maps, compared to when $\alpha$ is set at a higher value resembling higher ORness.

The Downsview CT illustrates the mechanics of MIN and MAX OWA at the local level. It can be observed in the standardized maps for each criterion that the Downsview CT performed well in the “cost” criteria but poorly in the “benefit” criteria, where the CT had low prevalence of poverty and unemployment rates, but also had low average household income, average dwelling values, and percentage of bachelor’s degree holders. In the MAX OWA calculation, the algorithm only counted the largest weighted criterion, in this case the unemployment rate, towards the OWA score, which resulted in the Downsview CT being placed in the highest class of socio-economic status. On the other hand, with the MIN OWA calculation, the algorithm only counted the smallest weighted criterion, here the average dwelling value, which drastically reduced the CT’s relative socio-economic status and relegated it to the bottom class in the MIN OWA map.

This localized example also demonstrates the “risk-taking”, or optimistic, characteristic of MAX/Higher ORness OWA calculations, and the “risk-averse”, or pessimistic, characteristic of MIN/Lower ORness OWA calculations (Malczewski & Rinner 2005). In the MAX OWA calculation at a location such as Downsview, as long as one criterion had a high weighted value, even if the all remaining weighted criteria were low, that location will have the opportunity to outcompete other locations just by using that single high weighted criterion value, hence being more “risk-taking” as only one criterion needed to be satisfied. In the MIN OWA calculation, the opposite is true, where the criterion with the lowest weighted criterion value gets counted; therefore, if even the lowest weighted criterion value performs better than most other locations’ lowest criteria, it meant that all criteria at that location were satisfied in order to receive the higher OWA score, thus making the calculation more “risk-averse”. The other $\alpha$ values of 0.3, 0.5 (WLC), and 0.7 show the spectrum of different risk strategies in between those two extremes that may be used by the user.
From examining these five maps, it appears that with the OWA approach where \( \alpha = 0.3 \) seemed to be the most useful in identifying neighbourhoods and clusters of higher socio-economic status using the five criteria. This approach identified key clusters where most of the criteria are met with higher weighted scores and showed more restraint in including CTs in the higher status category, which suits the purpose of the case study. However, given the proof of concept nature of this case study, note that the assessment of \( \alpha = 0.3 \) being potentially the most suitable ORness measure is merely a suggestion based on the premise of the study. The MAX and \( \alpha = 0.7 \) OWA maps showed maps that were more generalized, where simply the highest weighted criteria values get counted, and that design elevated many CTs to higher socio-economic status categories because of a singular high scoring criterion, when many of the other low scoring criteria would have contributed negatively overall. This characteristic with MAX OWA may therefore be less ideal for the point of this case study. The MIN OWA maps on the other hand showed high sensitivity to lower scoring criteria, and while in general the logical hotspots of higher socio-economic statuses were identified, preference was given overwhelmingly to midtown CTs where they scored highly for all five criteria used, which perhaps left behind some CTs with still higher socio-economic status by that metric. Therefore, while the MIN OWA with its strict, pessimistic approach may be good for subsequent analyses of specific clustered areas that were observed, for the purposes of identifying suitable neighbourhoods the MIN OWA approach looked to be too extreme in this case.

Through the analysis of the output maps it can be seen that different OWA weights can have profound impacts, and depending on the use cases the analysts who decide to use OWA can use the characteristics of ORness to help determine OWA strategies for desired outcomes. Particularly for the cases that were highly sensitive to weighted criteria scores such as the MIN OWA approach, the results also showed the practically endless variations of OWA analyses by having different assigned criterion weights and ordering weights. The differences optimistic and pessimistic approaches were also apparent in the maps, which showed that OWA applications on GIS do reflect their theoretical characteristics.
6 DISCUSSION AND CONCLUSION

One important factor to acknowledge in this research paper is that the calculations used for OWA and WLC are global, meaning that they are spatially implicit. This assumes that the parameters for the criteria do not change based on geographical space, where in many cases they do. An improvement to the current form of the plugin is to introduce a framework in the WLC and OWA calculator to incorporate the “local” versions of each, where the distances between each polygon (neighbours) can be accounted for in the localized approach for a spatially explicit analysis.

This research produced two QGIS Python plugins for quick standardization of fields and for carrying out WLC/OWA calculations, with intuitive interfaces and high user customizability. These two plugins in QGIS should hopefully help analysts conduct OWA and WLC analyses more easily, as well as offer stress-free standardization of data by relieving the user of the need to carry out multiple field calculations on the fly. Through the use of the plugins in the case study, it was also shown that the plugins run in a timely manner, where the OWA calculations were completed within seconds; however, given its usage of two for loops: one to iterate through each criterion and another to iterate through all locations for that criterion, larger datasets may take longer to compute. The ability for users to select as many field/criteria for analysis in both of the plugins not only improves the user experience with clear displays of available layers and fields for calculations, but also provide the ability to add more pre-sets or methods in the future if the need arises.

In the case study, the theoretical effects of OWA weighting schemes were demonstrated in the output maps, with the pessimistic and optimistic approaches each showing the expected categorizations of data. While the case study was focused on social geography using census data for urban neighbourhoods, the same process could be applied to site selection and risk assessment in the physical context, given that the data were provided in vector format and have suitable criteria for consideration. The case study also demonstrated the fact that the variations in criterion weights and order weights can have profound impacts on the resulting OWA scores in an analysis based on the methods selected, which allows high levels of user customization and discretion on a case-by-case basis.
Additional future developments on the plugins can include updated design choices, more standardization protocols, normalization options, and matrix calculations. For the design choices, there are two potential updates for the Field Standardizer. First, the current layout allows the users full latitude in terms of which standardization method to use for any of the selected criteria. A problem with this generous setting is that, especially for WLC and OWA calculations, when criteria are standardized, they should all be based on the same standardization scheme to avoid different re-scaled ranges and distributions, and that restriction was not enforced. Given the intended generalist usage of the Field Standardizer plugin, an improvement to this issue could be an additional checkbox that users can select to voluntarily “opt-in” to restrict to a singular standardization method. Second, the current algorithm overwrites the original fields directly, in order to avoid potentially unintelligible field names due to field name length restrictions. To improve this, perhaps a new box asking users to name the new field can be implemented, but it will come at the cost of making the plugin more complex. As to additional features, the Field Standardizer plugin can offer to normalize a criterion based on another field (e.g., in the case study, normalize the bachelor’s degree criteria by total population), and the WLC/OWA plugin can be further developed to carry out some of the specific criterion or order weight schemes. Criterion weight schemes may be easier to implement; for example, the pairwise comparison matrix using a scale of 1-9 for relative importance can be implemented in a separate window for criterion weight calculation within the plugin, which helps the user of the tool to quickly come up with a wide range of available criterion weights. In terms of order weights, however, given the complexities of the multitude of methods and complex calculations, might be more difficult to implement and take up more computing power, which may cause the plugins to run significantly longer. As open-source software, any potential user will be able to implement these and other improvements or contract out their implementation to an available GIS developer.
APPENDIX

A1. Repository for the Field Standardizer plugin on GitHub

https://github.com/greghuang8/standardizer_plugin

A2. Repository for the WLC/OWA Tool plugin on GitHub

https://github.com/greghuang8/owa_plugin
REFERENCES


