AN INTEGRATED GIS-BASED MULTI-CRITERIA DECISION ANALYSIS (MCDA) APPROACH TO MODEL COVID-19 VULNERABILITY IN GAUTENG SOUTH AFRICA

Master's Research Report

By

Jabulani Jele

2518740

Supervisor

Dr Iqra Atif

Submitted in partial fulfilment of the requirements for the degree of Master of Science in Geographic Information Systems and Remote Sensing in the School of Geography, Archaeology and Environmental Studies, Faculty of Science, University of the Witwatersrand

2021

DECLARATION

I declare that this research report is my own, unaided work. It is being submitted for the Degree of Master of Science in Geographic Information Systems and Remote Sensing at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

____

(Signature of candidate)

15th day of June 20 22 at Pretoria

ABSTRACT

Gauteng province is the epicentre of COVID-19 in South Africa and to date, the province accounts for 32.5% of the country's total COVID-19 cases. The severity of COVID-19 in Gauteng requires specialised methodologies that can help identify areas vulnerable to the disease and for support mechanisms to be directed within them. By integrating two (2) GIS-Based MCDA criteria indicator weight determining algorithms, namely; the subjective analytical hierarchy process (AHP) and the objective Entropy Weight Method (EWM), an integrated ward level COVID-19 vulnerability model for the province was developed. The receiver operating characteristic and area under the curve (ROC-AUC) method was used to validate the accuracy of the models. And with an AUC score of 0.703, the integrated COVID-19 vulnerability model was the most accurate for modelling COVID-19 vulnerability in Gauteng wards, as opposed to the individual AHP and EWM models with AUC scores of 0.634 and 0.686 respectively. The study demonstrates that the integrated COVID-19 vulnerability model is the most ideal model to help guide government intervention efforts in the fight against the disease in Gauteng.

Keywords: Analytical Hierarchy Process (AHP); Entropy Weight Method (EWM); Gauteng wards; spatial modelling; GIS analysis

ACKNOWLEDGEMENTS

I would like to acknowledge everyone who has supported me along the way towards completing this report.

TABLE OF CONTENTS

LIST OF ACRONYMS

AHP	Analytical Hierarchy Process
AUC	Area under Curve
EWM	Entropy Weight Method
GCRO	Gauteng City Region Observatory
GIS	Geographic Information System
MADM	Multiple Attribute Decision-Making
MCDA	Multi-Criteria Decision Analysis
MCDM	Multi-Criteria-Decision-Making
MCE	Multi-Criteria Evaluation
OWA	Ordered Weighted Averaging
ROC	Receiver Operating Characteristic
SAIMD	South African Index of Multiple Deprivation
STATSSA	Statistics South Africa
TOPSIS	Technique for Order of Preference by Similarity to Ideal Solution
WHO	World Health Organisation

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background

The first official case of COVID-19 was reported in the Chinese city of Wuhan in mid-December 2019 (Hengbo, Wei, & Ping, 2020; Mbunge, 2020). Since then, the World Health Organisation (WHO) declared COVID-19 a public health emergency of international concern on January 30th, before declaring it a global pandemic on March 11th 2020. Besides the global pandemic declaration, the WHO issued a list of recommendations of best practices that would protect the public against infection. The recommendations suggested the following:

(1) Social distancing, specifically, keeping a 1-meter distance between self and others; (2) Regular handwashing with water and soap or using an alcohol-based solution; (3) Avoiding spaces that are closed, crowded or involve close contact; and, (4) Avoid touching eyes, nose and mouth (WHO, 2020).

The March 11th declaration and subsequent release of the list of best practices by the WHO prompted governments of the world into action. The most common response taken by world governments to these pronouncements was the institutionalisation of full lockdowns within country borders alongside the emphasis on WHO recommendations for citizens. Reports indicate that more than 100 countries around the world instituted a form of lockdown (either full or partial) as a primary means to curb the rate of infections, with South Africa being one of these countries (Dunford, et al., 2020).

While regarded as a higher middle-income country, South Africa is also considered one of the most unequal in the world (IMF, 2020; Broadbent, Combrink, & Smart, 2020). The extent of inequality in South Africa can be underscored by official statistics indicating that approximately 30 million people (half the population) in the country live below the upper-bound poverty line which was recorded at R1335 per person per month in April 2021 (STATSSA, 2021; Broadbent, Combrink, & Smart, 2020). Official statistics also indicate that approximately one in seven households in the country are located in an informal settlement where overcrowding is common and that a third of households in the country lack access to a reliable water supply (SERI, 2018; Broadbent, Combrink, & Smart, 2020). The disparity in access to resources illustrated in the aforementioned figures points to a large proportion of South Africans being vulnerable to the effects of COVID-19 (Macharia, Joseph, & Okiro, 2020).

Under this context, it is important to appreciate that while the introduction of a national lockdown in South Africa will to a certain extent mitigate against the rapid spread of the disease, a more concerted effort driven through spatial targeting mechanisms needs to be undertaken to pinpoint and address areas identified as vulnerable to the disease (Hamann, 2020). A typical spatial targeting mechanism used to identify areas vulnerable to a disease involves the incorporation of Geographic Information Systems (GIS) and Multi-Criteria Decision Analysis (MCDA) algorithms.

GIS is described as a computer-supported system for collecting, recording, editing, storing and analysing spatial data (Kistemann, Dougendorf, & Schweikart, 2002). It provides a platform to integrate spatial data from multiple databases and scales for visualization, management and analysis (Kirby, Delmelle, & Eberth, 2016). The growing popularity of GIS in epidemiology is due to its capacity for basic mapping and robust geographical analysis (Rodriguez-Morales, et al., 2015). Interestingly, the development of disease vulnerability maps often involves multiple criteria that need to be considered and the framework used for investigating the effect of these criteria on disease vulnerability are the MCDA algorithms and their associated matrices.

The MCDA matrix is a structured framework for evaluating, investigating and analysing complex decisions using multiple and oftentimes conflicting criteria (Sarkar, 2020). In the case of disease risk or vulnerability mapping, GIS and MCDA can be used as toolkits to aggregate geospatial decision criteria and their weights into a vulnerability index map which can assist with the identification of areas to target for the necessary control mechanisms (Devarakonda, Sadasivuni, Nobrega, & Wu, 2020).

In MCDA matrix construction, different algorithms can determine the criteria weights assigned to the matrix. The most common approach or algorithm used to assign criteria weights to an MCDA matrix is the Analytical Hierarchy Process (AHP) and its associated pairwise comparison matrix. In the AHP method, ratio scales are obtained from ordinal scales which are derived from individual judgements using the pairwise comparison matrix (Akhmadullina, Vusilyeva, Yakovleva, Vopiyashina, & Kraineva, 2019). While the AHP method is the most widely used MCDA algorithm, its reliance on the judgements of experts to estimate criteria weights is a central point of criticism for the approach, as it is perceived to introduce an element of subjective bias in the results (Velasquez & Hester, 2013).

To circumvent the subjective bias associated with the AHP method, other MCDA algorithms like the Entropy Weight Method (EWM) have been utilised to estimate the criteria weights used in vulnerability maps (Wu, Li, Qian, & Chen, 2015). Unlike subjective weighting models like the AHP, the EWM methods biggest advantage lies in the fact that it does not rely on human factors i.e. expert opinions or judgements in the estimation of criteria weights, but it rather uses a strictly mathematical approach to assigning weights which enhances the overall objectivity of the model results (Zhu, Tian, & Yan, 2020). Although the model results acquired using EWM are objective, previous studies often critique this method by indicating that the weighted results acquired through its use do not always accurately reflect the importance of the observed criterion (Zhu, Tian, & Yan, 2020).

In this study, six spatially referenced vulnerability indicators: income poverty, education deprivation, living environment deprivation, number of healthcare facilities in the ward, population density and pre-existing health conditions; were used to develop an integrated COVID-19 vulnerability map for Gauteng wards based on AHP and EWM weight assigning algorithms. The fundamental principle toward integrating these two MCDA algorithms was to develop an accurate vulnerability map devoid of the typical criticisms labelled against AHP and EWM based models when used independently. This resulted in developing a vulnerability map that can guide decision-making regarding resource allocation and prioritisation efforts in the fight against COVID-19.

1.2. Research problem statement

The use of GIS-based MCDA algorithms such as the AHP and EWM have previously been successfully implemented to identify at-risk or vulnerable areas to a communicable disease (Dong, 2016). While using AHP and EWM methods have proven successful in vulnerability mapping, their respective criticisms (too subjective and too objective) when used independently are hard to avoid and often put doubt in the minds of those tasked to use these algorithms as decision-making tools.

To circumvent the criticism labelled toward these two GIS-based MCDA algorithms (AHP and EWM) when used independently, this research study integrated both algorithms into a hybrid COVID-19 vulnerability map for Gauteng. The rationale for integrating the two algorithms is that the resulting map will be devoid of the criticisms i.e. too subjective and too objective, resulting in an accurate map of COVID-19 vulnerability that harmonises both biases and can function as a reliable decision-making tool.

1.3. Aims and objectives

1.3.1. Aim

This study aims to integrate two GIS-based multi-criteria algorithms including the Analytical Hierarchy Process (AHP) and the Entropy Weight Method (EWM) to map COVID-19 vulnerability in Gauteng.

1.3.2. Objectives

1. To design a GIS-based MCDA matrix and evaluate criterion weights subjectively and objectively using AHP and EWM models.

2. To map the COVID-19 vulnerability in Gauteng South Africa based on AHP, EWM and the integrated model.

3. To validate the model results using Receiver Operating Characteristic (ROC) and Area Under Curve (AUC) method.

1.4. Research questions

1. Will the criteria weights used in the MCDA matrices for both the AHP and EWM models be given similar priority?

2. Which vulnerability map (AHP or EWM) is more accurate for mapping COVID-19 vulnerability in Gauteng?

3. Is the integrated approach of COVID-19 vulnerability mapping more accurate than the individual AHP and EWM models?

2. LITERATURE REVIEW

The approach taken towards constructing this literature review initially involved the identification and subsequent review of articles related to COVID-19 vulnerability mapping. The identification of common trends, patterns and techniques involved in COVID-19 vulnerability mapping supported the core write up in the first part of the review.

The second part of this literature review involved the review of articles with a focus on multiple criteria decision analysis (MCDA) and the algorithms associated with its use i.e. the Analytical Hierarchy Process (AHP) and Entropy Weight Method (EWM). The identification of common trends, patterns and techniques involved in MCDA and the associated algorithms supported the core write up in the second part of the review.

2.1. COVID-19 vulnerability mapping

The reviewed literature on COVID-19 vulnerability mapping illustrates evidence of a common understanding among authors of what vulnerability to COVID-19 entails. Typically, as explored by Macharia et al (2020), the most vulnerable groups to a particular disease or pandemic as is the case with COVID-19, are those with the reduced ability to avoid significant financial and human loss as a result of the disease (Macharia, Joseph, & Okiro, 2020). In a similar vein, Archarya and Porwal (2020) articulate that vulnerability to COVID-19 is a dynamic concept that not only applies to the epidemiologically vulnerable e.g. the elderly and those with comorbidities, but it also applies to people in socio-economic conditions that would make them less capable of coping with the crises (Acharya & Porwal, 2020).

The uniform understanding and interpretation of what vulnerability is among researchers translates into a uniform understanding of the criteria indicators to use when modelling COVID-19 vulnerability. While there is variation in the specific criteria indicators used by different authors, COVID-19 vulnerability mapping is typically modelled through the consideration of criteria indicators that can be grouped into three categories, namely: socio-economic, demographic and health-related factors (Rahman, Zatri, Ashik, Waliullah, & Khan, 2021). A vulnerability index was successfully modelled for Kenya by Macharia et al (2020) by considering criteria indicators such as car ownership, education attainment, water source, elderly population, access to hospitals, HIV, obesity, smoking etc. (Macharia, Joseph, & Okiro, 2020). Similarly, another was successfully developed for India by Malakar (2021) by considering criteria indicators such as population density, level of education, employment status and hospital access (Malakar, 2021).

While there is a level of consensus within the literature on the indicators to use for modelling COVID-19 vulnerability, there is a high level of variation regarding the number of criteria indicators to use for successfully modelling COVID-19 vulnerability. From the set of reviewed literature for this study, the number of criteria indicators used to model COVID-19 vulnerability ranged between six (6) and twenty-four (24) indicators. Six (6) criteria indicators were successfully used by Ali et al (2021) to model COVID-19 vulnerability in the United States of America (Ali, Mortula, & Sadiq, 2021). Whereas Macharia et al (2020) utilised twenty-four (24) indicators to model the COVID-19 vulnerability in Kenya (Macharia, Joseph,

& Okiro, 2020). The disparities in the number of criteria indicators used to model COVID-19 vulnerability among the articles reviewed for this study gives the impression that it can be successfully modelled with a minimum of six (6) criteria indicators provided that they are within the categories of health, demography and socio-economic factors (Rahman, Zatri, Ashik, Waliullah, & Khan, 2021).

The reviewed literature explores a myriad of methodologies used to model COVID-19 vulnerability around the world, and the range of methodologies adopted by scholars varies in their level of complexity. Some of the more simple methodologies adopted involved the normalization of criteria indicators to a common scale, followed by the spatial overlay of the indicators via an equal weight or a ranking system (Macharia, Joseph, & Okiro, 2020; Acharya & Porwal, 2020). Another fairly simple methodology employed by authors involved the use of a variety of GIS-based MCDA approaches to assess the independent impact that each criteria indicator has on COVID-19 vulnerability. These included studies by Sarkar (2020) which investigated COVID-19 susceptibility in Bangladeshi districts using the AHP (Sarkar, 2020). As well as studies that integrated two MCDA approaches in their modelling of COVID-19 vulnerability, such as the study by Malakar (2021) which integrated AHP and technique for order preference by similarity to ideal solution (TOPSIS) algorithms (Malakar, 2021). MCDA and its associated algorithms support this research study focus and will be explored in greater detail in the second section of the literature review.

The more complex methodologies employed by researchers in modelling COVID-19 vulnerability involved the use of machine learning algorithms as was the case in the study by Pourghasemi et al (2020) where a GIS-based machine learning algorithm was used for the assessment of COVID-19 outbreak risk and vulnerability in the Iranian province of Fars (Pourghasemi, et al., 2020). Other fairly complex methodologies employed by researchers involved the use of different variations of regression analysis. The study by Ali et al (2020) successfully modelled COVID-19 vulnerability on a five-level Likert scale using logistic regression (Ali, Mortula, & Sadiq, 2021), while Manda et al (2021) employed the use of spatial statistical regression models in modelling COVID-19 vulnerability across the African continent (Manda, Darikwa, Nkwenika, & Bergquist, 2021).

Overall, the review of literature on COVID-19 vulnerability mapping revealed that the majority of studies were conducted to assist with government prioritisation efforts in the fight against COVID-19. The resounding consensus among authors is that these studies provide a framework that can help guide the facilitation of where best to direct support mechanisms (Ali, Mortula, & Sadiq, 2021; Acharya & Porwal, 2020).

2.2. Multi-Criteria Decision Analysis (MCDA)

Literature on multi-criteria evaluation is inconsistent regarding the acronym used to describe the use of multiple criteria as a decision-making tool. The reviewed literature gives the impression that the phrases multiple criteria decision-making (MCDM), multiple criteria decision analysis (MCDA) and multiple attribute decision-making (MADM) can all be used interchangeably as they describe a framework to address problems that take into account the use of more than one criteria indicator (Nymbili & Erden, 2020; Alemdar, Kaya, Codur, Campisi, & Tesoriere, 2021).

The component that gives MCDA methodologies their spatial problem-solving attribute is their integration with GIS. GIS essentially provides a platform to merge data from multiple sources and spatial scales for visualization, management and analysis (Kirby, Delmelle, & Eberth, 2016). A common approach used by researchers when applying GIS-based MCDA methodologies in their studies, involves them overlaying spatially referenced criteria indicators onto a GIS platform (Macharia, Joseph, & Okiro, 2020). This approach allows the researcher to produce thematic or colour-shaded maps that are useful in identifying areas of interest i.e. ideal sites for settlement planning as well as areas at risk or vulnerable to certain environmental phenomena or diseases (Kirby, Delmelle, & Eberth, 2016).

There are various examples of studies that effectively merged GIS and MCDA approaches in their pursuit to address study objectives. From the reviewed literature, the majority of articles could be classified into those that merged GIS and MCDA with an epidemiological basis, and those that merged GIS and MCDA with an environmental sciences basis. Some of those that were analysed with an epidemiological basis included the study by Boyaci and Sisman (2021) that used this methodological approach to identify ideal sites for locating a pandemic hospital in Istanbul, Turkey (Boyaci & Sisman, 2021). Similarly, the study by Sarkar (2020) which identified districts in Bangladesh most susceptible to COVID-19 was undertaken with an epidemiological basis in mind (Sarkar, 2020). On the other hand, the studies reviewed using this methodology with an environmental sciences basis included one by Gigovic et al (2019) which successfully managed to produce a landslide susceptibility map for the western region of the Republic of Serbia (Gigovic, Drobnjak, & Pamucar, 2019). Meanwhile, a study by Yariyan et al (2020) managed to effectively model an index that improved the accuracy of detecting seismic event vulnerability in Sanandaj City Iran (Yariyan, Avand, Soltani, Ghorbanzadeh, & Blaschke, 2020).

While the majority of articles reviewed for this study included those in the fields of epidemiology and environmental sciences, it is important not to lose sight of the fact that GIS-based MCDA methodologies are typically multidisciplinary and therefore cut across a wider range of disciplines than the ones explored in this study. The multidisciplinary nature of GIS-based MCDA methodologies affects how certain aims or study purposes are framed. For instance, Yue (2010) framed the purpose of MCDA approaches as processes that are used to find the most optimal alternative from a set of feasible alternatives with respect to a finite set of attributes (Yue, 2010). The framing used by Yue (2010) is befitting to using MCDA methodologies in the fields of economics and the built environment. On the other hand, Devarakonda et al (2020) framed MCDA methodologies with an environmental and health sciences approach in mind, by highlighting that MCDA methodologies identify the most vulnerable areas that require strong prevention and control measures (Devarakonda, Sadasivuni, Nobrega, & Wu, 2020).

A key attribute involved in MCDA analysis is the algorithm used to evaluate the criteria associated with the study. The MCDA algorithm used in a study plays a pivotal role in determining how the relative importance of the various criteria indicators in the study will be evaluated (Nymbili & Erden, 2020). There are a plethora of MCDA algorithms utilised in many study contexts throughout the literature. Some of the most explored include the analytical hierarchy process (AHP), the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), the Entropy Weight Method (EWM) and Ordered Weighted Aggregating (OWA).

Most researchers have demonstrated an ability to achieve their study aims using a single MCDA algorithm. This was exemplified in the study conducted by Shadeed and Alawna (2021) who used the AHP as the primary algorithm in deriving a COVID-19 vulnerability model for Palestine (Shadeed & Alawna, 2021), as well as the study by Devarakonda et al (2020) which illustrated the effectiveness of using the EWM in identifying regions prone to vector-borne diseases in Illinois, Chicago (Devarakonda, Sadasivuni, Nobrega, & Wu, 2020).

While the use of singular MCDA algorithms has proven successful in achieving study aims (as illustrated in the aforementioned studies), researchers have developed techniques that integrate two MCDA algorithms to form hybrid weight assigning criteria models (Nymbili & Erden, 2020; Boyaci & Sisman, 2021). Typically, hybrid models are developed to mask any potential shortcomings that may exist when using an MCDA algorithm independently (Das & Pal, 2020).

The following three sections of this literature review explore in greater detail literature on the two MCDA algorithms featured in this study and the concept of MCDA algorithm integration.

2.2.1. The Analytical Hierarchy Process (AHP)

The Analytical Hierarchical Process (AHP) is a well-known and commonly used MCDA algorithm for assigning criteria weights which are determined through the pairwise comparison matrix (Saaty, 2008). The method factors in both quantitative and qualitative information to inform a decision with the opinions of expert decision-makers (DM) playing a pivotal role in the criteria weighting (Alemdar, Kaya, Codur, Campisi, & Tesoriere, 2021).

The AHP's popularity among researchers from a wide range of disciplines is due to its relative ease of use, the reliability of results generated and its overall flexibility as a weight assigning algorithm (Velasquez & Hester, 2013; Mishra, Gayen, & Haque, 2020). Evidence of its flexibility and multidisciplinary application is demonstrated in its use in studies by Alemdar et al (2021), Shaheed and Alawna (2021) as well as Mishra et al (2020) who all used the algorithm in health and epidemiology related studies (Alemdar, Kaya, Codur, Campisi, & Tesoriere, 2021; Shadeed & Alawna, 2021; Mishra, Gayen, & Haque, 2020). On the other hand, Das et al (2021), as well as Das and Pal (2020), used the AHP in social sciences and environmental sciences related studies respectively (Das, Das, Giri, Sarkar, & Saha, 2021; Das & Pal, 2020).

While the flexibility of the AHP is demonstrated in its use in the aforementioned studies, its reliability as a decision-making tool was demonstrated by a comparative study conducted by Dong (2016). This study looked to compare the accuracy of the models derived using AHP, TOPSIS and OWA to assess Dengue fever susceptibility in Ecuador and Iran (Dong, 2016). The results from the comparison indicated that the AHP offered the best and most accurate predictive capacity in both scenarios (Dong, 2016).

The methodological approach used by researchers when applying the AHP as a weight assigning algorithm involves four (4) key steps namely; 1) Ascertaining and defining the criteria; 2) Creating a pairwise matrix for comparing the relative importance between criteria indicators with the assistance of experts; 3) Normalization of the pairwise matrix; and 4) Checking the level of consistency of the normalized matrix by calculating its consistency ratio

(CR) which must be less than 0.1 (Alemdar, Kaya, Codur, Campisi, & Tesoriere, 2021; Mishra, Gayen, & Haque, 2020).

Expert feedback is a crucial component in AHP criteria weight assigning and plays a big role in ensuring the accuracy of the model output (Sarkar, 2020). While expert feedback is crucial to AHP model accuracy, the reliance upon it relegates the AHP to being labelled a subjective weight assigning algorithm that is prone to internal bias (Velasquez & Hester, 2013; Nymbili & Erden, 2020; Devarakonda, Sadasivuni, Nobrega, & Wu, 2020). The subjective bias associated with AHP model outputs is a common deterrent against its use. To circumvent this bias, research studies such as the ones done by Devarakonda (2020), as well as Panchal and Shrivastava (2021), opted to use an alternative and less subjective MCDA algorithm called the Entropy Weight Method (EWM), which is known for its objective approach to criteria weight determination (Devarakonda, Sadasivuni, Nobrega, & Wu, 2020; Panchal & Shrivastava, 2021).

2.2.2. The Entropy Weight Method (EWM)

The EWM is a weight assigning algorithm that calculates weights among a set of criteria indicators based on the amount of information contained within them (Devarakonda, Sadasivuni, Nobrega, & Wu, 2020). The appeal of the EWM as a weight assigning algorithm lies in its elimination of subjective factors such as the influence of expert decision-makers in its determination of criteria weights (Wang & Zhang, 2015; Devarakonda, Sadasivuni, Nobrega, & Wu, 2020).

The literature indicates that the EWM has seen wide use as a decision-making algorithm in recent years. It also indicates that the method has found extensive use and support in environmental studies research (Zhu, Tian, & Yan, 2020). Studies that have successfully utilised its objective approach to weight allocation of decision-making criteria include the study by Wu et al (2017) where the algorithm was used to develop a water quality index in North-West China (Wu, Xue, Tian, & Wang, 2017), as well as a study by Wang and Zhang (2015) who successfully used it to evaluate stress factors that may threaten water resource security in Chongqing China (Wang & Zhang, 2015). The algorithm has also seen successful implementation in landslide studies particularly those by Sharma et al (2010) where it was used to evaluate its influence on landslide causing parameters (Sharma, Ghose, Patel, & Debnath, 2010), as well as the study by Panchal and Shrivastava (2021), where it was one of the featured algorithms in a comparative study of landslide susceptibility models (Panchal & Shrivastava, 2021).

While its use in environmental-related studies is well documented, the algorithm has not seen as prolific use in studies related to epidemiology. From the set of the reviewed literature, evidence of its use in epidemiological related research is seen in studies by Devarakonda et al (2020) where it was used to assign criteria weights to factors related to vector-borne disease (Devarakonda, Sadasivuni, Nobrega, & Wu, 2020), as well as the study by Zafar et al (2021) that used the algorithm to develop a Dengue vulnerability index for Lao People's Democratic Republic and Thailand (Zafar, et al., 2021).

The aforementioned studies give evidence of the effectiveness of the EWM approach as a weight assigning algorithm, however, the literature indicates that the model is not flawless and that certain caveats exist with its use (Zhu, Tian, & Yan, 2020). The main caveat that a researcher needs to take into account when using the EWM as a primary algorithm is the fact that the weights derived by the EWM, do not always accurately reflect the importance of the criteria indicator in question (Zhu, Tian, & Yan, 2020). It is due to this reason that some researchers opt to use the EWM as a hybrid model than as the sole independent weight assigning algorithm in their studies. The study by Nyimbili and Erden (2020) deliberately integrated EWM and AHP algorithms as a means to circumvent the shortcomings or caveats associated with the use of each model independently (Nymbili & Erden, 2020).

2.2.3. MCDA algorithm integration

The integration of two MCDA algorithms into a hybrid model is an approach to criteria indicator weight determination that seeks to reduce the shortcomings associated with their independent use (Das & Pal, 2020). Additionally, researchers adopt this approach to derive more accurate and robust results from their models. A large body of research promotes the concept of integrating MCDA algorithms to derive more accurate model results. The study by Jam et al (2021) for instance, concludes by indicating that the use of the AHP in combination with another MCDA algorithm is the most ideal methodological approach to use for generating accurate landslide susceptibility models (Jam, Mosaffaie, Sarfaraz, Shadfar, & Akhtari, 2021).

Across the reviewed literature, the AHP seems to be the algorithm of choice when the methodology of MCDA algorithm integration is adopted. The preference of the AHP as one half of an integrated model could be due to its relative ease of use, the reliability of results generated and its overall flexibility as a weight assigning algorithm (Velasquez & Hester, 2013;

Mishra, Gayen, & Haque, 2020). Evidence of integrating the AHP and EWM algorithms (which is the methodological approach of interest for this study) can be observed in the study by Nyimbili and Erden (2020). In their study, this methodology was used to assign criteria weights related to the assessment of suitable areas for planning urban emergency facilities in Istanbul Turkey (Nymbili & Erden, 2020).

Interestingly, evidence of the integration of these two algorithms in an epidemiological study or to identify disease vulnerability does not seem to exist. As such, this research study was undertaken explicitly to potentially fill this void.

3. MATERIALS AND METHODS

3.1. Study area

Gauteng, the smallest but most populous province in South Africa with a land area coverage of 18,178km^2 is the study area of interest (see figure 1). The province is on a latitude of -26.270760°S and a longitude of 28.11268°E, and it is bordered by Limpopo, the North-West, Mpumalanga, and the Free State province. 2020 population estimates indicate that the province had a population size of 15.5 million people, which equates to 26% of South Africa's total population (STATSSA, 2020). Gauteng is divided into three metropolitan municipalities, namely The City of Ekurhuleni, City of Johannesburg and City of Tshwane. It is also home to two district municipalities, namely Sedibeng and The West Rand.

The province serves as South Africa's economic hub and it is often referred to as the "heartbeat" of the country's commercial and industrial sectors (Municipalities, 2020). Being an economic hub for the country, the province not only attracts a lot of inward migration, but it also sees a fair share of immigrants from other parts of the continent and the world setting up residency within it. Like all other provinces in South Africa, Gauteng experiences many challenges. These challenges include high levels of poverty, unemployment and crime among many others. The challenges faced in the province are underscored by the extreme levels of inequality evident within its cities and towns (GCRO, 2021). The province is associated with pockets or spatial concentrations of huge wealth alongside bigger pockets or spatial concentrations of extreme poverty and deprivation in the form of townships and informal settlements (GCRO, 2021).

Gauteng has been the epicentre of the COVID-19 pandemic in South Africa, with record numbers of daily cases (approximately 11000 per day) being experienced within the province during the countries 3rd wave of COVID-19 infections (ENCA, 2021). Currently, the Omicron fuelled 4th wave of COVID-19 infections has compounded Gauteng provinces position as the epicentre of the disease, with figures indicating that as of the 11th of December 2021, the province had 1.1 million cumulative COVID-19 cases (Meterlerkamp, 2022). This figure equates to 32.5% of South Africa's total COVID-19 cumulative cases tally of 3.5 million.

The state of affairs in Gauteng enhances the relevance of the study methodology for the province. The study can play a crucial role in assisting government stakeholders to identify wards most vulnerable to COVID-19 so the necessary control mechanisms can be introduced within them.

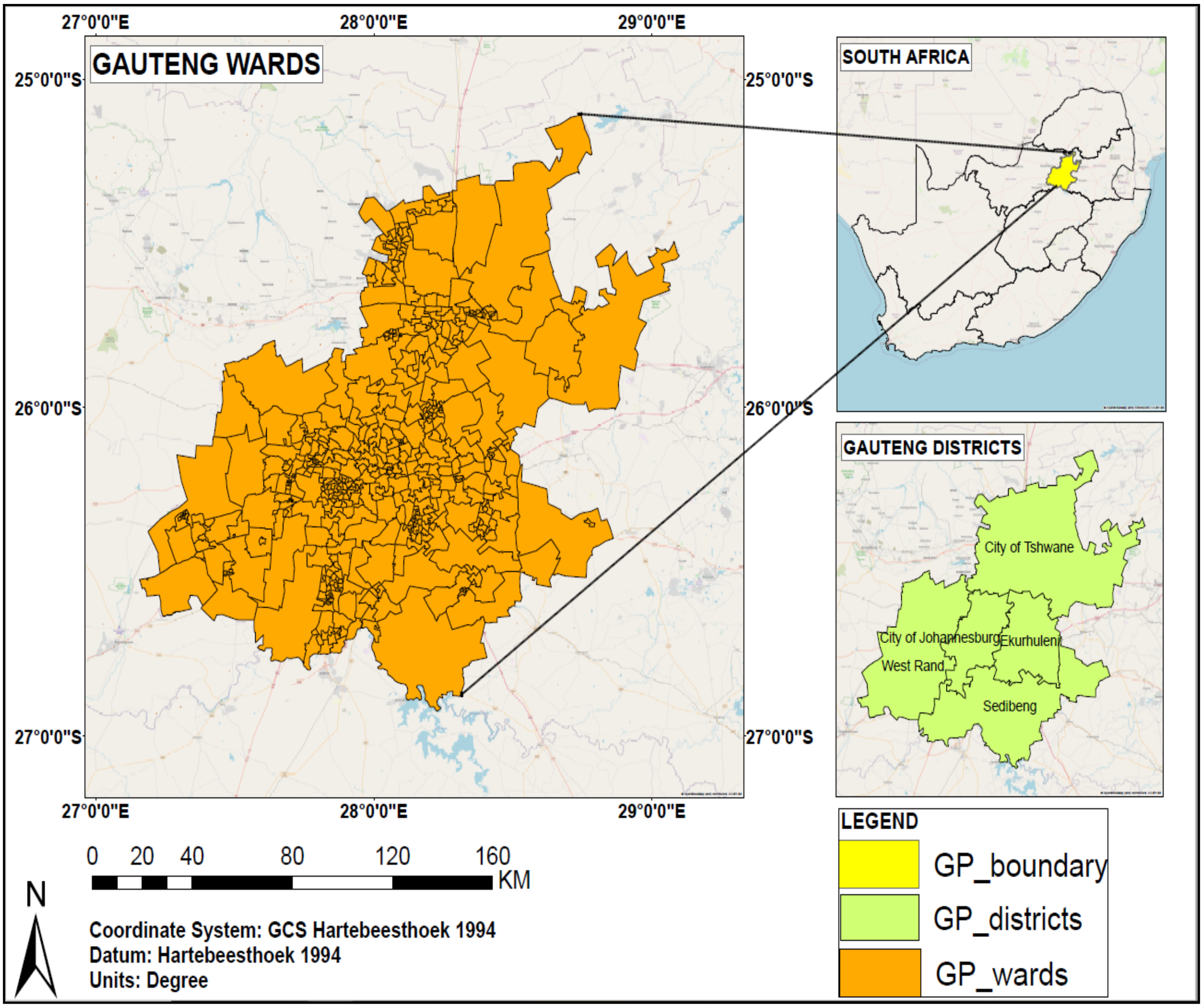

Figure 1: Study Area Map

3.2. Research design

The detailed and multi-phased workflow implemented for this study is illustrated and described in figure 2.

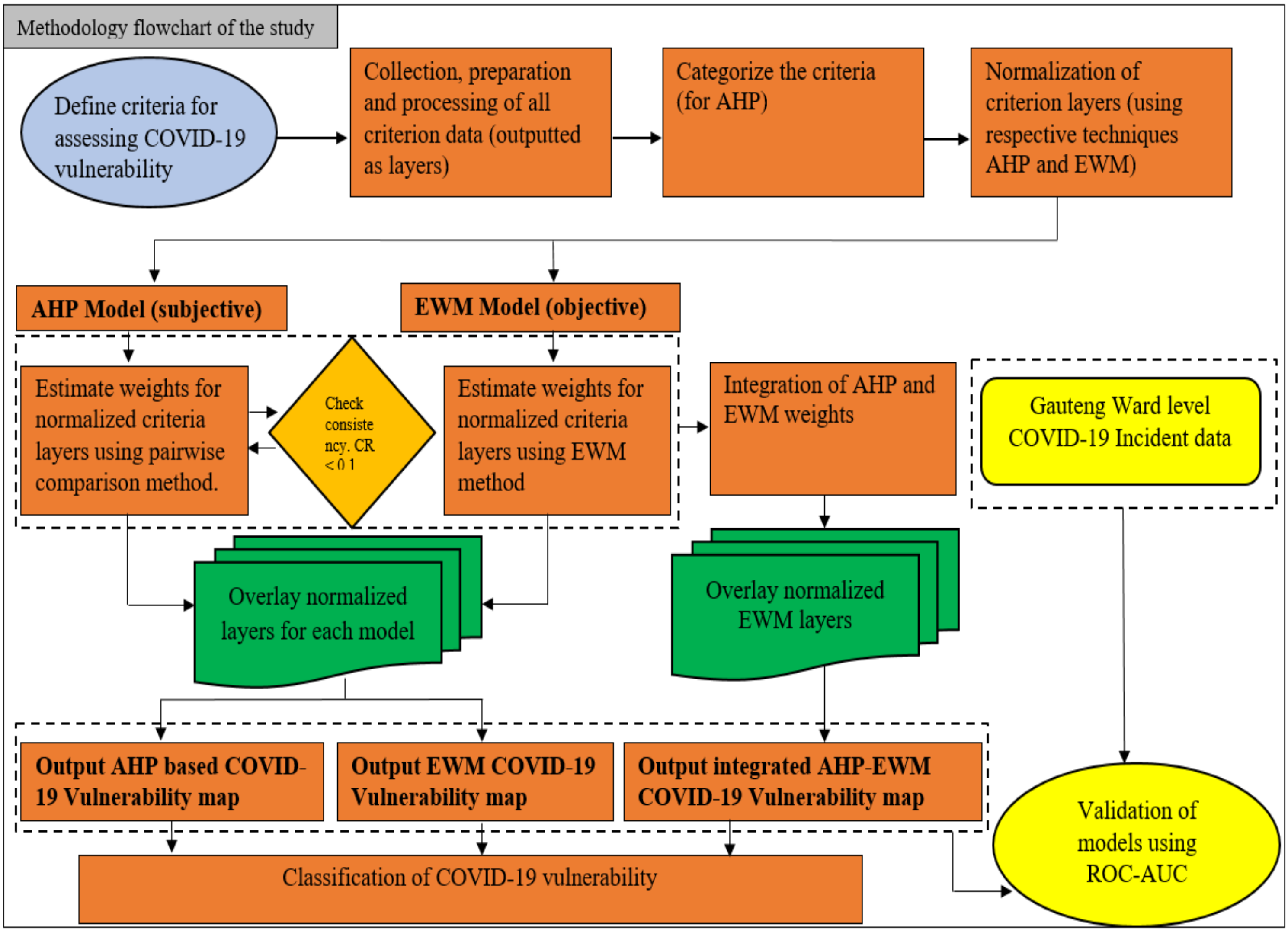

Figure 2: Methodological flowchart

3.3. Data requirements

Table 1 provides information on the dataset, sources and details for the individual pieces of data used to complete the study.

Table 1: Datasets used for the research

Dataset	Data source	Details
GCRO Quality of life V survey data (2017/2018)	Gauteng City Region Observatory (GCRO)	Used for the ward level data on the following COVID-19 vulnerability indicators; • Population density • Number of people with one or more pre-existing health conditions
South African Index of Multiple Deprivation 2011 (SAIMD 2011)	Southern African Social Policy Research Institute (SASPRI)	Used for the ward level deprivation measures on these criteria; • Income poverty • Education deprivation • Living environment deprivation
Health facility database managed by public sector	World Health Organisation (WHO)	Used for the ward level criteria data on the following; • The number of public health facilities at ward level.
Gauteng ward level COVID-19 incident data	National Institute for Communicable Disease	This dataset was used for validating the final AHP, EWM and integrated models.
National and subnational administrative boundaries of SA	Municipal Demarcation Board (MDB)	Administrative boundary layers covering the study area extent.
Participation from field expert	Experts in social sciences, policymaking, public health etc.	Three experts were recruited for completing the AHP's pairwise comparison matrix.

3.4. Methods

The process toward developing the COVID-19 vulnerability maps for Gauteng was broken up into four (4) steps namely; (1) Defining the criteria, (2) Normalization, (3) Assigning criteria indicator weights and mapping (4) Validation.

3.4.1. Defining the criteria

The literature review indicated that COVID-19 vulnerability is a dynamic concept that not only applies to the epidemiologically vulnerable (elderly and sick), but that it also encapsulates those stuck in socio-economic conditions that would make them less capable of coping with the crises over the long term (Macharia, Joseph, & Okiro, 2020; Acharya & Porwal, 2020). Additionally, the literature review also indicated that the criteria indicators used in COVID-19 vulnerability mapping, are typically those that can be categorized into socio-economic, demographic and health-related factors (Rahman, Zatri, Ashik, Waliullah, & Khan, 2021).

With this knowledge, the six criteria indicators used to assess COVID-19 vulnerability in Gauteng wards are defined and illustrated below.

1. **Income Poverty**

Income poverty describes the ward level number of Gauteng residents living below the upper-bound poverty line (R779 per person per month) in 2011 (Noble, Zembe, Wright, & Avenell, 2013). Income poverty is crucial in assessing COVID-19 vulnerability because it determines the services and the quality of environment an individual has access to. The poorer the individual, the less ideal their living environment is likely to be towards COVID-19 prevention and the less likely the individual is to afford quality medical care. Income poverty is also considered a COVID-19 vulnerability indicator because it considers individuals who could not withstand the financial burden caused by prolonged COVID-19 isolation campaigns. Figure 3 illustrates the distribution of income poverty in Gauteng wards.

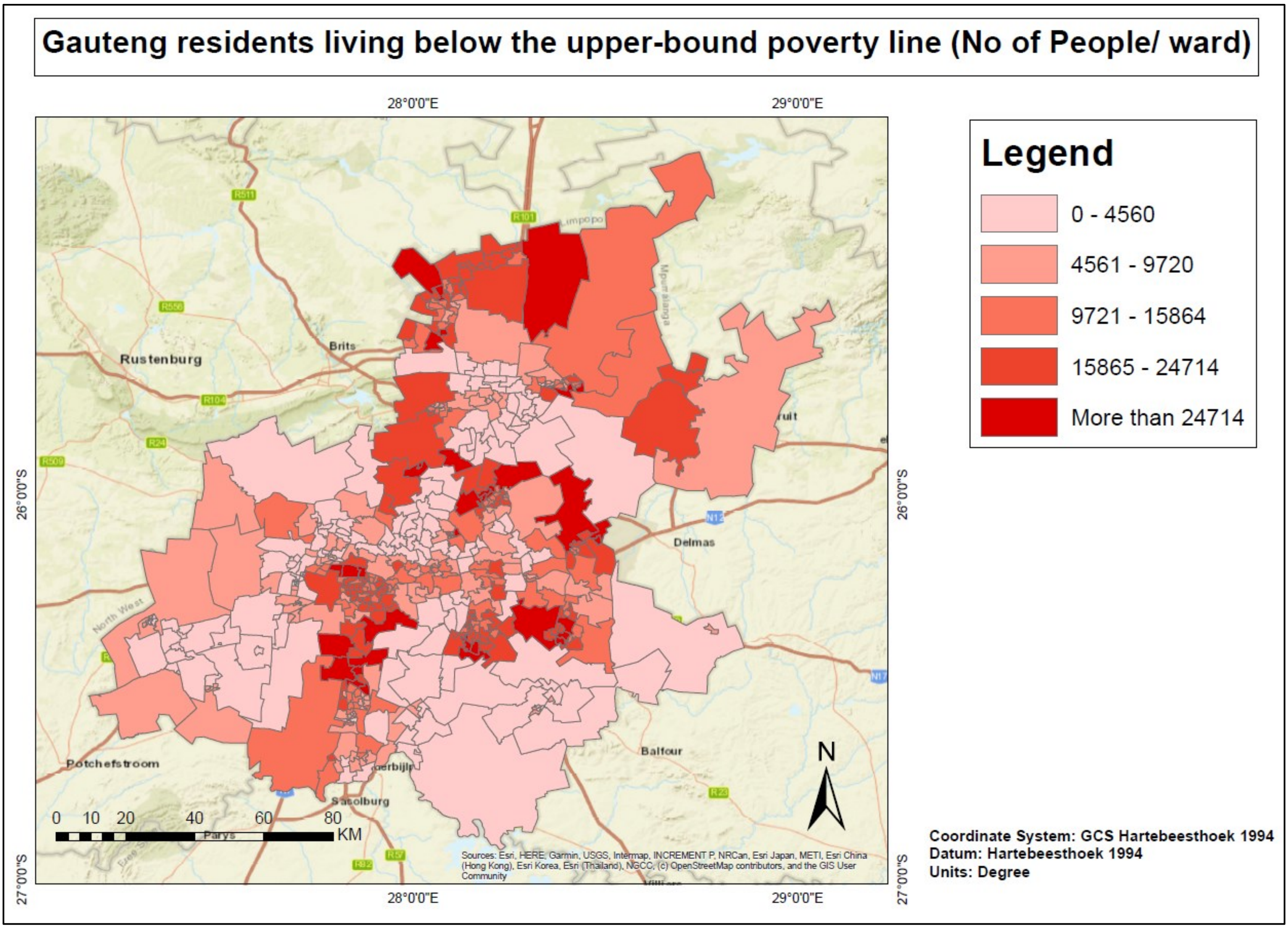

Figure 3: Distribution of income poverty in Gauteng wards

2. Education deprivation

Education deprivation is an indicator that profiles the ward level number of Gauteng residents who are between 18-64 years old that have no secondary education (Noble, Zembe, Wright, & Avenell, 2013). Education deprivation in this study is a proxy for the level of awareness an individual will likely have regarding COVID-19 and the prevention strategies recommended for it. The assumption is that people with lower levels of education will be more vulnerable to COVID-19. Figure 4 illustrates the distribution of education deprivation in Gauteng wards.

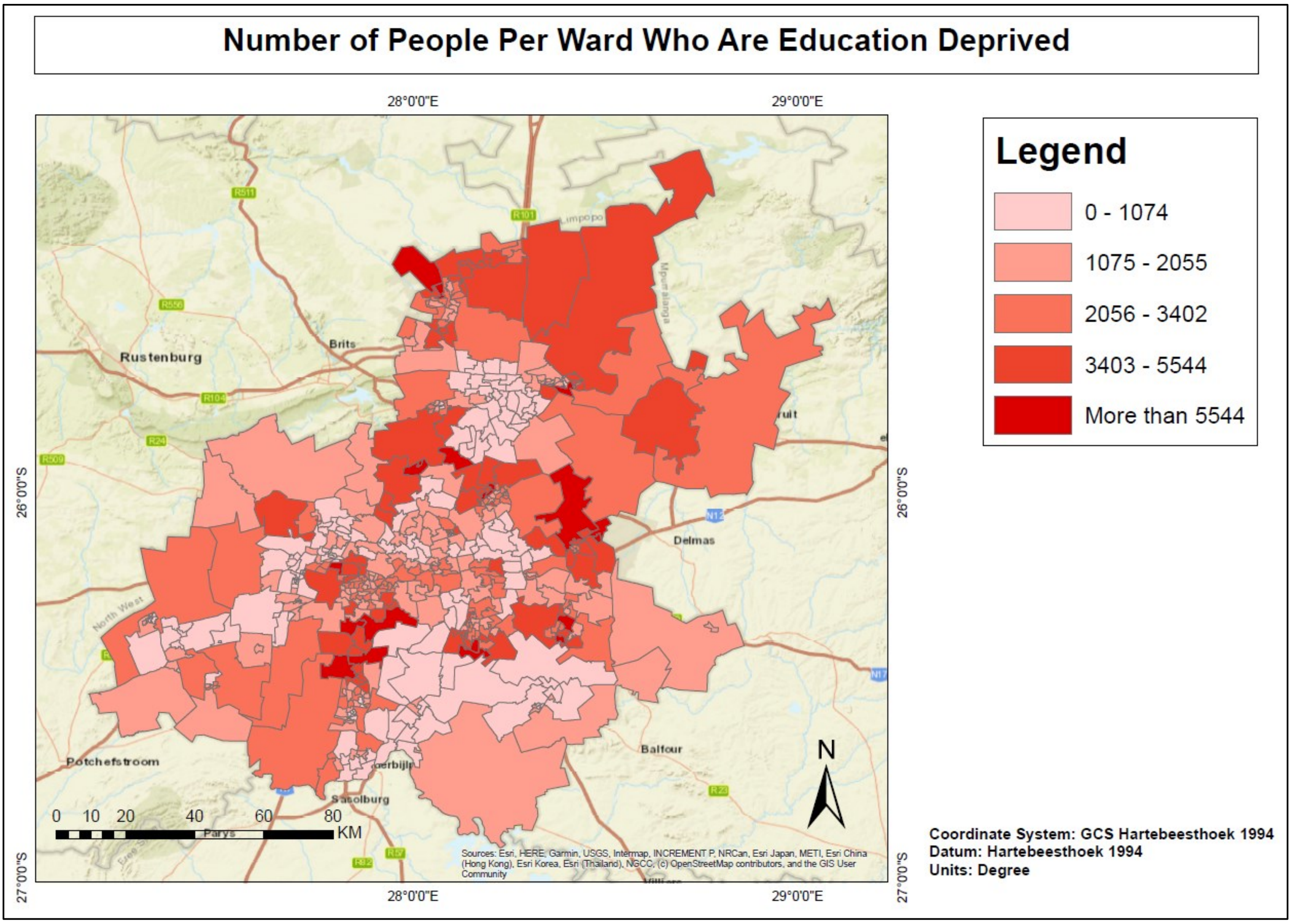

Figure 4: Distribution of education deprivation in Gauteng wards

3. Living environment deprivation

Living environment deprivation is an indicator that profiles the ward level number of Gauteng households per municipality which have inadequate water supply, or which have inadequate sanitation, or which do not use electricity as the main source for lighting, or which are housed in a shack (Noble, Zembe, Wright, & Avenell, 2013). Living environment deprivation is an indicator for COVID-19 vulnerability because it profiles those living in conditions that make them less capable of coping with the crises long term. Figure 5 illustrates the distribution of living environment deprivation in Gauteng wards.

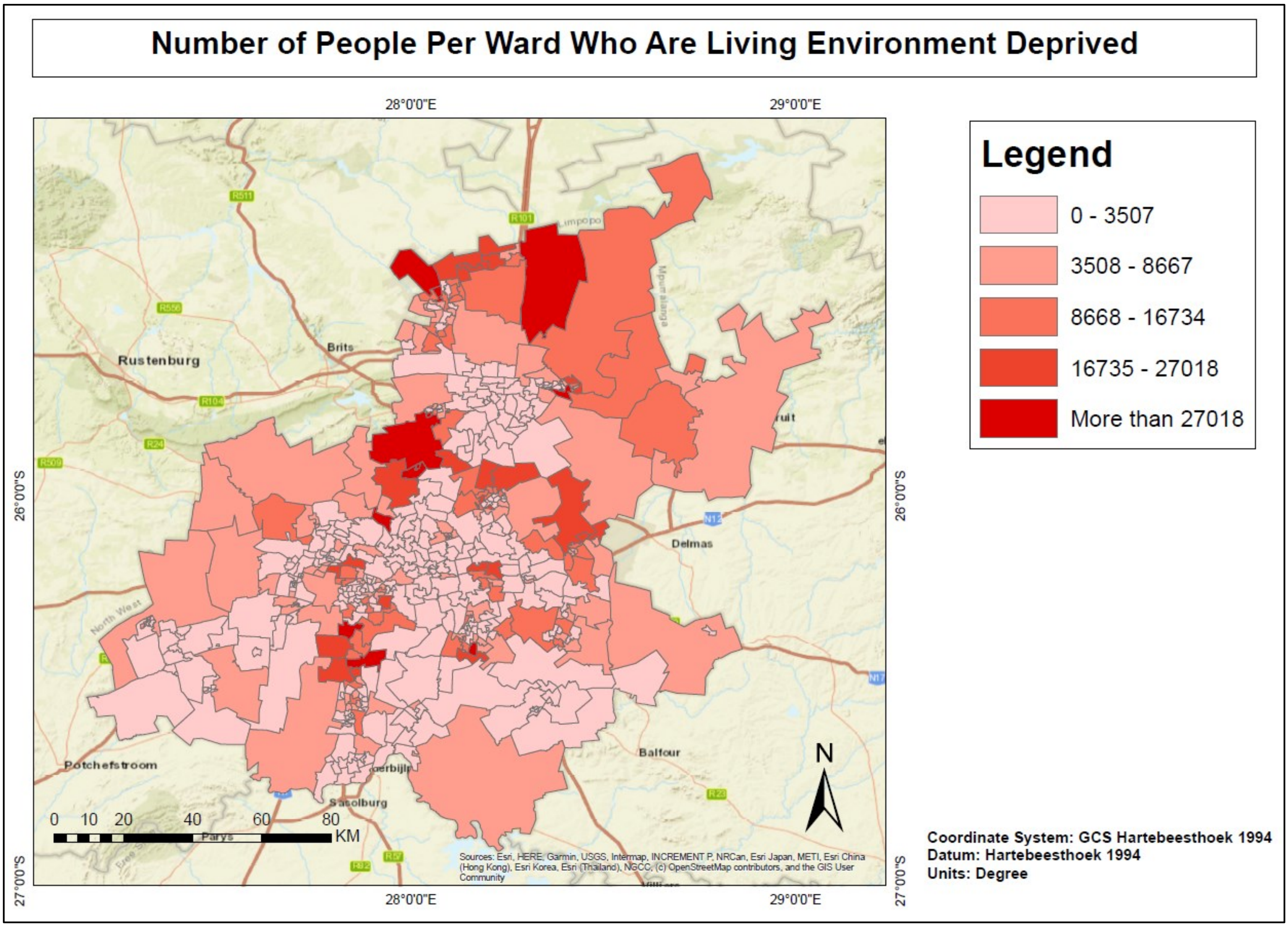

Figure 5: Distribution of living environment deprivation in Gauteng wards

4. The number of healthcare centres in the ward

To effectively manage a pandemic, the ease at which people access the healthcare system plays an important role (Acharya & Porwal, 2020). In other words, having immediate or readily accessible healthcare facilities is important to COVID-19 mitigation. With this logic, Gauteng wards with fewer healthcare facilities (clinics and hospitals) would be regarded as more vulnerable to COVID-19. Figure 6 illustrates the distribution of healthcare facilities in Gauteng wards.

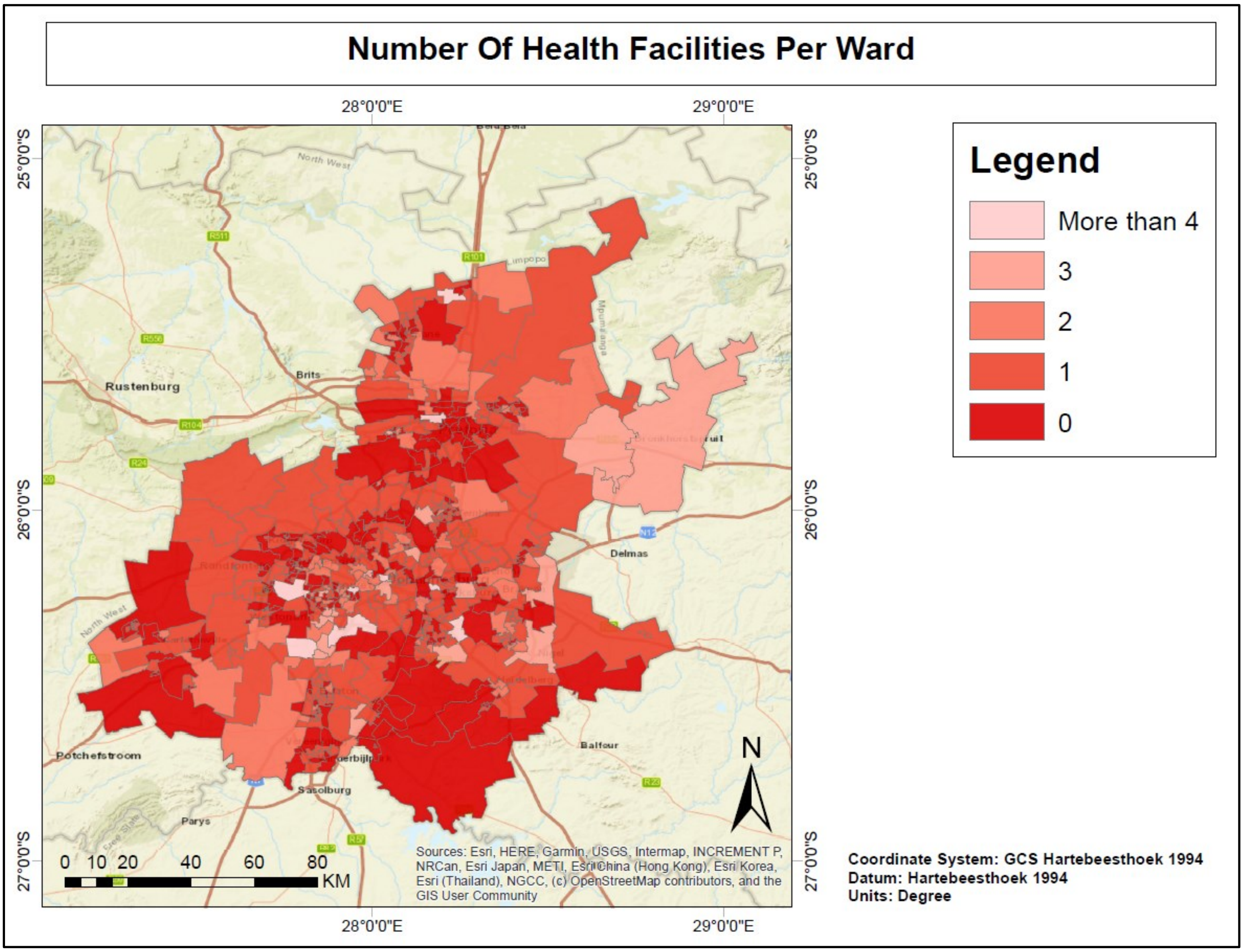

Figure 6: Distribution of healthcare facilities in Gauteng wards

5. Population density

The rationale behind population density being a vulnerability indicator for COVID-19 is based on the idea that when people live close to each other, the potential to spread the disease is heightened (Alam, 2021). With this logic, areas with high population densities would be regarded as more vulnerable to COVID-19. Figure 7 illustrates the distribution of population density in Gauteng wards.

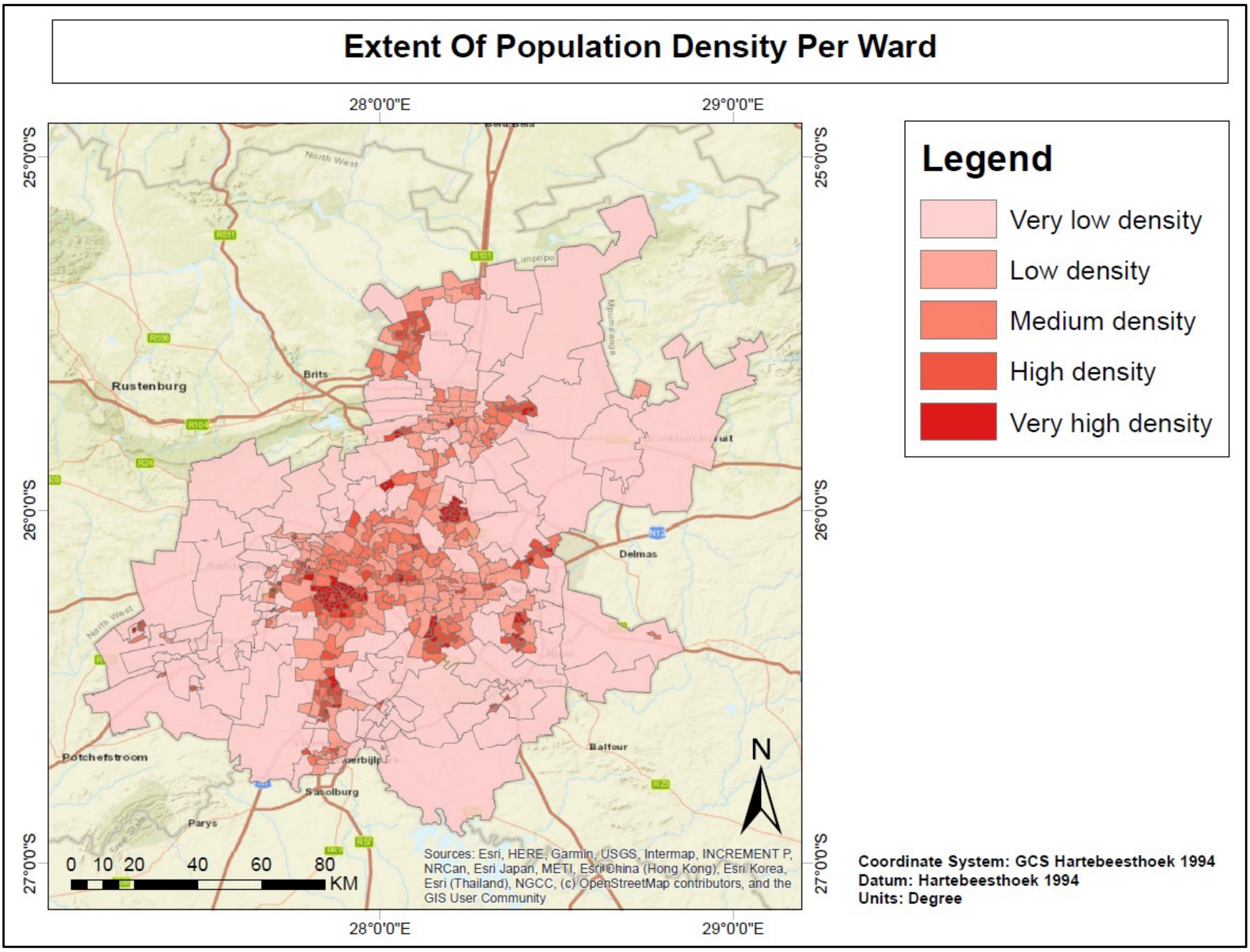

Figure 7: Distribution of population density in Gauteng wards

6. Number of people with pre-existing health conditions

It is indicated that people who suffer from certain pre-existing health conditions such as cancer, diabetes, heart conditions, HIV, Asthma etc. are more prone to severe COVID-19 illness (CDC, 2021). With this logic, wards with a higher proportion of individuals living with pre-existing health conditions can be regarded as more vulnerable to COVID-19. Figure 8 illustrates the distribution of the proportion of people in Gauteng wards living with pre-existing health conditions.

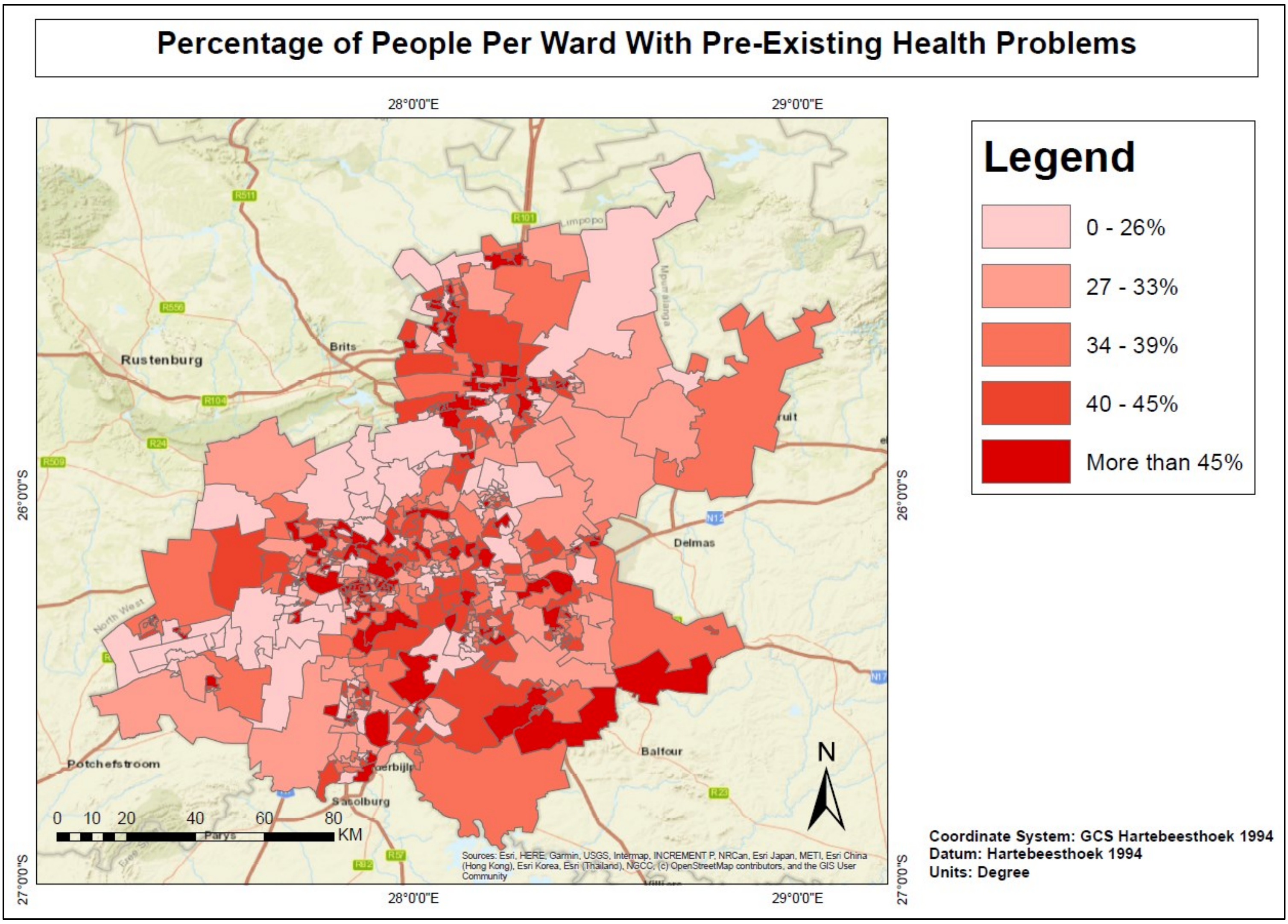

Figure 8: Distribution of the proportion of Gauteng residents living with pre-existing health problems

3.4.2. Normalization

After all criteria indicators were defined, each indicator was spatially referenced and outputted onto a GIS environment as a layer. Criteria layers were then categorized as per requirements of the AHP method into beneficial (criteria whose higher value is desired i.e. the number of health facilities in a ward) and non-beneficial (criteria whose lower value is desired i.e. education deprivation).

Following the categorization of all criteria layers, AHP based normalization was performed using equations (1) and (2). For the EWM, normalization of the criteria indicators was not preceded by categorization, it was simply performed on each indicator using equation (3). All normalization was performed on each spatially referenced criteria indicator layers attribute table via the field calculator on ArcGIS 10.6.

The three equations used for normalizing the criteria layers are illustrated below. The first one was used for the beneficial criteria and the second one was for the non-beneficial criteria (as per AHP requirements).

$$n_{ij} = \frac{r_{ij}}{r_{max}} \qquad (1)$$

$$n_{ij} = 1 - \frac{r_{ij}}{r_{max}} \qquad (2)$$

Where, r_{ij} is the performance value in each individual cell and r_{max} is the maximum value in the criteria group (Nymbili & Erden, 2020).

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}}, \quad (i = 1,2,\ldots,m; j = 1,2,\ldots,n) \qquad (3)$$

Where, x_{ij} is the performance value in each individual cell and $\sum_{i=1}^{m} x_{ij}$ is the sum of the performance scores in the criteria group (Nymbili & Erden, 2020).

3.4.3. Assigning criteria indicator weights and mapping

The three algorithms used for assigning weights to the criteria indicators were the AHP, the EWM and the integrated (AHP-EWM) approach. Once all criteria weights were derived using their respective algorithms, an AHP, EWM and integrated COVID-19 vulnerability model was developed on a GIS platform,

1. AHP

The AHP is an MCDA algorithm that uses stratified structures to analyse decision-making criteria through a pairwise comparison matrix that assigns criteria weights (Nymbili & Erden, 2020). Since the AHP method requires subjective opinions of experts to derive weights during the pairwise comparison stage of the process (Zhu, Tian, & Yan, 2020), it was important for study completion that experts get recruited as participants to help derive criteria weights.

In this study, the views of three (3) experts were consulted for constructing the pairwise comparison matrix and subsequent weight allocation for the criteria indicators. A questionnaire was designed (see annexure A) for experts to evaluate the relative importance between sets of criteria identified as important in determining COVID-19 vulnerability. The scale of relative importance proposed by Saaty (1980) and shown in table 2, was used to construct the pairwise comparison matrix as shown in table 3. The ratings involved in the scale of relative importance adhere to a nine (9) point continuous scale ranging from one (1) to nine (9) and their reciprocals

(Mishra, Gayen, & Haque, 2020; Sarkar, 2020). Values greater than one (1) represent more importance, those that are less than one (1) i.e. the reciprocals, represent less importance and those equal to one (1) represent equal importance (Saaty, 1980; Sarkar, 2020).

Table 2: Scale of relative importance

CHOICE	IMPORTANCE VALUE
Equally Important	1
Moderately Important	3
Strong Importance	5
Very Strongly Important	7
Extreme Importance	9
Values In Between Importance	2,4,6,8

Table 3: Pairwise matrix used for the study

Pairwise Comparison Matrix						
	Income Poverty	Education	Living environment deprivation	Number of healthcare centres	Population density	People with pre-existing conditions
Income Poverty	1					
Education		1				
Living environment deprivation			1			
Number of healthcare centres				1		
Population density					1	
People with pre-existing conditions						1

Once the pairwise matrix was complete, the results were checked for consistency using the consistency ratio recommended by Saaty (1980), see equation 4. The value of the consistency ratio has to be less than or equal to 0.1 for the consistency of the judgements to be deemed acceptable (Saaty, 1980; Jam, Mosaffaie, Sarfaraz, Shadfar, & Akhtari, 2021). If the value of the consistency ratio is greater than 0.1, then consistency re-analysis will have to be conducted (Saaty, 1980; Jam, Mosaffaie, Sarfaraz, Shadfar, & Akhtari, 2021).

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} \qquad \textit{(4)}$$

Where CI the Consistency Index is calculated using equation 5 and RI is the Random Consistency Index obtained using table 4 and determined by the number of criteria indicators used in the study, which was six (6).

$$\mathrm{CI} = \frac{\lambda_{max} - n}{n-1} \qquad (5)$$

Where n is the number of criteria indicators and λ_{max} is the maximum eigenvalue (Saaty, 1980; Jam, Mosaffaie, Sarfaraz, Shadfar, & Akhtari, 2021).

Table 4: Random Index table

n	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
RI	0.00	0.00	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49

Before the criteria weights derived from the pairwise matrix were used, the normalized AHP criteria layers were first overlaid with one another using the overlay analysis tool in ArcGIS version 10.6. This resulted in creating a single polygon output that contained all six (6) criteria indicator parameters in its attribute table.

Once overlaid, it was only then that the derived criteria weights were the basis for calculating the vulnerability score for each Gauteng ward. The calculation of the vulnerability score for each ward was performed by multiplying the normalized performance scores of each ward by their associated criteria indicator weight. Once this was achieved, the total of these values was added together to compute the vulnerability score.

Vulnerability score per ward $= (np_{ij} * W_{(\text{Income poverty})}) + (np_{ij} * W_{(\text{Education deprivation})}) + (np_{ij} * W_{(\text{Living environment deprivation})}) + (np_{ij} * W_{(\text{Number of healthcare centres})}) + (np_{ij} * W_{(\text{Population density})}) + (np_{ij} * W_{(\text{Pre-existing health conditions})})$

Where np_{ij} is the normalized performance score per criteria indicator and W is the associated criteria indicator weight.

With the vulnerability scores calculated, the basis of an AHP COVID-19 vulnerability model for Gauteng was obtained and to visualize the AHP vulnerability model developed, GIS was used to classify the vulnerability values into five classes through the quantile classification method.

2. EWM

The Entropy Weight Method is described as a concept that measures the extent of variation in a dataset using probability theory (Nymbili & Erden, 2020). Whereas the AHP approach uses a subjective measure for assigning weights, the EWM uses a strictly objective assessment for weight allocation (Zhu, Tian, & Yan, 2020).

In this method, no additional data was needed to compute the criteria weights. The criteria weights were simply derived from the normalized EWM criteria layers and calculated using equations (6) and (7) below.

Computing the entropy value using the following equation;

$$E_j = -k \sum_{i=1}^{m} p_{ij} \ln p_{ij} \qquad (i = 1,2,\dots,m; j = 1,2,\dots,n) \tag{6}$$

Where, $-k$ is the entropy constant which is equal to k = 1/ ln (m), $\sum_{i=1}^{m}$ is the sum of the performance scores and p_{ij} is the individual performance scores (Lotfi & Fallahnejad, 2010).

Computing the entropy weights using the following equation;

$$w_j = \frac{1-E_j}{\sum_{j=1}^{n}(1-E_j)} \qquad j = 1,2,\dots,n \tag{7}$$

Where, E_j is the entropy value, $1 - E_j$ is Diversification value and $\sum_{j=1}^{n} (1 - E_j)$is the sum of Diversification (Lotfi & Fallahnejad, 2010).

Before the derived criteria weights were used, the normalized EWM criteria layers were first overlaid with one another using the overlay analysis tool in ArcGIS version 10.6. This resulted in creating a single polygon output that contained all six (6) criteria indicator parameters in its attribute table.

Once overlaid, it was only then that the derived criteria weights were the basis for calculating the vulnerability score for each Gauteng ward. The calculation of the vulnerability score for each ward was performed by multiplying the normalized performance scores of each ward by their associated criteria indicator weight. Once this was achieved, the total of these values was added together to compute the vulnerability score.

***Vulnerability score per ward* = (np$_{ij}$ * *W*$_{(Income\ poverty)}$) + (np$_{ij}$ * *W*$_{(Education\ deprivation)}$) + (np$_{ij}$ * *W*$_{(Living\ environment\ deprivation)}$) + (np$_{ij}$ * *W*$_{(Number\ of\ healthcare\ centres)}$) + (np$_{ij}$ * *W*$_{(Population\ density)}$) + (np$_{ij}$ * *W*$_{(Pre\text{-}existing\ health\ conditions)}$)**

Where np_{ij} is the normalized performance score per criteria indicator and W is the associated criteria indicator weight.

With the vulnerability scores calculated, the basis of an EWM COVID-19 vulnerability model for Gauteng was obtained and to visualize the EWM vulnerability model developed, GIS was used to classify the vulnerability scores into five classes through the quantile classification method.

3. Integrated model

The weights derived from the AHP and EWM approaches were combined and used as the basis for plotting an integrated AHP-EWM COVID-19 vulnerability map for Gauteng. Equation (8) is the recommended formulae for integrating AHP and EWM weights.

$$w_j^* = \frac{s_j w_j}{\sum_{j=1}^{n} s_j w_j} \qquad (8)$$
$$j = 1,2,\dots,n$$

Where, S_j is the subjective weight calculated from the AHP, and w_j is the objective weight derived from the entropy method (Nymbili & Erden, 2020).

Before the derived criteria weights were used, duplicates of the normalized EWM criteria layers were first overlaid with one another using the overlay analysis tool in ArcGIS version 10.6. This resulted in creating a single polygon output that contained all six (6) criteria indicator parameters in its attribute table.

Once overlaid, it was only then that the derived criteria weights were the basis for calculating the vulnerability score for each Gauteng ward. The calculation of the vulnerability score for each ward was performed by multiplying the normalized performance scores of each ward by their associated criteria indicator weight. Once this was achieved, the total of these values was added together to compute the vulnerability score.

***Vulnerability score per ward* = (np$_{ij}$ * *W*$_{(Income\ poverty)}$) + (np$_{ij}$ * *W*$_{(Education\ deprivation)}$) + (np$_{ij}$ * *W*$_{(Living\ environment\ deprivation)}$) + (np$_{ij}$ * *W*$_{(Number\ of\ healthcare\ centres)}$) + (np$_{ij}$ * *W*$_{(Population\ density)}$) + (np$_{ij}$ * *W*$_{(Pre\text{-}existing\ health\ conditions)}$)**

Where np_{ij} is the normalized performance score per criteria indicator and W is the associated criteria indicator weight.

With the vulnerability scores calculated, the basis of an Integrated COVID-19 vulnerability model for Gauteng was obtained and to visualize the integrated vulnerability model developed, GIS was used to classify the vulnerability scores into five classes through the quantile classification method.

3.4.4. Validation

The steps taken towards validating the models first involved randomly selecting 158 wards (30%) to be used as testing variables from the cumulative COVID-19 cases dataset which was originally in excel format. The random selection of the sample validation dataset was conducted via Excel 2013 before the cumulative COVID-19 case dataset was spatially referenced onto a GIS platform. Once on a GIS, the cumulative COVID-19 case data was converted from a polygon to a point dataset with each ward and its associated COVID-19 cases information represented as a point on the centre of the ward.

The number of COVID-19 cases represented by each point was then classified into two (2) groups. The first group identified points (wards) with less than or equal to 1482 cumulative COVID-19 cases which were labelled as validation points with low COVID-19 cases. While the second group identified points (wards) with more than or equal to 1483 cumulative COVID-19 cases, labelled as validation points with high COVID-19 cases.

The cumulative COVID-19 cases point data was then joined onto each COVID-19 vulnerability model layer using the "Ward ID" as the unique identifier. This process was then followed by the conversion of the associated attribute tables into excel documents to be used for computing the ROC-AUC in SPSS.

The ROC-AUC was the validation technique used in this study to validate the three (3) COVID-19 vulnerability models generated. ROC-AUC values are widely used validation techniques for verifying the overall adequacy of generated models (Pourghasemi, et al., 2020). ROC-AUC operate by juxtaposing a validation dataset against the model output generated (Pourghasemi, et al., 2020; Mishra, Gayen, & Haque, 2020).

A model is regarded as perfect (1.0-0.9), very-good (0.9-0.8), good (0.8-0.7), moderate (0.7-0.6) and poor (0.6-0.5) with these respective ROC-AUC values (Pourghasemi, et al., 2020).

The ROC-AUC for this study was calculated for each COVID-19 vulnerability model by using their vulnerability scores as the ‘testing variable’ and the cumulative COVID-19 case data as the ‘state variable’ in SPSS.

4. RESULTS

The results are broken down into four (4) sections and organised by the model output developed. Detailed results were revealed for the steps taken toward the generation of the AHP, EWM and Integrated COVID-19 vulnerability models. The final section reveals the results of the validation process.

4.1. AHP COVID-19 vulnerability model

The six spatially referenced criteria indicators were categorized into beneficial and non-beneficial criteria as it is a pre-requisite step to be taken before normalizing them. From the list of indicators, the only beneficial criteria and its respective $r_{\max}$ value was the number of healthcare facilities in the ward (8). The remaining criteria indicators were categorized as non-beneficial criteria with their respective $r_{\max}$ values as follows;

- Pre-existing health conditions (77.9)
- Population density (95527.31)
- Living environment deprivation (44361)
- Income poverty (51996)
- Education deprivation (11283)

Once all criteria indicators were successfully normalized, criteria weighting through pairwise comparison followed. Three experts (a social policy and public health expert, a research and policymaking expert, and a deprivation and inequality expert) were consulted for completing the pairwise comparison matrix which was used to determine the criteria weights assigned to each indicator. The pairwise comparison matrix output for each expert is illustrated in tables 5, 6 and 7.

Table 5: Pairwise matrix of social policy and public health expert

Pairwise Comparison Matrix						
	Income Poverty	Education	Living environment deprivation	Number of healthcare centres	Population density	People with pre-existing conditions
Income Poverty	**1**	8	1	3	2	1
Education	1/8	**1**	1/8	1/3	1/3	1/7
Living environment deprivation	1	8	**1**	3	2	1/2
Number of healthcare centres	1/3	3	1/3	**1**	2	1/2
Population density	1/2	3	1/2	1/2	**1**	1/3
People with pre-existing conditions	1	7	2	2	3	**1**

Table 6: Pairwise matrix of research and policy making expert

Pairwise Comparison Matrix						
	Income Poverty	Education	Living environment deprivation	Number of healthcare centres	Population density	People with pre-existing conditions
Income Poverty	**1**	6	1	4	3	1/6
Education	1/6	**1**	1/3	1	1/2	1/6
Living environment deprivation	1	3	**1**	4	3	1/6
Number of healthcare centres	1/4	1	1/4	**1**	1/4	1/6
Population density	1/3	2	1/3	4	**1**	1/3
People with pre-existing conditions	6	6	6	6	3	**1**

Table 7: Pairwise matrix of deprivation and inequality expert

Pairwise Comparison Matrix						
	Income Poverty	Education	Living environment deprivation	Number of healthcare centres	Population density	People with pre-existing conditions
Income Poverty	**1**	2	1/2	5	4	1/3
Education	1/2	**1**	1/2	3	2	1/4
Living environment deprivation	2	2	**1**	5	5	1/3
Number of healthcare centres	1/5	1/3	1/5	**1**	1/3	1/5
Population density	1/4	1/2	1/5	3	**1**	1/3
People with pre-existing conditions	3	4	3	5	3	**1**

The judgements of all three (3) experts from their respective pairwise matrices were aggregated by their mean to derive the final (grouped) pairwise matrix as shown in table 8. The grouped pairwise matrix was the matrix used to derive the final weights of the AHP model.

Table 8: Grouped pairwise matrix

Pairwise Comparison Matrix						
	Income Poverty	Education	Living environment deprivation	Number of healthcare centres	Population density	People with pre-existing conditions
Income Poverty	**1**	5.33	0.83	4	3	0.5
Education	0.19	**1**	0.32	1.44	0.94	0.19
Living environment deprivation	1.20	3.13	**1**	2.67	3.33	0.33
Number of healthcare centres	0.25	0.70	0.37	**1**	0.86	0.29
Population density	0.33	1.06	0.30	1.16	**1**	0.33
People with pre-existing conditions	2	5.26	3.03	3.45	3.03	**1**

Based on the grouped pairwise matrix, the criteria weights derived and the ranks for each indicator are illustrated in table 9.

Table 9: AHP weights and rankings

Criteria	Weight (%)	Rank
Income Poverty	23	2
Education Deprivation	7	5
Living Environment Deprivation	20	3
Number of Healthcare Facilities	7	5
Population Density	8	4
Pre-Existing Health Conditions	35	1

To check the reliability of the derived criteria weights, a λ_{max} value of 6.2183, a consistency index (CI) value of 0.04366 and a consistency ratio (CR) value of 0.035 were computed based on the grouped matrix. Given that the CR value is less than 0.1 ($0.035 < 0.1$), it was determined that the derived grouped matrix and its associated criteria indicator weights were reliable.

With a reliable set of subjective criteria indicator weights derived, the vulnerability scores for each ward were calculated using the vulnerability score equation via the field calculator on ArcGIS 10.6.

Vulnerability score per ward = (np_{ij} * 0.23) + (np_{ij} * 0.07) + (np_{ij} * 0.20) + (np_{ij} * 0.07) + (np_{ij} * 0.08) + (np_{ij} * 0.35)

The vulnerability scores were then classified into five (5) classes using the quantile classification method, where the least vulnerable ward was classified as level 1 and the most vulnerable ward to COVID-19 was classified as level 5. The output derived from this exercise was an AHP COVID-19 vulnerability map shown in figure 9.

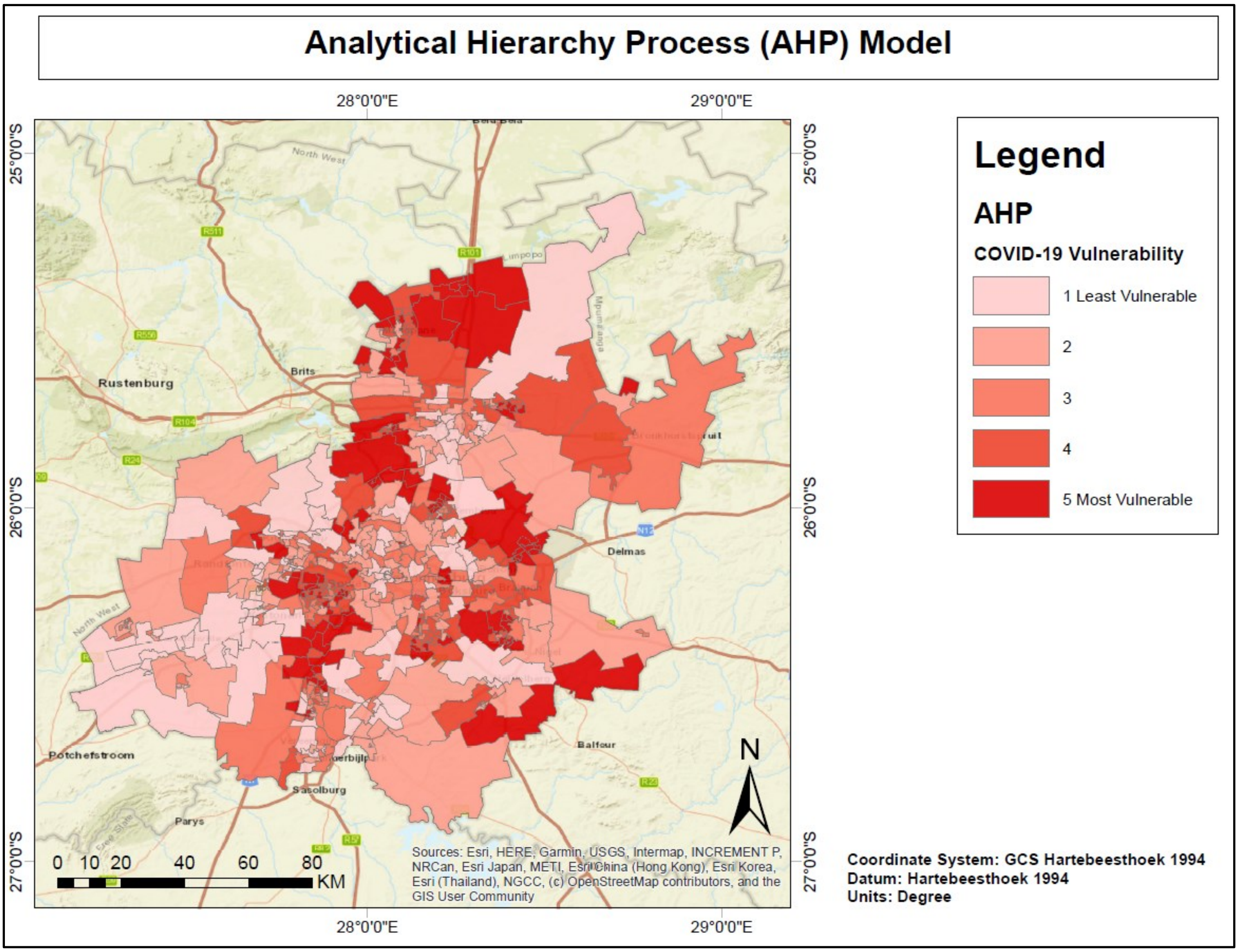

Figure 9: AHP COVID-19 vulnerability map for Gauteng wards

4.2. EWM COVID-19 vulnerability model

The normalization process of the spatially referenced criteria indicator layers was conducted using equation three (3) as mentioned in section 3.4.2 of the methodology chapter, via the field calculator in ArcGIS 10.6.

Upon normalization, equations six (6) and seven (7) as mentioned in section 3.4.3 of the methodology chapter, were used to compute the entropy value (E_j), Diversification (1 - E_j) and the objective weights (w_j) for each criteria indicator. Each of these values is illustrated in table 10.

Table 10: EWM weights and rankings

Criteria	E_j	(1 - E_j)	Weight (%)	Rank
Income Poverty	0,991858	0,008142	12	3
Education Deprivation	0,908893	0,091107	11	4
Living Environment Deprivation	0,896664	0,103336	26	1
Number of Healthcare Facilities	0,951396	0,048604	26	1
Population Density	0,957318	0,042682	23	2
Pre-Existing Health Condition	0,89797	0,10203	2	5

With a set of objective criteria indicator weights derived, the vulnerability scores for each ward were calculated using the vulnerability score equation via the field calculator on ArcGIS 10.6.

Vulnerability score per ward = (np_{ij} * 0.12) + (np_{ij} * 0.11) + (np_{ij} * 0.26) + (np_{ij} * 0.26) + (np_{ij} * 0.23) + (np_{ij} * 0.02)

The vulnerability scores were then classified into five (5) classes using the quantile classification method, where the least vulnerable ward was classified as level 1 and the most vulnerable ward to COVID-19 was classified as level 5. The output derived from this exercise was an EWM COVID-19 vulnerability map shown in figure 10.

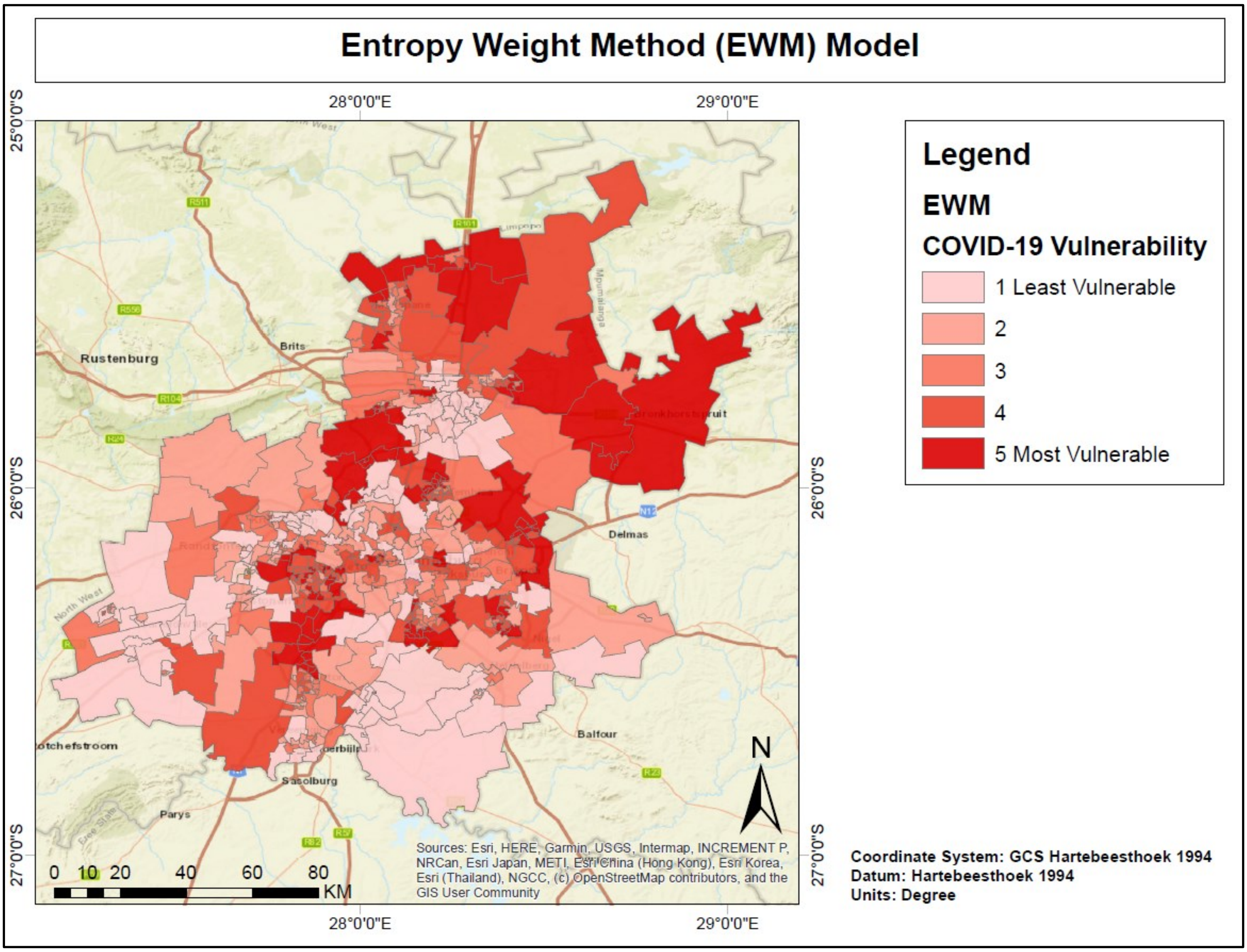

Figure 10: EWM COVID-19 vulnerability map for Gauteng Wards

4.3. Integrated COVID-19 vulnerability model

In this study, integrating the AHP and EWM weights was performed to derive the weights needed to develop an integrated COVID-19 vulnerability map. Equation eight (8) as noted in section 3.4.3 of this study was used to derive the integrated criteria indicator weights and the results for this are shown in table 11.

Table 11: Integrated weights and rankings

Criteria	AHP weight (%)	EWM weight (%)	Integrated weight (%)	Ranking
Income Poverty	23	12	21	2
Education Deprivation	7	11	6	4
Living Environment Deprivation	20	26	40	1
Number of Healthcare Facilities	7	26	14	3
Population Density	8	23	14	3
Pre-Existing Health Conditions	35	2	5	5

With a set of subjective-objective (integrated) criteria indicator weights derived, the vulnerability scores for each ward were calculated using the vulnerability score equation via the field calculator on ArcGIS.

Vulnerability score per ward $= (np_{ij} * 0.21) + (np_{ij} * 0.06) + (np_{ij} * 0.40) + (np_{ij} * 0.14) + (np_{ij} * 0.14) + (np_{ij} * 0.05)$

The vulnerability scores were then classified into five (5) classes using the quantile classification method, where the least vulnerable ward was classified as level 1 and the most vulnerable ward to COVID-19 was classified as level 5. The output derived from this exercise was an Integrated COVID-19 vulnerability map shown in figure 11.

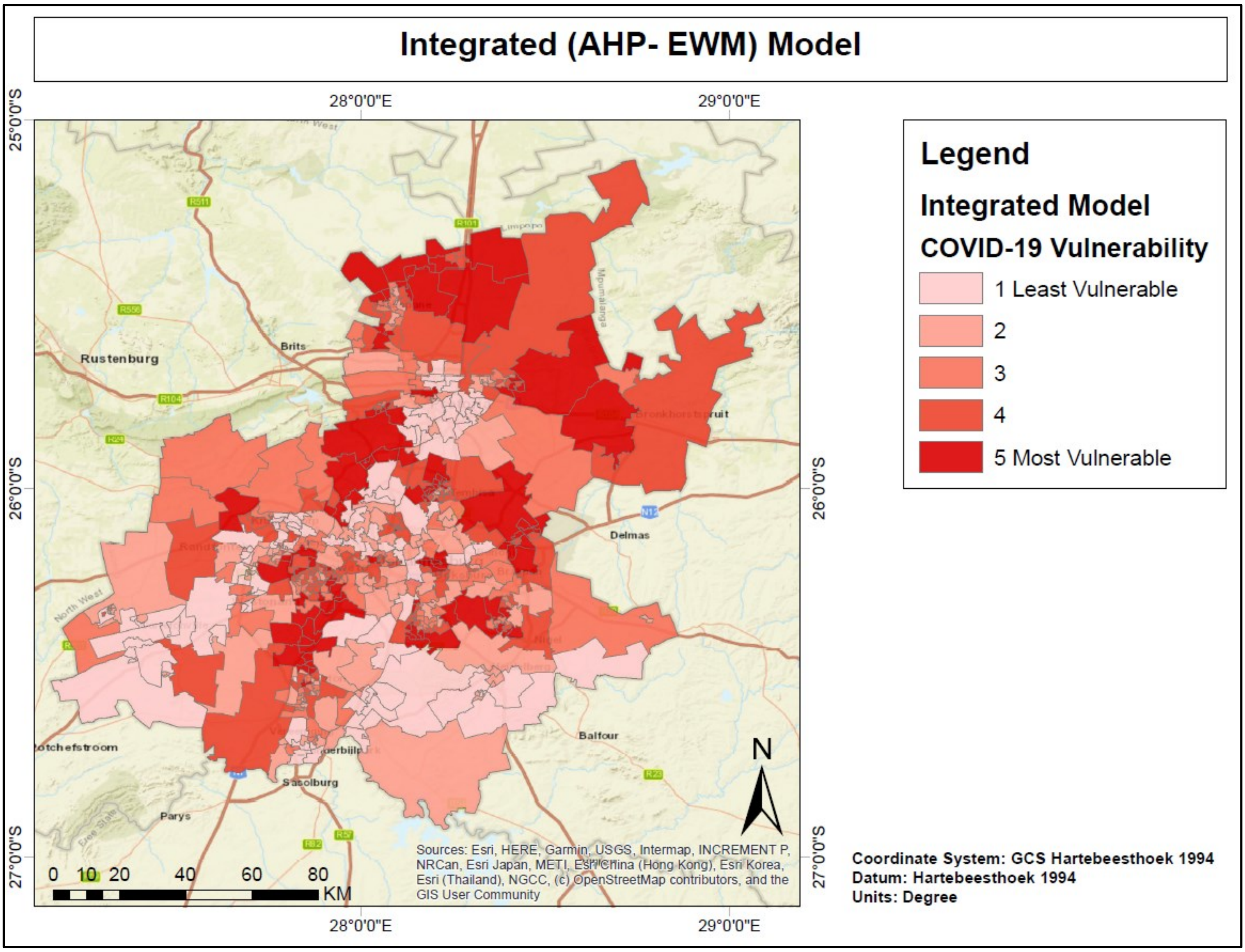

Figure 11: Integrated COVID-19 vulnerability map for Gauteng wards

4.4. Model validation

Model validation was conducted using the ROC-AUC where the validation dataset (randomly selected cumulative COVID-19 case data) was juxtaposed by the model output (the three COVID-19 vulnerability maps). While SPSS was used to verify the adequacy of the generated models by calculating their ROC-AUC values, GIS was used to visualize the validation dataset with the model outputs.

Figures 12, 13 and 14 illustrate the GIS output derived from overlaying the validation dataset over the model output. The visual fidelity of the GIS output was enhanced by converting the derived model outputs into continuous surfaces before smoothing them using a filter.

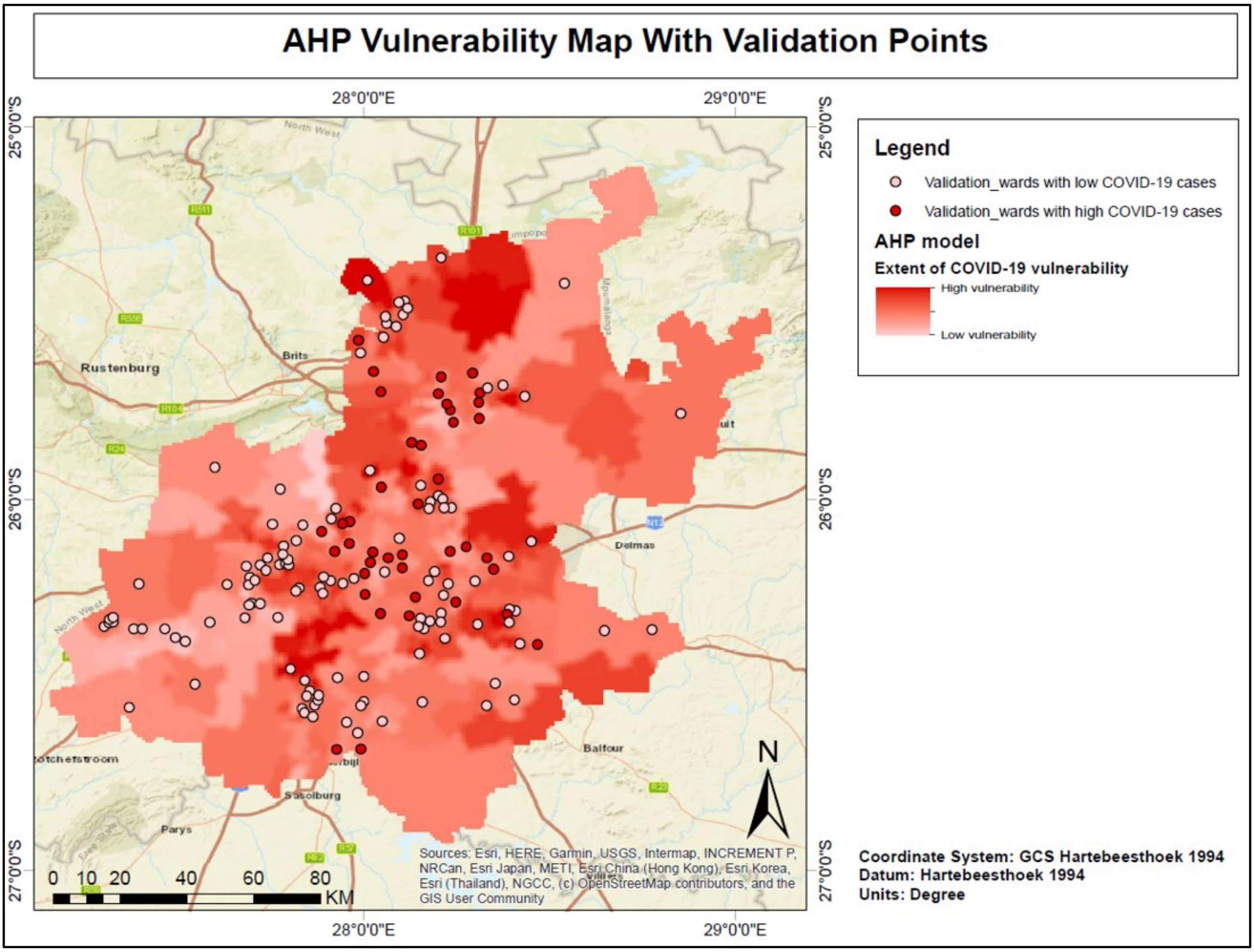

Figure 12: Validation plot for the AHP COVID-19 vulnerability map

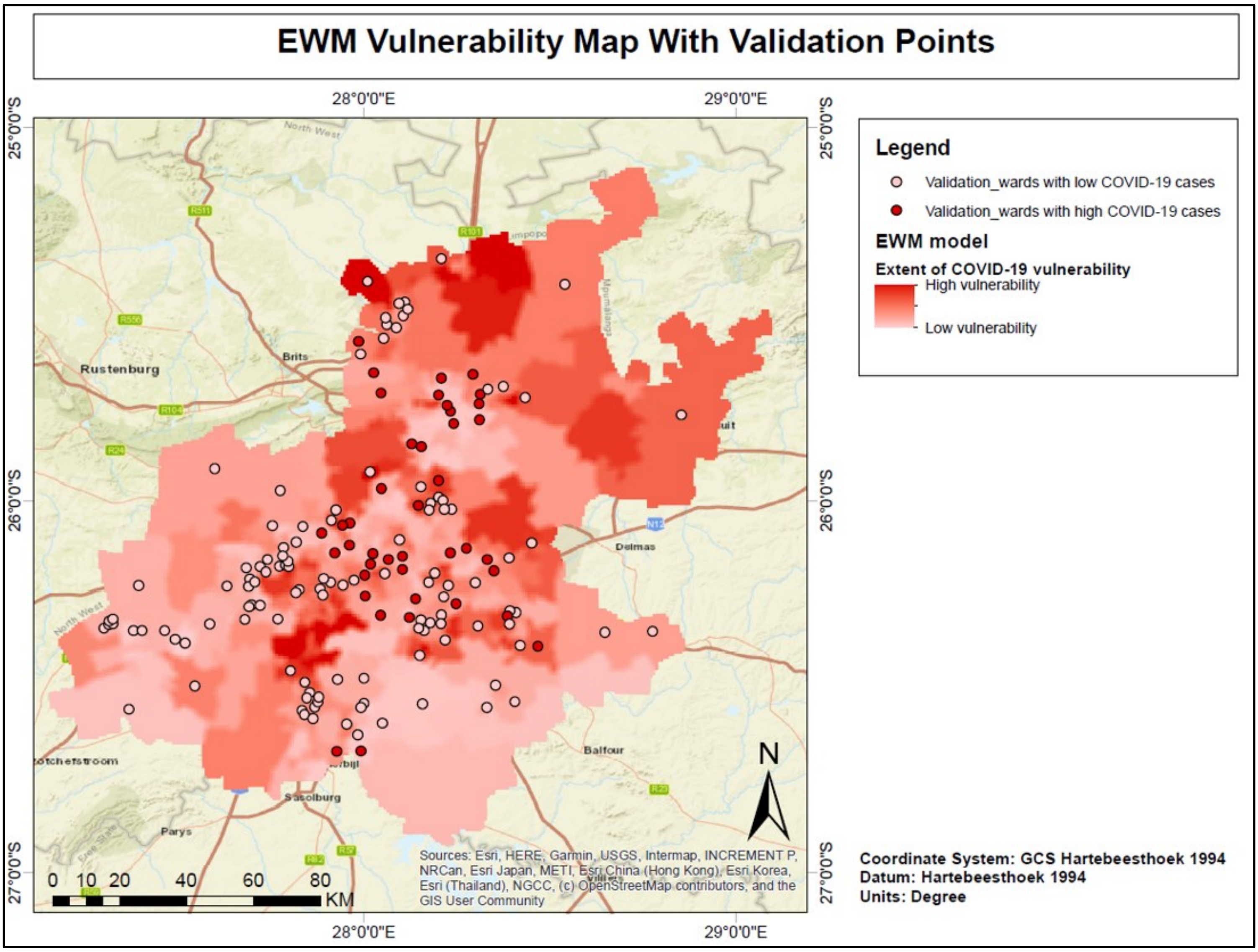

Figure 13: Validation plot for the EWM COVID-19 vulnerability map

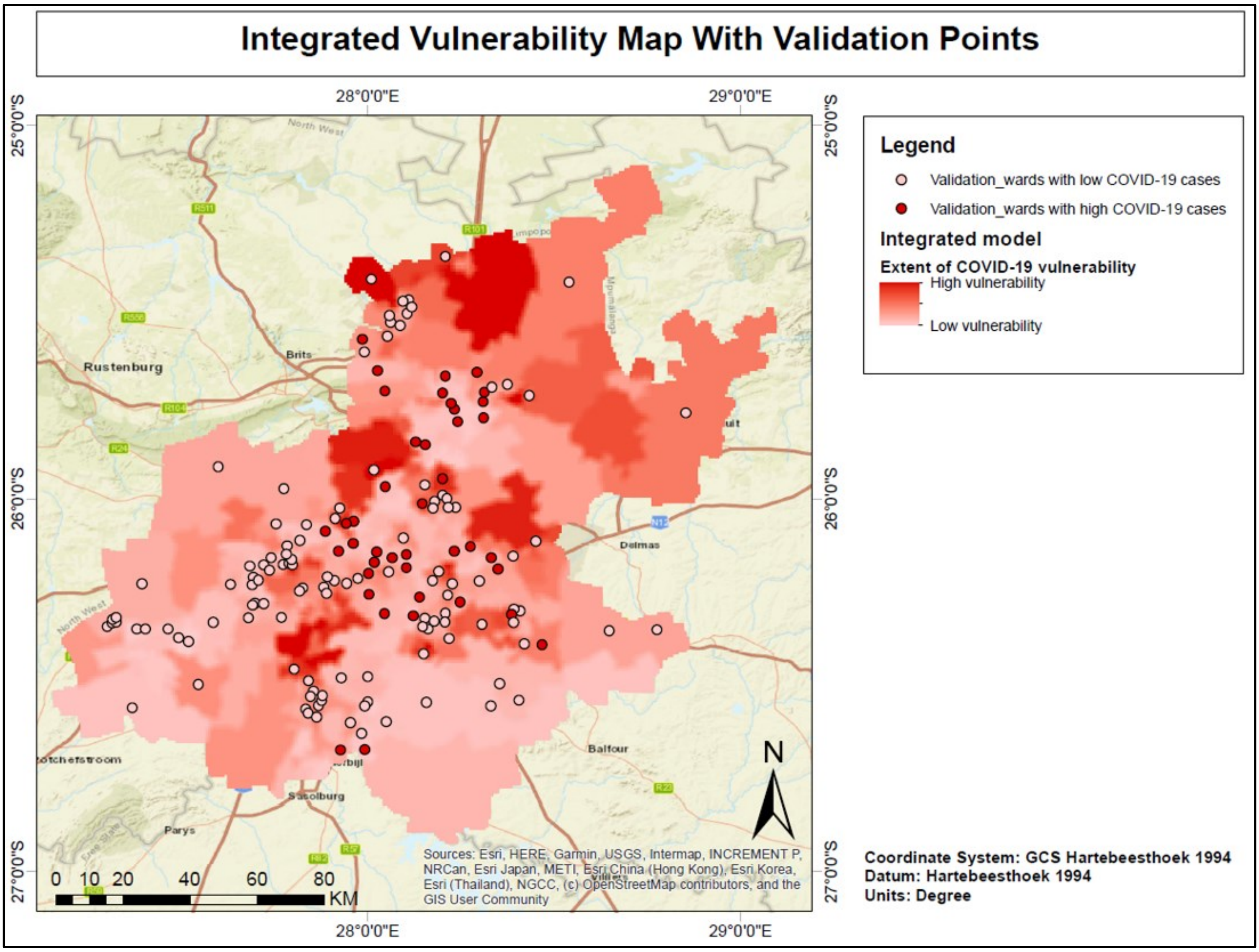

Figure 14: Validation plot for the integrated COVID-19 vulnerability map

The ROC-AUC validation output derived using SPSS is illustrated in figures 15, 16 and 17. The output from the results indicates that the AHP and EWM models had AUC scores of 0.634 and 0.686 respectively, which are both regarded as results that are moderately accurate for a model because they fall within the range of 0.6 and 0.7 (Pourghasemi, et al., 2020). On the other hand, the integrated models results indicate that the model had an AUC score of 0.703, which is a result that is regarded as good for a model because it falls within the range of 0.7 and 0.8 (Pourghasemi, et al., 2020).

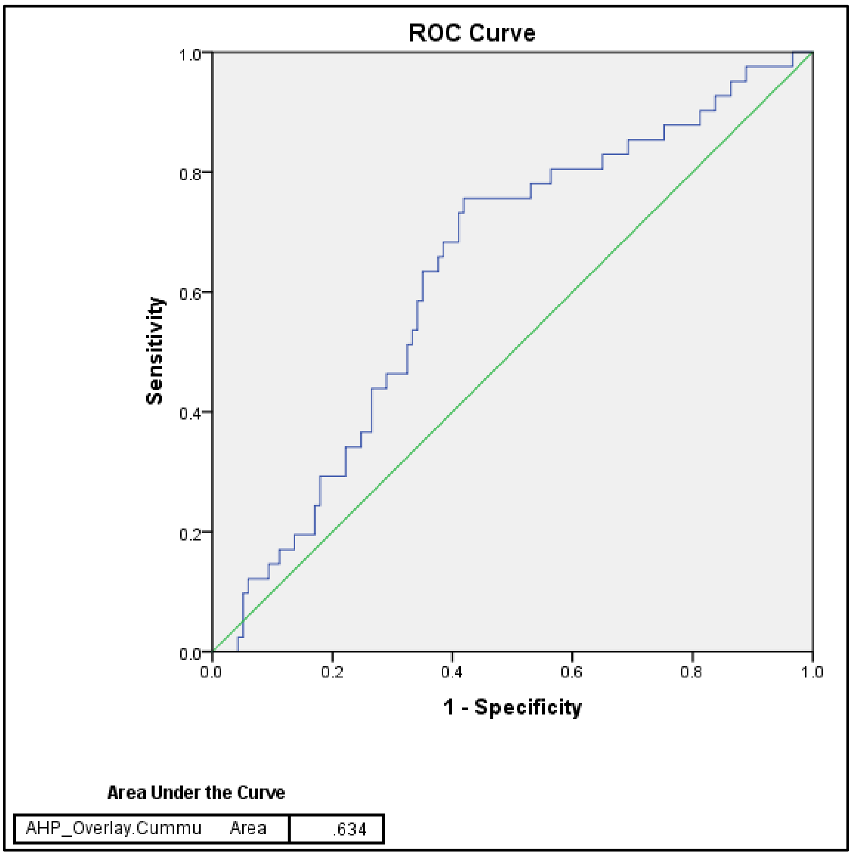

Figure 15: ROC-AUC results for AHP model

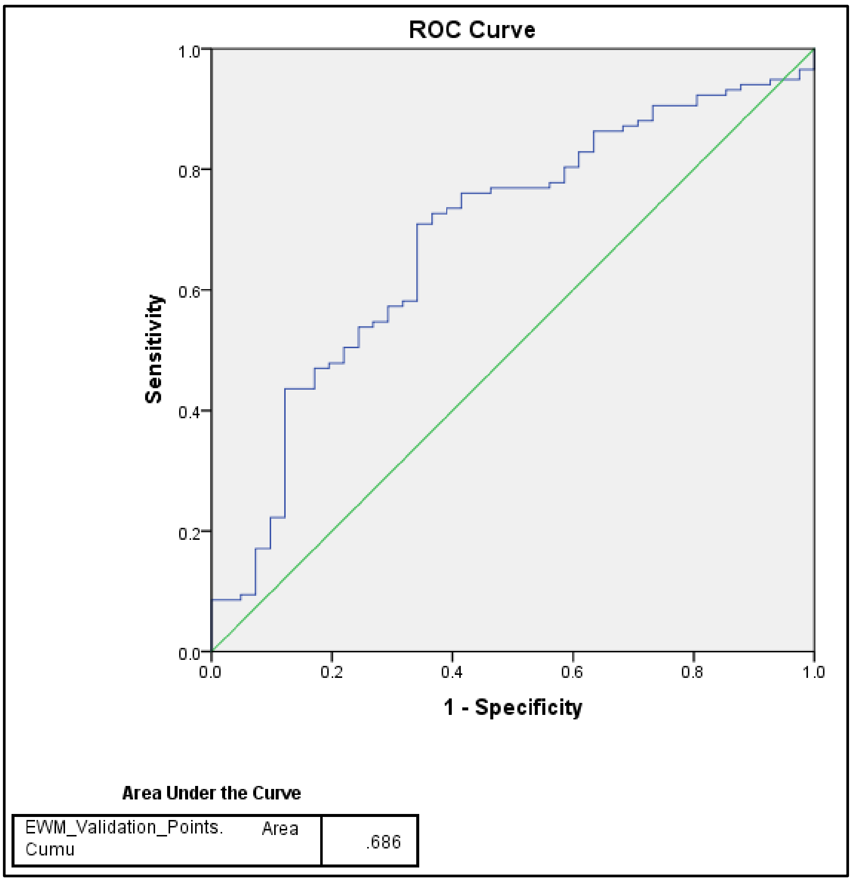

Figure 16: ROC-AUC results for EWM model

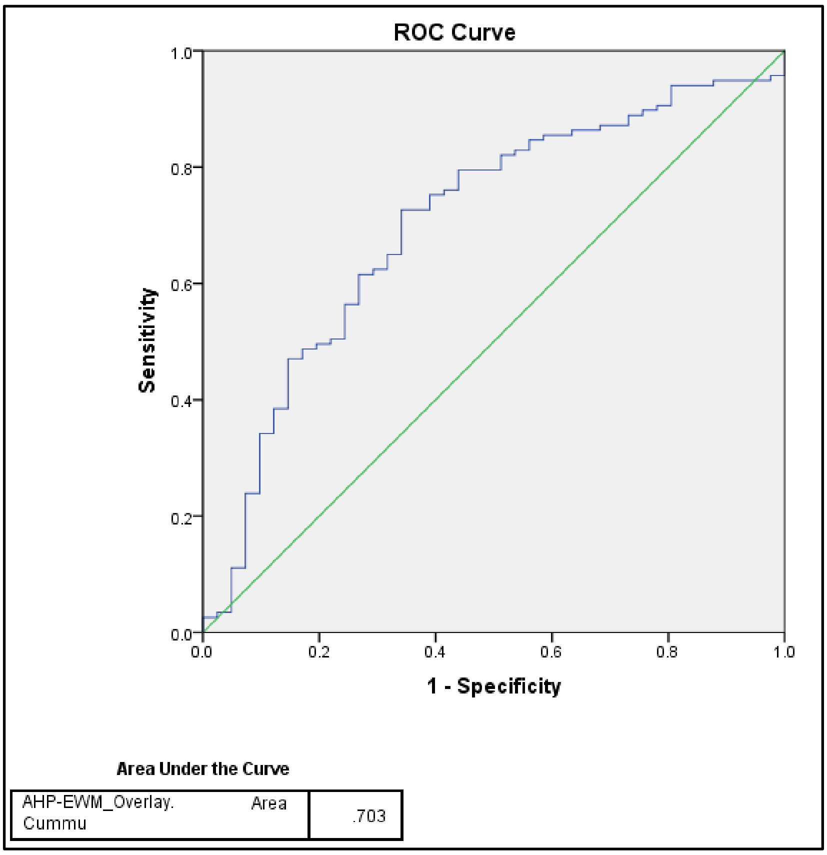

Figure 17: ROC-AUC results for integrated model

5. DISCUSSION

At the heart of this research study and the reason the specific methodological approach was used, were three fundamental research questions that propelled scientific inquiry. 1) Will the criteria weights used in the MCDA matrices for both the AHP and EWM models be given similar priority? 2) Which vulnerability map (AHP or EWM) is more accurate for mapping COVID-19 vulnerability in Gauteng? 3) Is the integrated approach to COVID-19 vulnerability mapping more accurate than the individual AHP and EWM models?

In this section of the report, responses to each of these research questions will be given based on study results, and a discussion unpacking the implications of these responses will be had.

5.1. Will the criteria weights used in the MCDA matrices for both the AHP and EWM models be given similar priority?

The study results revealed that the subjective (AHP) and objective (EWM) approach to criteria weight determination allocated different priorities to the criteria indicator weights. The subjective AHP weight assigning algorithm gave the highest priority (ranking of 1) to pre-existing health conditions with an indicator weight of 35%. It then gave the least priority (ranking of 5) to two criteria indicators which were education deprivation and the number of healthcare facilities. In both cases, these criteria indicators were given a criteria weight of 7%. Meanwhile, the objective EWM weight assigning approach gave the highest priority (ranking 1) to both living environment deprivation and the number of healthcare facilities with a criteria indicator weight of 26% each. It then in contrast to the subjective weight allocation approach, gave the least priority (ranking of 5) to pre-existing health conditions with a criteria indicator weight of 2%.

The differences observed in how criteria indicator weights are prioritised between the subjective (AHP) and objective (EWM) algorithms, justifies the decision to use an integrated algorithm in the study. The integrated algorithm helps find the middle ground between the criteria indicator weights allocated by these two polar opposite algorithms.

In integrating the criteria indicator weights of the two algorithms, the study results revealed that the criteria indicator weights gave the highest priority (ranking of 1) to living environment deprivation with a weight of 40% and least priority (ranking of 5) to pre-existing health conditions with a weight of 5%.

5.2. Which vulnerability map (AHP or EWM) is more accurate for mapping COVID-19 vulnerability in Gauteng?

In comparing the validation results via ROC-AUC for the AHP and EWM COVID-19 vulnerability models. The results revealed that although both models can be regarded as moderately accurate with AUC values between 0.6 and 0.7 (Pourghasemi, et al., 2020), the EWM model happened to be the more accurate one for modelling COVID-19 vulnerability with an AUC score of 0.686 compared to the AHP's AUC score of 0.634.

As stated in the literature review, the lack of known studies that have used the combination of the AHP and EWM as algorithms to model COVID-19 vulnerability makes it impossible to make direct model accuracy comparisons of this study with other COVID-19 vulnerability studies. However, if we based the comparison on studies that have used both algorithms to model landslide susceptibility for instance, then having a result that saw the EWM model being more accurate than the AHP model is unsurprising. This is the case because a similar result was acquired in the study by Panchal and Shrivastava (2021), where it was found that the EWM showed greater model accuracy with an AUC score of 0.883 than the AHP model with an AUC score of 0.732 for modelling landslide susceptibility in India (Panchal & Shrivastava, 2021).

While the two COVID-19 vulnerability models were regarded as moderately accurate with AUC scores in the range of 0.6 to 0.7 (Pourghasemi, et al., 2020). The overall accuracy of all derived models in this study can be improved by updated datasets. Given that some criteria indicators such as living environment deprivation, education deprivation and income poverty are all based on 2011 census information, compromises the accuracy of derived models. The discrepancy of having a validation dataset based on cumulative COVID-19 cases from March 2020 to December 2021 and juxtaposing it to model outputs derived based on outdated criteria indicator information, limits the accuracy of the models derived and therefore lowers the AUC scores associated with these models.

5.3. Is the integrated approach to COVID-19 vulnerability mapping more accurate than the individual AHP and EWM models?

The integrated model had an AUC score of 0.703 which is regarded as a good model as it is within the range of 0.7 and 0.8 (Pourghasemi, et al., 2020). This model result was revealed to be higher and therefore more accurate than the individual AHP and EWM models. The literature review indicated that the integration of two weight assigning algorithms is an approach that is used to develop hybrid models that are more accurate and overcome the shortcomings associated with the isolated use of weight assigning algorithms (Das & Pal, 2020). In this regard, the outcome derived from the study results indicated that the integrated model yielded a model accuracy higher than the individual models is consistent with the literature.

The results derived from this study demonstrate that the integrated model approach is the best approach to use for modelling COVID-19 vulnerability in Gauteng. The integrated model not only outperformed the other individual models, but in its essence, it harmonises the individual criticisms and shortcomings associated with the other models when used in isolation.

5.4. Study implications for policy making

This research study provides a basis for informing policy making with regards to understanding the nuances involved in defining COVID-19 vulnerability. The study reveals that COVID-19 vulnerability is a dynamic concept that needs to be understood not only in the epidemiological sense i.e. by considering those who are elderly, or those affected by certain comorbidities but to instead look at COVID-19 vulnerability more holistically by also considering those in socio-economic conditions that would make them less capable of dealing with the disease over the long term i.e. in the case of prolonged hard lockdown measures. In addition to providing a nuanced understanding for defining COVID-19 vulnerability, this research study also advocates for the use of integrated or hybrid models in the mapping of COVID-19 vulnerability. Integrated models as shown in this study, have proven to be the best suited for accurately modelling COVID-19 vulnerability and should therefore be the standard models of choice to use for guiding government intervention efforts. This would ensure that the limited resources allocated toward the mitigation of the disease are optimally utilised.

6. CONCLUSION

South Africa's fight against COVID-19 still rages on, with the country most recently being on the tail end of the fourth wave of COVID-19 infections that were propelled by the Omicron variant of the disease (Meterlerkamp, 2022). Gauteng remains the hardest hit province by COVID-19 with 32.5% of South Africa's total COVID-19 cases being reported within it. The extent of COVID-19 cases in Gauteng justifies the need for developing a vulnerability model that can help guide government planning and intervention efforts. The integrated COVID-19 vulnerability model derived in this study is the most appropriate model to use to help guide government planning in the fight against COVID-19, as it most accurately identifies Gauteng wards most vulnerable to the disease. The integrated COVID-19 vulnerability model not only offers greater accuracy in identifying Gauteng wards vulnerable to the disease, but it offers the added advantage of being a hybrid model that harmonises the criticisms labelled against the subjective (AHP) and objective (EWM) models that helped derive it.

While the integrated model has been proven to be the best-suited vulnerability model to help guide intervention and support mechanism efforts in Gauteng's COVID-19 fight, its iteration as derived in this study is not suited to be a decision-making tool yet. The reason for this is due to using outdated criteria indicator information such as those that were based on 2011 census information that no longer accurately capture the realities of modern-day life. If the integrated COVID-19 vulnerability model was to be the spatial targeting mechanism of choice to address COVID-19, then it would be ideal for it to be based on more recent criteria indicator information.

The upcoming 2022 South African population census and its follow up dataset offers the opportunity for the integrated COVID-19 vulnerability model to be revised with recent criteria indicator information that reflects the realities of modern-day life. This exercise would ensure that the integrated COVID-19 model derived from it, would be the definitive version that can function as a tool to support government decision-making and prioritisation efforts.

Overall, this study has demonstrated the potential of the integrated approach to modelling COVID-19 vulnerability, and its results indicate that the integrated approach is the most suited to be a supporting tool for the government.

7. BIBLIOGRAPHY

Acharya, R., & Porwal, A. (2020). A vulnerability index for the management of and response to the COVID-19 epidemic in India: An ecological study. *The Lancet Global Health*, 1142-1151.

Akhmadullina, A., Vusilyeva, S., Yakovleva, T., Vopiyashina, S., & Kraineva, R. (2019). Matrix of pairwise comparison: decision making with the analytic network process. *EDP Sciences Journal*, 1-9.

Alam, Z. (2021). Is Population Density a Risk Factor for Communicable Disease Like COVID-19? A Case of Bangladesh. *Asia Pacific Journal of Public Health*.

Alemdar, K. D., Kaya, O., Codur, M. Y., Campisi, T., & Tesoriere, G. (2021). Accessibility of vaccination centres in COVID-19 outbreak control: A GIS-based multi-criteria decision making approach. *International journal of Geo-information*, 1-19.

Ali, T., Mortula, M., & Sadiq, R. (2021). GIS based vulnerability analysis of United States to COVID-19 occurrence. *Journal of risk research*, 416-431.

Boyaci, A. C., & Sisman, A. (2021). Pandemic hospital site selection: A GIS based MCDM approach employin Pythagorean fuzzy sets. *Environmental science and pollution research*, 1985 - 1997.

Broadbent, A., Combrink, H., & Smart, B. (2020). COVID-19 in South Africa. *Global Epidemiology*, 1-4.

CDC. (2021, May 12). *Centre for Disease Control*. Retrieved from Science Brief: Evidence used to update the list of underlying medical conditions that increase a persons risk of severe illness from COVID-19: https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/underlying-evidence-table.html

Das, B., & Pal, S. C. (2020). Assessment of groundwater vulnerability to over-exploitation using MCDA, AHP, Fuzzy logic and novel ensemble models: A case study of Goghat-I and II blocks of West Bengal, India. *Environmental Earth Sciences*, 1-16.

Das, M., Das, A., Giri, B., Sarkar, R., & Saha, S. (2021). Habitat vulnerability in slum areas of India - What we learnt from COVID-19. *International Journal of Disaster Risk Reduction*, 1-16.

Devarakonda, P., Sadasivuni, R., Nobrega, R. A., & Wu, J. (2020). Application of spatial multicriteria decision analysis in healthcare: Identifying drivers and triggeres of infectious disease outbreaks using ensemble learning. *GeoJournal*, 1151-1174.

Dong, S. (2016). *Comparisons between different Multi-Criteria Decision Analysis Techniques for Disease Susceptibility Mapping.* Lund: Lund University.

Dunford, D., Dale, B., Stylianou, N., Lowther, E., Ahmed, M., & de la Torre Arenas, I. (2020, April 6). *Coronavirus: The world in lockdown in maps and charts*. Retrieved from BBC NEWS: https://www.bbc.com/news/world-52103747

ENCA. (2021, June 24). *Watch| COVID-19| Gauteng didn't prepare properly MAC Chair*. Retrieved from ENCA: https://www.enca.com/analysis/watch-covid-19-gauteng-didnt-prepare-properly-mlisana

GCRO. (2021). *Gauteng City - Region Observatory*. Retrieved from The Gauteng City - Region : https://gcro.ac.za/about/the-gauteng-city-region/

Gigovic, L., Drobnjak, S., & Pamucar, D. (2019). The Application of the Hybrid GIS spatial multi-0criteria decision analysis best-worst methodology for landslide susceptibility mapping. *International Journal of Geo-Information*, 1-29.

Hamann, C. (2020). *COVID-19 risk indices per municipality and metropolitan municipality planning region.* Johannesburg: GCRO.

Hengbo, Z., Wei, L., & Ping, N. (2020). The novel coronavirus outbreak in Wuhan, China. *Global Health Research and Policy*, 1-3.

IMF. (2020). *International Monetary Fund (IMF); Staff report for the 2019 Article IV Consultation.*

Jam, A. S., Mosaffaie, J., Sarfaraz, F., Shadfar, S., & Akhtari, R. (2021). GIS-Based landslide susceptibility mapping using hybrid MCDM models. *Natural Hazards*, 1-22.

Kirby, R. S., Delmelle, E., & Eberth, J. (2016). Advances in spatial epidemiology and geographic information systems. *Annals of epidemiology*, 1-9.

Kistemann, T., Dougendorf, F., & Schweikart, J. (2002). New perspectives on the use of Geopgraphical Information Systems (GIS) in environmental health sciences. *International journal of hygiene and environment health*, 169-181.

Lotfi, F. H., & Fallahnejad, R. (2010). Imprecise Shannons Entropy and Multi Attribute Decision Making. *Entropy*, 53-62.

Macharia, P. M., Joseph, N. K., & Okiro, E. A. (2020). A vulnerability index for COVID-19: Spatial analysis at the subnational level in Kenya. *BMJ Global Health*, 1-12.

Malakar, S. (2021). Geospatial modelling of COVID-19 vulnerability using an integrated fuzzy MCDM approach. *Modelling Earth systems and environment*, 1-14.

Manda, S. O., Darikwa, T., Nkwenika, T., & Bergquist, R. (2021). A spatial analysis of COVID-19 in African countries: Evaluating the effects of socio-economic vulnerability and neighbouring. *Environmental research and public health*, 1-15.

Mbunge, E. (2020). Effects of COVID-19 in South Africa health system and society: An explanatory study. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews* , 1809-1814.

Meterlerkamp, T. (2022, January 18). *Omicron and beyond: South Africa's fourth wave is waning, but don't celebrate just yet, says experts*. Retrieved from DAILY MAVERICK:

https://www.dailymaverick.co.za/article/2022-01-18-south-africas-fourth-wave-is-waning-but-dont-celebrate-just-yet-say-experts/

Mishra, S. V., Gayen, A., & Haque, S. M. (2020). COVID-19 and urban vulnerability in India. *Habitat Internation*, 1-11.

Municipalities. (2020). *Municipalities of South Africa*. Retrieved from Gauteng municipalities: https://municipalities.co.za/provinces/view/3/gauteng

Noble, M., Zembe, W., Wright, G., & Avenell, D. (2013). *Multiple Deprivation and Income Poverty at Small Area Level in South Africa in 2011.* Cape Town: SASPRI.

Nymbili, P. H., & Erden, T. (2020). A hybrid approach integrating Entropy - AHP and GIS for suitability assessment of urban emergency facilities. *International journal of Geo-Information*, 1-29.

Panchal, S., & Shrivastava, A. K. (2021). A comparative study of frequency ratio, Shannon entropy and analalytical hierachy process (AHP) models for landslide susceptibility assessment. *International journal of Geo-Information*, 1-20.

Pourghasemi, H. R., Pouryan, S., Farajzadeh, Z., Sadhasivam, N., Heifari, B., Babaei, S., & Tiefenbacher, J. P. (2020). Assessment of the outbreak risk mapping and infection behaviour of COVID-19: Application of the autoregressive integrated-moving average (ARIMA) and polynomial models. *PlosONE*, 1-25.

Rahman, H., Zatri, N. M., Ashik, F. R., Waliullah, M., & Khan, A. (2021). Identification of risk factors contributors to COVID-19 incidence rates in Bangladesh: A GIS-based spatial modelling approach. *Heliyon*, 1-11.

Rodriguez-Morales, A. J., Orrego-Acevedo, C. A., Zombrano-Munoz, Y., Garcia-Folleco, F. J., Herrera-Giraldo, A. C., & Lozardo, C. O. (2015). Mapping malaria in municipalities of the Coffee Triangle region of Colombia using Geographic Information Systems (GIS). *Jounral of infection and public health*, 603-611.

Saaty, T. L. (1980). *The Analytical Hierarchy Process: Planning, Priority Setting, Resource Allocation.* New York: McGraw.

Saaty, T. L. (2008). Decision making with the analytical hierarchy process. *International Journal of Services Sciences*, 83-98.

Sarkar, S. K. (2020). COVID-19 Susceptibility mapping using multicriteria evaluation. *Disaster Medicine and Public Health Preparedness*, 521-537.

SERI. (2018). *Socio-Economic Rights Institute of South Africa (SERI). Informal Settmenets and Human Rights in South Africa:Submission to the United Nations Special Rapporteur on Adequate Housing as a Component of the Rights to an Adequate Standard of Living.*

Shadeed, S., & Alawna, S. (2021). GIS-Based COVID-19 vulnerability mapping in the West Bank, Palestine. *Internation Journal of Disaster Risk Reduction*, 1-8.

Sharma, L. P., Ghose, M. K., Patel, N., & Debnath, P. (2010). Influence of Shannon entropy landslide causing parameters for vulnerability study and zonation- A case study in Sikkim, India. *Arab journal of geosciences*, 1-12.

STATSSA. (2017). *Statistics South Africa: Poverty trends in South Africa.*

STATSSA. (2020, July 9). *Statistics South Africa*. Retrieved from 2020 Mid year population estimates: http://www.statssa.gov.za/?p=13453

Velasquez, M., & Hester, P. T. (2013). An analysis of multi-criteria decision making methods. *International Journal of Operations Research*, 56-66.

Wang, L. C., & Zhang, J. Y. (2015). Assessment of water resource security in Chongqing City of China: What has been done and what remains to be done? *Natural Hazards*, 2751-2772.

WHO. (2020). *Novel Coronavirus (2019-nCoV) SITUATION REPORT - 1.* Geneva: WHO.

Wu, J. H., Xue, C. Y., Tian, R., & Wang, S. (2017). Lake water quality assessment: A case study of Shakhu lake in the semi-arid loess area of north west China. *Environmental Earth Sciences*, 1-15.

Wu, J., Li, P., Qian, H., & Chen, J. (2015). On sensitivity of entropy weight to sample statisitics in assessing water quality: statisitical analysis based on large stochastic samples. *Environmental Earth Science*, 2185 - 2195.

Yariyan, P., Avand, M., Soltani, F., Ghorbanzadeh, O., & Blaschke, T. (2020). Earthquake vulnerability mapping using different hybrid models. *Symmetry*, 1-31.

Yue, Z. (2010). *A method for group decision-making based on determining weights of decision makers using TOPSIS.* Zhanjiang: Elsevier.

Zafar, S., Shipin, O., Paul, R. E., Rocklov, J., Haque, U., Rahman, M. S., . . . Overgaard, H. J. (2021). Development and comparison of Dengue vulnerability indices using GIS-based multi-criteria decision analysis in Lao PDR and Thailand. *Environmental research and public health*, 1-25.

Zhu, Y., Tian, D., & Yan, F. (2020). Research Article: Effectiveness of Entropy Weight Method in Decision- Making. *Hindawi. Mathematical Problem in Engineering*, 1-5.

8. ANNEXURES

ANNEXURE A: Research questionnaire for the AHP component of the research

The AHP is a multi-criteria decision analysis (MCDA) method that uses stratified structures to analyse decision-making criteria through a pairwise comparison matrix that assigns criteria weights (Nymbili & Erden, 2020). The scale of relative importance (see table 1) as proposed by Thomas Saaty is used to determine the relative importance between sets of criteria for the pairwise matrix.

Annexure table 1

CHOICE	IMPORTANCE VALUE
Equally Important	1
Moderately Important	3
Strong Importance	5
Very Strongly Important	7
Extreme Importance	9
Values In Between Importance	2,4,6,8

This questionnaire is designed for an expert to evaluate the relative importance between sets of criteria identified as important in determining COVID-19 vulnerability. Feedback from this questionnaire will construct a pairwise matrix and subsequent criteria weights to be used for the COVID-19 vulnerability map. The criteria used for determining COVID-19 vulnerability are described in table 2 below.

Annexure table 2

CRITERIA FACTOR	DESCRIPTION FOR COVID-19 VULNERABILITY
1) INCOME POVERTY	Income poverty is crucial in assessing COVID-19 vulnerability because it determines services and the quality of environment an individual has access to. The poorer the individual, the less ideal their living environment is likely to be towards COVID-19 prevention and the less likely the individual is to afford quality medical care.
2) EDUCATION DEPRIVATION	Education deprivation is a proxy for the level of awareness an individual will likely have regarding COVID-19 and the prevention strategies recommended for it. The assumption is that people with lower levels of education will be more vulnerable to COVID-19.
3) LIVING ENVIRONMENT DEPRIVATION	Living environment deprivation is a pivotal domain in determining COVID-19 vulnerability because it is a domain that encompasses individuals living in

	impoverished conditions i.e. informal settlements with minimal access to basic services such as water and sanitation with the former (water) being critical in COVID-19 prevention.
4) NUMBER OF HEALTHCARE CENTRES IN THE WARD	Having immediate or readily accessible healthcare facilities is important to COVID-19 mitigation. With this logic, areas with fewer healthcare centres could be regarded as more vulnerable to COVID-19.
5) POPULATION DENSITY	The rationale behind population density being a risk factor for COVID-19 is based on the idea that when people live close to each other, the potential to spread the disease is heightened (Alam, 2021).
6) PRE-EXISTING HEALTH CONDITIONS	It is indicated that people who suffer from certain pre-existing health conditions such as cancer, diabetes, heart conditions, HIV, Asthma etc. are more prone to severe COVID-19 illness (CDC, 2021).

Below are example scenarios of how questionnaire responses should be rationalised and interpreted.

Example 1: In comparing the relative importance between income poverty and education deprivation, table 3 illustrates that **income poverty** is considered 3 times more important than **education deprivation** in determining COVID-19 vulnerability.

		Extreme importance				Strong importance				Equal importance				Strong importance				Extreme importance	
		9	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	
1	Income Poverty							✓											Education Deprivation

Example 2: In comparing the relative importance between income poverty and education deprivation, table 4 illustrates that **income poverty** is considered equally important (1) to **education deprivation** in determining COVID-19 vulnerability.

Extreme importance Strong importance Equal importance Strong importance Extreme importance

		9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	
1	Income Poverty									✓									Education Deprivation

Example 3: In comparing the relative importance between income poverty and education deprivation, table 5 illustrates that **education deprivation** is considered 3 times more important than **income poverty** in determining COVID-19 vulnerability.

Table 5: Income poverty vs Education deprivation

Extreme importance Strong importance Equal importance Strong importance Extreme importance

		9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	
1	Income Poverty											✓							Education Deprivation

Using these examples as guidelines, please complete the questionnaire below by placing a tick (**✓**) in the importance value of your choosing.

Please compare the relative importance of each set of criteria factors used to determine COVID-19 vulnerability in Gauteng using the scale below.

Extreme importance Strong importance Equal importance Strong importance Extreme importance

9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9

1	Income Poverty																		Education Deprivation
2	Income Poverty																		Living Environment Deprivation
3	Income Poverty																		Number of Healthcare Centres in the Ward
4	Income Poverty																		Population Density
5	Income Poverty																		Pre-Existing Health Condition
6	Education Deprivation																		Living Environment Deprivation
7	Education Deprivation																		Number of Healthcare Centres in the Ward
8	Education Deprivation																		Population Density
9	Education Deprivation																		Pre-Existing Health Condition
10	Living Environment Deprivation																		Number of Healthcare

																			Centres in the Ward
11	Living Environment Deprivation																		Population Density
12	Living Environment Deprivation																		Pre-Existing Health Condition
13	Number of Healthcare Centres in the Ward																		Population Density
14	Number of Healthcare Centres in the Ward																		Pre-Existing Health Condition
15	Population Density																		Pre-Existing Health Condition

Participant Field of Expertise:

Number of Years in this field: