Modelling the likelihood of wetland occurrence in KwaZulu-Natal, South Africa: a Bayesian approach

by

Jens Hiestermann

206507774

Submitted in fulfilment of the academic requirements of
Masters of Science Degree in Environmental Science

School of Agriculture, Earth and Environmental Sciences
University of KwaZulu-Natal
Pietermaritzburg

February 2014

DEDICATION

This thesis is dedicated to my parents, Gerald and Angelika Hiestermann.

DECLARATION 1

This dissertation was conducted at the School of Environmental Science, Discipline of Geography, University of KwaZulu-Natal, Pietermaritzburg, from February 2011 to January 2013, under the supervision of Dr Nick Rivers-Moore and Professor Onisimo Mutanga.

This dissertation represents the original work by the researcher, and has not been submitted to any other university. Where work of other authors has been used, it has been duly acknowledged in both the text and the reference list.

JENS HIESTERMANN (Candidate): _________________________

PROFESSOR ONISIMO MUTANGA (Supervisor): _________________________

DR NICK RIVERS-MOORE (Co-supervisor): _________________________

DECLARATION 2

Publications that form part of and/or include research presented in this thesis:

Hiestermann, J. and Rivers-Moore, N.A. 2012. Mapping wetlands in KwaZulu-Natal: The Bayesian approach. This was submitted to the journal *Wetland Ecology and Management* and is currently under review.

The work was done by the first author under the guidance and supervision of the second author.

Discipline of Geography, School of Environmental Sciences, Faculty of Science and Agriculture, University of KwaZulu-Natal, Pietermaritzburg Campus, Private Bag X01, Scottsville, 3209, South Africa.

ACKNOWLEDGEMENTS

Firstly I would like to thank the Lord for the opportunity to do my Masters of Science degree and being by my side in the difficult times. A special thank you goes to my parents, who have helped through my years of tertiary education, even though it was sometimes financially taxing.

I would like to thank the Wildlife and Environment Society of South Africa (WESSA) and the Mondi Wetlands Programme for providing me with an opportunity to do my MSc while gaining valuable work experience. My sincere thanks go to Damian Walters, Vaughan Koopman, Michelle Hiestermann and David Lindley for your guidance, motivation, patience and encouragement throughout the two years of this research.

Thank you to Dr Boyd Escott from Ezemvelo KZN Wildlife for helping me with many GIS-related questions, data requests and unexpected visits during these years. Your experienced assistance has been greatly appreciated.

Finally, with gratitude, a special thank you to Dr Nick Rivers-Moore, who has supervised me throughout this thesis and who has generously given his time, patience and valuable advice throughout the research process. Special thanks also go to Professor Onisimo Mutanga, for your support and for always keeping me within the academic realm in this research.

ABSTRACT

Global trends of transformation and loss of wetlands to other land uses has deleterious effects on surrounding ecosystems, and there is a resultant increasing need for improved mapping of wetlands. This is because wetland conservation and management depends on accurate spatial representation of these systems. Current approaches to mapping wetlands through the classification of satellite imagery typically under-represent actual wetland area, and the importance of ancillary data in improving the accuracy in mapping wetlands is recognized. This study uses likelihood estimates of wetland occurrence in KwaZulu-Natal (KZN), South Africa, using a number of environmental surrogate predictors (such as slope, rainfall, soil properties etc.). Using statistical information from a set of mutually independent environmental variables in known wetland areas, conditional probabilities were derived through a Bayesian network (BN) from which a raster layer of wetland probability was created. The layer represents the likelihood of wetlands occurring in a specific area according to the statistical conditional probability of the wetland determinants. Probability values of 80% and greater also accounted for approximately 6% of the KZN area (5 520 km²), which is substantially more than the previously documented wetland area in KZN (4% of the KZN area or 4 200 km²). Using an independent test dataset, Receiver Operating Characteristic (ROC) curves with the Area Under Curve (AUC) analysis verified that the final model output predicted wetland area well (AUC 0.853). Based on visual comparisons between the probability layer and ground verified wetland systems, it was shown that high wetland probability areas in the final output correlated well with previously highlighted major wetland and wetland-rich areas in KZN. Assessment of the final probability values indicated that the higher the probability values, the higher the accuracy in predicting wetland occurrence in a landscape setting, irrespective of the wetland area. It was concluded that the layer derived from predictor layers in a BN has the potential to improve the accuracy of the KZN wetland layer by serving as valuable ancillary data. Application of the final probability layer could extend into the development of updated spatial freshwater conservation plans, potentially predicting the historical wetland extents, and as input into the land cover classification process.

Keywords: ancillary data, Bayesian network, GIS, modelling, probability, wetland mapping

Table of Contents

List of Figures

List of Tables

Abbreviations

AMSL	Above Mean Sea Level
BN	Bayesian Network
BP	Before Present
DEA	Department of Environmental Affairs
DEAT	Department of Environmental Affairs and Tourism
DEM	Digital Elevation Model
DWE	Department of Water and Environment
DWAF	Department of Water Affairs and Forestry
EIA	Environmental Impact Assessment
CPT	Conditional Probability Table
EKZNW	Ezemvelo KZN Wildlife
ESRI	Environmental Systems Research Institute
ETM	Enhanced Thematic Mapper
FEPA	Freshwater Ecosystem Priority Areas
GIS	Geographic Information System/s
KZN	KwaZulu-Natal
LWP	Landscape Wetness Potential
NBSAP	National Biodiversity Strategy and Action Plan
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Wetness Index
NEMA	National Environmental Management Act
NFEPA	National Freshwater Ecosystem Priority Areas
NGO	Non-Governmental Organisation
NWI	National Wetland Inventory
NWM	National Wetland Map
MBVI	Multi-band Vegetation Indices
MODIS	Moderate-resolution Imaging Spectroradiometer
MESMA	Member Spectral Mixture Analysis
MVSP	Multivariate Statistical Package
NSBA	National Spatial Biodiversity Assessment
PCA	Principle Component Analysis
RSA	Republic of South Africa
SPOT	Système Pour l’Observation de la Terre
SRTM	Shuttle Radar Topography Mission
SWB	Standing Water Body
TCWI	Tasseled Cap Wetness Index
TBVI	Two-band Vegetation Indices
UKZN	University of KwaZulu-Natal

CHAPTER ONE

1 Introduction

Wetlands constitute a resource of great economic, cultural, scientific and recreational value to human life. Wetlands and people are ultimately interdependent. As such, the progressive encroachment on, and loss of, wetlands needs to be stopped and measures must be taken to conserve and make wise use of wetland resources (Ramsar, 2011:11). *Despite the economic and ecological importance of wetlands, there are many uncertainties regarding their extent, distribution, and ecological and physical functions, and if you cannot measure wetlands you cannot manage them* (Bwangoy *et al.*, 2010:73).

Wetlands are highly productive systems supporting unique flora and fauna, and are important storehouses of genetic plant material (Chhokar *et al.*, 2004; De Voogt *et al.*, 2000; Kidd, 2011). Wetlands are described as the ‘kidneys of the landscape,’ and being downstream receivers of waste from both natural and human sources, they cleanse polluted waters, prevent floods, protect shorelines and recharge groundwater aquifers (Islam *et al.*, 2008; Mitsch and Gosselink, 1993). Habitat loss, fragmentation, ecosystem disruption and global warming severely threaten the diversity of wetlands and their species (May *et al.*, 2002; Olhan *et al*. 2010; Woodhouse *et al*., 2000). Because of the close relationship that exists between freshwater systems and catchment conditions, the conversion of wetlands to other land uses has deleterious effects on the surrounding ecosystems and results in the loss of important ecosystem services (Rivers-Moore *et al.*, 2010). As a result, wetlands have gained considerable recognition globally over the past 20 years as people have begun to realise the importance of managing and conserving them.

1.1 Wetland mapping and inventory

To prevent further loss, and to conserve existing wetland ecosystems for their biodiversity and ecosystem goods and services, it is important to develop an inventory of wetlands (Ozesmi and Bauer, 2002; Ramsar, 2010). Wetland inventories form the baseline data layer, which can be used for many purposes, including comprehensive resource management plans, environmental impact assessments, natural resource inventories, habitat surveys, and the trend analysis of wetland status (Adam *et al.*, 2010; Ozesmi and Bauer, 2002; Wilen *et al.*, 2002). A wetland inventory in this study refers to ‘the collection and/or collation of core information for wetland management, including the provision of an information base for specific assessment and monitoring activities’ (Ramsar, 2007: 7). Critical to building a wetland inventory is the process of mapping wetlands and building the necessary information, such as the wetland type, size and location. The contracting parties to the Ramsar Convention on Wetlands have recognised the importance of identifying the location and characteristics of wetlands (Ramsar, 2010). Conservation usually requires a blueprint of the biodiversity patterns and processes, and a good understanding of biodiversity and the factors influencing species distribution and status. Creating an accurate inventory of wetlands, in terms of wetland type and location, can help to achieve this (Rivers-Moore and Goodman, 2010). A wetland inventory is useful in understanding the spatial distribution of different wetlands and their linkages with other land units. It is also important in the conservation and management of valuable ecosystems. Thus, the identification of wetlands not only helps to establish a baseline for measuring future change in wetland area, function and values, but also facilitates surveys to provide wetlands with a condition and conservation status.

Historical events, technological innovations and the changing values of society have shaped the advancements in the way wetlands are inventoried. In the seventeenth and eighteenth centuries, colonial settlers detailed wetlands in their sketch maps; this now provides valuable information of the historical extent of the wetlands at those times (Dahl and Allord, 1997). Wetlands inventoried in the early 1900s were used to identify which wetland areas could be drained and converted to other land uses (Wilen and Tiner, 1993). The First and Second World Wars saw the development and advancement of aerial imagery for military purposes, but following the development of aerial photograph interpretation methods in WWII, aerial photographs were used in numerous fields, primarily land use and soil maps. The emergence of landscape ecology led to aerial photographs being used to map wetlands (Johnson, 2003). In the late 1960s this progressed to the use of satellite remote sensing and Geographic Information Systems (GIS), which have since been at the forefront in wetland mapping exercises globally.

Wetland systems often cover large spatial domains and recording and delineating all wetlands in an area cannot always rely on field surveying methods, because these methods are time consuming and costly. Remote sensing of satellite imagery is the most cost-effective, least time consuming and most consistent alternative to field surveys and allows wetland mapping across larger geographical areas (Kulawardhana *et al.,* 2007; Ozesmi and Bauer, 2002; Rebelo *et al.*, 2009). There is a plethora of literature highlighting the successful approaches and advantages of mapping wetlands using remote sensing and GIS (Islam *et al.,* 2008; Knight *et al.*, 2009; Kulawardhana *et al.*, 2007; Landmann *et al.*, 2010; Li and Chen, 2005; Lunetta *et al.*, 1999; Ozesmi and Bauer, 2002; Ryo *et al.*, 2012; Thompson *et al.*, 2002). The advancements in temporal resolution (the time the satellite takes to revisit the same geographical area) of satellite imagery allows wetlands to be monitored seasonally or yearly. In recent years there has been a rapid advancement in the spectral, spatial and temporal resolution of satellite imagery, which has led to ongoing improvements in wetland classification methods (O'Hara, 2002).

In spite of these technological advancements, and the practical advantages of using remote sensing to record and delineate wetlands, the extant literature elucidates certain limitations associated with the classification of satellite imagery. Spectral confusion in satellite imagery results in the misclassification of pixels and is a common limitation in the classification of satellite imagery. Spectral confusion can cause the misclassification of pixels, where, for example, fluctuating water levels (which alter the spectral reflectance of the vegetation), or fire scars and hill shading, are classified as wetland vegetation or open water (Ozesmi and Bauer, 2002).

The literature also highlights the importance of using ancillary data to increase the accuracy of wetland mapping (Kulawardhana *et al.*, 2007). For example, O'Hara (2002: 3) states that spectral analysis (classification of satellite imagery) must be followed by data-fusion techniques combining and analysing spectral products with topographic, hydrologic information and soils data. Ancillary data can be in the form of topological variables (e.g. slope, elevation, flow accumulation, etc.), environmental characteristics (e.g. soil characteristics, geology, rainfall, evaporation, etc.), and probability models (e.g. terrain-based hydrological models). These have been used to improve the accuracy of many satellite-image classification techniques, including approaches to wetland mapping (Kulawardhana *et al.*, 2007; Ricchetti 2000).

The existing wetland map for KwaZulu-Natal (KZN), South Africa (Scott-Shaw and Escott, 2011) is a compilation of the best available wetland data covering KZN, received from various sources and organisations. The wetland map of KZN is continually updated as field surveys of wetlands are completed or following provincial land cover classification updates using the latest satellite SPOT (Système Pour l'Observation de la Terre) imagery. The use of reliable ancillary data in future land cover classification efforts could improve the current KZN wetland layer substantially in terms of reliability and accuracy. An accurate spatial representation of wetland location on a provincial scale is valuable not only in building a complete wetland inventory, but also in the development of plans, measures and targets in terms of freshwater conservation. Conservation of freshwater systems is achieved by selecting and protecting key areas of habitat, especially wetlands. These key areas can then be prioritised to maximise potential freshwater and biodiversity gains, and a wetland inventory would help to achieve this (Woodhouse *et al.*, 2000).

This research study employs a non-image-based wetland mapping approach to model the likelihood of wetland occurrence in KZN. The approach uses probabilities derived from a Bayesian network (BN) to produce a wetland probability map at a regional scale. Since the literature highlights the importance of using ancillary data in wetland mapping approaches, this study aims to develop a probability model using other defining variables (topological and environmental parameters) that could then improve the identification of wetlands in the landscape. This research differs from traditional probability models because it uses Bayesian conditional probabilities (see section 2.2.3 for further description). Using environmental parameters in a BN, estimates of those parameters corresponding with identified wetland areas are used as evidence and knowledge to model probabilities of wetlands occurring in other areas.

BNs are considered probabilistic network-graphical models and are used for probabilistic inference. They are based on Bayes' Theorem, a mathematical formula used to calculate probabilities among several variables that are causally related. Bayesian statistics combine a priori probabilities with the probabilities of occurrence conditional to the value (or class of values) of each environmental predictor (Guisan and Zimmermann 2000). Dlamini (2010) used a BN to integrate the biotic, abiotic, and human factors for predicting the likelihood of wildfire activity in Swaziland. Using variable nodes including fire activity, land cover, elevation, temperature, distance to settlements, and road density, Dlamini (2010) found the BN to have a high predictive accuracy. The distribution of wetlands across a landscape is influenced by uncharacterised spatial interactions between many environmental variables. Wetlands occur in different climatic, geomorphic, hydrologic or physiochemical environments, and are not defined by a particular combination of interactions. This study therefore uses a BN to integrate uncharacterised spatial interactions between environmental variables in wetlands in order to predict the locality and occurrence of wetlands successfully over a large spatial domain.

Research involving the use of Bayesian conditional probabilities to model the occurrence of wetlands over a large spatial domain has not yet been published. This study therefore constitutes a novel approach to wetland mapping.Where previously the use of BNs was largely restricted to the fields of mathematics and computer sciences (Aguilera *et al.*, 2011; Kocabas and Dragicevic 2007), they are becoming an increasingly popular method of modelling uncertain and complex domains such as those encountered in ecosystems and environmental management (Uusitalo, 2007). As in this study, BNs are mainly applied as a technique for inference, using discretised continuous variables (continuous data that has been condensed into a number of qualitative classes). The outcomes of BNs in environmental studies are usually not represented spatially, yet a spatial representation and understanding of these outcomes is essential for informed decision-making in land-use management and planning (Grêt-Regamey and Straub, 2006). This research study therefore aims to represent its results spatially, in order for the results to be accurately understood and easily applied by decision makers.

1.2 Aim and objectives

The aim of this study was to model the likelihood of wetland occurrence in KZN using conditional probabilities derived through a BN. The usefulness of this wetland probability layer was assessed for its ability to act as ancillary data to improve and refine the current wetland layer for KZN. The following objective were identified to achieve this aim:

- Identification of a suite of mutually independent environmental variables that correlate with wetland presence;
- Development of a likelihood of wetland occurrence image using a Bayesian network model; and
- Assessment of the usefulness of the abovementioned layer as ancillary data.

1.3 Thesis structure

Chapter Two sets out to review the existing literature on the key theory informing this research. The chapter begins with a section on the definition, importance and conservation of wetlands, and the associated legislation. Next, methods and approaches to mapping wetlands are evaluated in an international and local context, and the chapter ends with a consideration of the theory on the applicability of probability modelling in mapping wetlands.

Chapter Three explores the methods and approaches used to achieve the aims and objectives of this study. This chapter outlines the building of the inventory dataset, the extraction of a statistical dataset for the Bayesian statistics used in generating conditional probabilities, and spatially represents the derived conditional probabilities in GIS.

Chapter Four outlines the results of the final probability layer of modelling the likelihood of wetland occurrence in KZN obtained using the methods described in Chapter Three. Furthermore, the results obtained through visual assessment and accuracy assessment are assessed in this chapter.

Chapter Five provides a discussion of the results obtained and critically examines the techniques used to obtain the final probability layer. The accuracy of the output using this method and approach and the applicability of such an output is also discussed. Results obtained in this study will be related to previous research. Recommendations for future research are discussed.

CHAPTER TWO

2 Literature Review

Since this study focuses on the identification of wetlands, the literature review begins with an overview that covers the definition of wetlands, their formation, legislation and regulations pertaining to them, and their conservation. The literature review then revisits past research on wetland mapping approaches, both internationally and locally, and identifies certain limitations and gaps that this study aims to address. The idea of using Bayesian statistics in modelling the occurrence of wetlands is unique; therefore, the review also builds an understanding of the conceptual probability approach by highlighting some examples relating to the idea.

2.1 Introduction

Globally, the conservation of wetlands began in the mid-1970s. Prior to this, the destruction and drainage of these systems was common practice, and wetlands were seen as a nuisance, as health hazards, and as good potential agricultural and commercial development areas once drained (Mitsch and Gosselink, 2000). Wetlands were seen as flat open areas that would, once drained, provide land that could easily be modified for agriculture, and the perceived benefits included accessibility, good soils and the possibility of a water source for irrigation (Heimlich *et al.*, 1998). These perceptions led to the destruction of wetlands and the loss of these valuable ecosystems. The disappearance of wetlands led to undesirable consequences, such as the loss of groundwater reserves and the consequent need for irrigation, flash floods, shoreline destruction, and the accumulation of pollutants (Ramsar, 1993). These undesirable consequences resulted in the loss of a great deal of wetland fauna and flora that relied on those systems. It was only due to the combined concern and activities of hunters, scientists, engineers, lawyers and environmentalists that a case for preserving wetlands as a valuable resource was made. Through such collective efforts, society at large began to realise that the destruction of these systems had serious widespread economic as well as ecological and aesthetic consequences, and research studies began to highlight this (Mitsch and Gosselink, 2000).

Wetlands provide important provisioning services (e.g. food, freshwater, fibre and fuel, biochemical products, genetic materials), regulating services (e.g. climate regulation, water regulation, water purification and waste treatment, erosion regulation, natural hazard regulation, pollination), cultural services (e.g. spiritual and inspirational, recreational, aesthetic, educational) and supporting services (e.g. soil formation and nutrient cycling) (MEA, 2005). It is beyond the scope of this thesis to provide an in-depth review of these goods and services, but they deserve some attention as they provide a strong foundation for why wetlands are important and why mapping these systems is necessary. The transboundary nature of wetland systems and the realisation of their importance resulted in the need for international agreements and arrangements to use wetlands wisely and preserve them. In 1971 this gave rise to the Convention on Wetlands of International Importance (more commonly known as the Ramsar Convention). The Ramsar Convention recognised the importance of wetlands as valuable ecosystems, and the importance of compiling national wetland inventories as a key tool for informing policies and other actions to achieve the conservation and wise use of wetlands (Ramsar, 2010).

2.1.1 Defining wetlands

The Ramsar Convention (Ramsar, 2006: 7) defined wetlands as "areas of marsh, fen, peat land or water, whether natural or artificial, permanent or temporary, with water that is static or flowing, fresh, brackish or salt, including areas of marine water the depth of which at low tide does not exceed six meters". Wetlands are areas that are seasonally or permanently waterlogged or saturated, and are characterised by vegetation that has adapted to survive in saturated soil conditions. Heimlich *et al.* (1998: 10) define a wetland as "land that (1) has a predominance of hydric soils; (2) is inundated or saturated by surface or groundwater at a frequency and duration sufficient to support a prevalence of hydrophytic vegetation typically adapted for life in saturated soil conditions; and (3) under normal circumstances does support a prevalence of such vegetation." A complicating factor in defining wetlands is that they occur in a wide variety of hydrologic conditions and also vary in geographical extent, making it difficult to delineate or spatially map and define them. Scientists and researchers realise the complex nature of wetlands and have even suggested treating wetland ecology as a distinct field of study (Mitsch and Gosselink, 2000). For the purposes of this study, wetlands will be defined as per the South African National Water Act (Act 36 of 1998: 18) (RSA, 1998a), where a wetland is defined as "land which is transitional between terrestrial and aquatic systems where the water table is usually at or near the surface, or the land is periodically covered with shallow water, and which land in normal circumstances supports or would support vegetation typically adapted to life in saturated soil."

2.1.2 The formation of wetlands

"Every wetland is a unique actualization of many abiotic and biotic factors, including geologic and geomorphic history, topography, connections to the local and regional hydrologic system, connections to local and regional ecosystems, time since formation, and disturbance history. Wetlands defy simple classification" (Batzer and Sharitz, 2006: 67).

The physical factors that control the distribution of wetlands in the landscape along these paths interact at a variety of scales: from regional (climate), to the watershed (hydrology, water chemistry), to the local/wetland (basin morphology and soil properties). The formation of wetlands in the landscape are directly linked to the global cycling of water, where large quantities of water evaporated off the ocean are carried over land and eventually fall to the ground as rain or snow. Through gravity, water naturally flows back into the ocean either through surface flow or groundwater flow. It is along these myriad flow paths that freshwater wetlands are found (Van Der Valk, 2006), and it is the silent work of water that shapes the landscape. "Water in the hydrological system weathers, transports and deposits rock and soil material in continental settings or coastal areas, shaping the landscape in fundamental ways" (Ellery *et al.*, 2008: 28). The formation, size and persistence of wetlands are controlled ultimately by these hydrological factors, which cause prolonged and shallow saturation.

The three key factors in the formation of the hydrologic conditions found in wetlands are climate, topography and geology (Figure 2.1). These cause the favourable conditions for wetland hydrology and soils to form, which competitively favour hydrophilic vegetation. Wetlands tend to be more prevalent in cool and wet climates than in hot and drier climates, because cool climates have less water loss from the land via evapotranspiration and wet climates have excess precipitation (see section 2.1.3) (Mitsch and Gosselink, 2007). With regard to the geomorphology of the landscape and basin, flat and gently sloping terrains tend to have far more wetlands than steep sloping landscapes, and therefore wetlands occur in a relatively limited set of positions (Junhua and Wenjun, 2005).

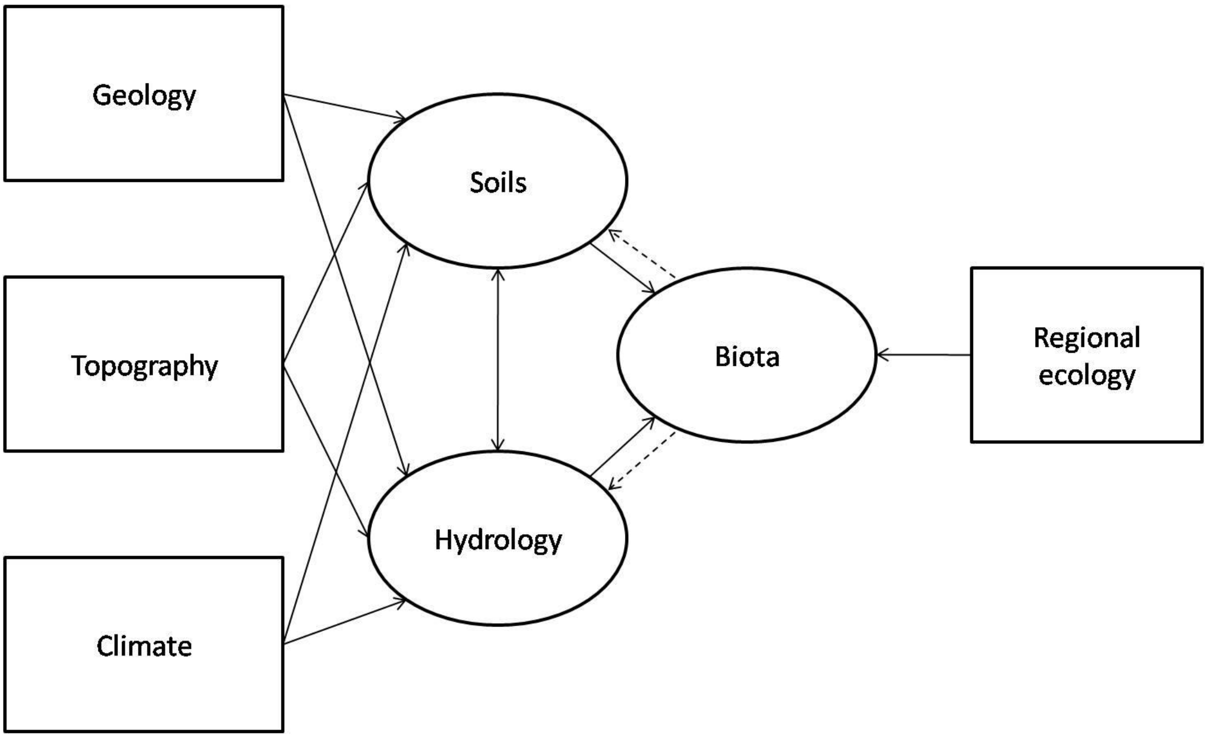


Figure 2.1: Abiotic and biotic influences on wetlands (Batzer and Sharitz, 2006).

The physical constitution of geology plays an important part in determining whether a wetland forms at a particular site (Moore, 2006). Wetlands form where water accumulates and therefore it is important that the ground underlying the basin be impervious to water, or at least limit the rapid loss of surface water from an area via stream flow. Due to the constantly changing river and sediment dynamics, wetlands are constantly being destroyed, shifting location, and being created (Batzer and Sharitz, 2006).

The most basic requirement for wetland formation is the sustained saturation of the upper part of the substrate. The importance of hydrology in determining the presence or absence of a wetland is well accepted, but the threshold conditions that satisfy the hydrologic criteria are complex and continually in need of study. "The hydrologic state of any wetland is a result of the balance between the water that flows into and out of a wetland, the surface contours of the landscape, the subsurface soil, geology, and groundwater conditions producing enough saturation to maintain substrate and biota that are characteristic of wetlands" (Mitsch and Gosselink, 2007; NRC, 1995). The frequency of saturation is dependent on the regional differences in physiographic, climatic and antecedent moisture conditions.

The major components that influence the hydrologic budget for wetlands are precipitation, surface inflows and outflows, groundwater, evapotranspiration and tides (Mitsch and Gosselink, 2007). Regions where precipitation exceeds the loss of water through evapotranspiration and surface runoff, are usually areas where wetlands occur most extensively (exceptions to this are wetlands fed by tides or riparian wetlands dependent on river flow). Patterns of surface inflows and outflows of a wetland usually correlate with the precipitation patterns in the same area. A study in the Okavango Delta in southern Africa indicated that the river input into the delta was approximately equivalent to the rainfall averaged over the entire delta, but only 1% of the inflows left the wetland region because the water budget is essentially balanced through the loss of water by evapotranspiration in this semi-arid climate. Meteorological conditions such as solar radiation and surface temperature enhance evapotranspiration by increasing the value of the vapour pressure at the evaporating surface, and therefore it is mostly where the surface inflows into an area exceed the evapotranspiration values that wetlands are found (Mitsch and Gosselink, 2007).

2.1.3 Wetlands in South Africa

The climate experienced in the southern regions of the African continent ensures that South Africa is predominantly dry. For most of the year, a high pressure system prevails over the interior of the country and its divergent upper air masses tend to circulate eastwards, causing minimal precipitation and clear dry weather. Southerly and southwesterly winds bring rain to the southern coastal plains and mountains in the winter months, while southeasterly and northeasterly winds bring moisture off the Indian Ocean, bringing rain onto the east coast during the summer months (Hughes and Hughes, 1992). South Africa is an arid country, with an average annual rainfall of about 497 mm for the country, which is well below the world average of 860 mm (Cowan, 1995). The mountains off the southern and eastern coastal plains cause a rain shadow across the interior, causing this area to be predominantly dry, but also create high water yield areas to the leeward side of the escarpment. High water yield areas are important in the overall water supply of the country and are sub-quaternary catchments where the mean annual runoff (mm per year) is at least three times the average for the related primary catchment.

The rainfall patterns in South Africa are key in sustaining inland wetlands, because the rainfall becomes runoff (surface flow or river flow), or contributes to the groundwater recharge that enters a wetland (subsurface) flow, or because it falls directly on the wetland (Ellery *et al.*, 2008). The water balance in a wetland is determined by its inputs as well as the atmospheric demand. Potential evapotranspiration is an indicator of atmospheric demand, and solar radiation provides the energy that drives evapotranspiration. Due to South Africa’s climatic conditions, potential evapotranspiration increases as one moves westwards across the country, due to the presence of clear skies (clear skies result in higher levels of solar radiation). Given the rainfall patterns across the country, and the potential for evapotranspiration being the greatest in the western parts, it would be expected that the greatest abundance of wetlands should be in the eastern parts of South Africa (Ellery *et al.,* 2008). South Africa has a relatively low mean annual rainfall, therefore it can’t be expected that wetlands rely primarily on rainfall. Areas such as the Lesotho Highlands and the eastern and southern coastal regions have a positive water balance in terms of the difference between water gain and water loss. Alternatively, wetlands can be sustained by groundwater flow and hill-slope seepage but these wetlands are likely to be small and situated at or near to sea level, where the groundwater table is close to the land surface (i.e. the coastal plains of KwaZulu-Natal and along the southern Cape coast). Most wetlands in South Africa occur as part of the country’s drainage system (i.e. they are fed by a river network).

The geomorphology and climate found in South Africa today was largely influenced by a subcontinental uplift that occurred in two pulses around 20 million and 5 million years ago. The uplift tilted southern Africa, raising the eastern regions by approximately 1 150 m and the western regions by approximately 250 m (Ellery *et al.,* 2008). This uplift resulted in the western regions becoming even drier as air masses originating from the warm Indian Ocean lost most of their moisture as they rose over the newly elevated escarpment. The uplift was responsible for increased river gradients, which resulted in new pulses of erosion that created new erosional plains at a lower elevation and a striking topography along the coastal rivers (e.g. the Valley of a Thousand Hills in KZN).

Wetlands in South Africa are the products of these erosional (degradational) and depositional (aggradational) processes, as well as the presence of geological controls. Resistance to erosion varies in different geological settings, but has a profound effect on wetland formation. The subcontinental uplift caused a greater river energy to degrade the valleys, but this was only due to the presence of erosion-resistant geology that retarded the downcutting of the rivers further, inhibiting the movement of water away from a site, and reducing the slope gradient (and thus reducing the rivers' energy). A dolerite dyke is a common example of a geological control, where its erosive resistance causes a river's energy to be used to widen the valley because erosional incision is impossible; this then creates a setting conducive to the formation of wetlands, because the barrier to downward erosion has an effect similar to that of a dam wall. The loss of energy in a river due to the geological control causes sediment to be lost from the water column, therefore causing aggradation/sedimentation along its path. These areas of sedimentation along a geological control are the areas most likely to be wetlands because sedimentation fills a valley and causes an upstream reduction in gradient (Ellery *et al.,* 2008). As mentioned in section 2.1.2, wetlands are more likely to exist in flat and gently sloping areas than steep-sloping landscapes (Junhua and Wenjun, 2005).

As discussed above, abiotic factors such as climate, geology, and topography are key determinants in wetland form and function (Batzer and Sharitz, 2006), and can explain the distribution of wetlands across South Africa (Table 2.1). Cowan (1995) detailed the wetland regions of South Africa, and in grouping wetland regions into areas found that similar wetlands develop in locations that have similar nutrient regimes, topography and hydrology. Cowan (1995) divided South Africa into different wetland regions based on geomorphological provinces, climate and geology. In establishing the wetland regions, Cowan (1995) found it was important to identify the factors that have an influence on the topography, hydrology and nutrient regimes of wetlands in South Africa. The topography of South Africa was determined by the underlying geology and terrain morphological map; hydrology was determined by a combination of climatic factors that included rainfall, run-off and topography; and the nutrient regime, was determined by climates effect on water chemistry in terms of temperature zones as well as potential evapotranspiration. Four main wetland regions were identified: the plateau wetland group, the mountain wetland region, the coastal slope and rimland wetland region, and the coastal plains. Although the distribution of wetlands in South Africa can be attributed to the differences in geology, drainage (geomorphology) and climate, it is not the intention here to provide a detailed description of South Africa's wetlands but rather to build an understanding of the driving forces behind the formation of wetlands. Readers are directed to studies by Rogers (1997), Dini *et al.* (1998) and Ewart-Smith *et al.* (2006), who have provided comprehensive discussions on the distribution of these wetland systems in southern Africa.

McCarthy and Hancox (2000) list the factors important in determining the distribution of wetlands in the region as follows:

- changes in sea level
- fluvial sedimentation in deltas, alluvial fans and floodplains
- present climate and climate change
- chemical sedimentation from groundwater
- neotectonic activity (motions and deformations of the Earth's crust; which include geological and geomorphological processes)
- vegetational succession and plant-water interactions
- aeolian processes (the wind's ability to shape the landscape), including deflation and dune formation
- hydrogeological factors (factors involved in the interaction between groundwater movement and geology), and
- anthropogenic factors (impacts of humans on nature).

In South Africa, the climate and the geomorphology have resulted in a landscape characterised by certain wetland features (Cowan, 1995: 3):

- small rivers of erratic flow, and very few with a strong perennial flow
- very few well-developed floodplains, but most catchments having vleis
- endorheic pans being a feature of the plateau
- except for one small rockfall lake, there are no other natural inland lakes
- coastal and estuarine lakes at places along the coasts of the southwestern and southern Cape and northern KwaZulu-Natal and
- estuaries (those along the east coast had well-developed mangrove forests).

Table 2.1: Case studies of wetland origin and evolution (Ellery *et al.,* 2008)

Wetland System	Location	Factors that led to its formation
Mfolozi floodplain	Subtropical eastern seaboard, 250 km north of Durban (KwaZulu-Natal)	Geological and climatic history: During the separation of the continents the rifting events led to localised subsidence, which was filled by Zululand Group depositions that slowly aggraded into marine sediments as the ocean filled the expanding basin. Marine transgressions led to periods of erosion and deposition. The major uplift events of over 1000 m exposed marine sediments in the area, and with the dropping sea level resulting from the last ice age, a new erosional phase that carved deep coastal valleys began. The subsequent rise in sea level, following the last ice age, led to the drowning and infilling of previously incised coastal valleys, which is largely responsible for the formation of the Mfolozi floodplain wetland. The sea level still exercises a strong control to the existence of the wetlands in the lower areas of the floodplain. Geomorphological factors: The shape of the upper region of the floodplain causes the Mfolozi River to follow a course that progresses over the course of just 1.15 km from an incised confined valley 950 m wide into a broad floodplain setting 6 km wide. The rapid change in geomorphology results in a reduction of the carrying capacity of the river, creating a large-scale deposition in the form of an alluvial fan. The alluvial fan has created lakes along its boundary where tributaries have been cut off by the fan's progradation.
Nylsvlei	Along the Nyl/Mogalakwena River, Waterberg (North West)	Geomorphological factors: The longitudinal profile of the river shows that the wetland occupies the valley where the longitudinal slope is gradual (0.05%). The gradual profile of the valley is responsible for the origin of the Nyl floodplain and wetland. Although underlying geology may be partly responsible for the gradual slope experienced in the valley, it is mainly due to the extensive alluvial deposits brought in by two tributaries (the Rooisloot River and the Dorps River) that are drowning the Nylsvlei Valley from the southeast, lowering the gradient upstream and leading to wetland formation.
Stillerust wetland	Foothills of the Drakensberg mountains, northwest of Durban (KwaZulu-Natal)	Geological factors: The geological control of this wetland is erosion-resistant dolerite sill outcrops. They are responsible for a step-change in slope (from 4.5% to 0.2%) found in the wetland. Most of South Africa's rivers are in a long-term state of active incision following the continental-scale upliftment 20 million and 5 million years BP, but this incision is impeded on river courses where these dolerite intrusions exist. In the Stillerust wetland the dolerite sills act as new base levels in a longitudinal profile and in these cases the river uses its available energy to laterally plane the less-resistant rocks upstream, which result in wide gently-sloping valleys allowing floodplain wetlands to form.

Although the formation of wetlands is driven by the processes described above, the role of humans can deliberately or accidentally modify many of these driving variables. Although not all human interventions have a negative impact on the wetland, there are those that do lead to wetland degradation and wetland loss. Human disturbance can be categorised into four broad impacts (Kotze and Breen, 1994): impacts that change the flow pattern within the wetland; disturbances of wetland soils; changes in the surface roughness and vegetation cover; and replacement of the natural vegetation. These four points generally affect a wetland's hydroperiod, i.e. the timing, frequency, duration and extent of flooding.

In South Africa, the key pressures on wetlands that contribute to their loss and degradation are cultivation (sugarcane, orchards, wheat, etc.), urban development, dam construction and poor grazing management (Driver *et al.,* 2012). Other on-site pressures include road construction, forestry plantations, dumping of solid waste, burning, mining, and waste treatment.

Prior to wetland conservation in the 1970s, it was common practice to drain wetlands for the production of crops and planted pastures. The draining of wetlands is detrimental to a wetland's hydrology, because effectively the wetland's function of regulating stream flow is bypassed, speeding up the movement of water through the wetland, and thus increasing the erosive power of the water (Kotze and Breen, 1994). In South Africa, the Conservation of Agricultural Resources Act of 1983 now protects wetlands, by prohibiting land-users from cultivating or draining wetlands. In KwaZulu-Natal, wetlands were drained for the production of sugarcane crops, and the coastal areas still bear the scars of this practice in the form of herring-bone drains still visible today (Grenfell *et al.*, 2007). The loss of wetlands that resulted from cultivation practices cannot entirely be reversed.

The condition of the rivers in South Africa can be directly linked to the condition of its wetlands because most of the wetlands in South Africa form part of a drainage/river network (Ellery *et al.*, 2008, Driver *et al.,* 2012). Therefore, off-site causes of wetland degradation include (Driver *et al.*, 2012: 78):

- The change in flow regime of a river system (changes to the amount and timing of flows of freshwater to the wetland, for example as a result of water abstraction, effluent discharge, and dams in the catchment).
- Deterioration of water quality in associated rivers because of various polluting activities in the surrounding catchment.
- Poor grazing management or poor crop production practices in the catchment that result in an increased sediment load being deposited in the wetland.

It is impossible to assess accurately how much wetland area in South Africa has been irreversibly lost due to anthropogenic forces, but analysis has revealed that approximately 45% of the remaining wetlands are in a critically modified condition because of human interaction (Driver *et al.,* 2012). The fact is that the degradation of wetlands results in a loss of their socio-economic functions and services. They are critical, for example, in storing water, regulating water supply, and improving water quality. The formation of wetlands involves the complex interaction of a number of abiotic and biotic factors (section 2.1.2, Figure 2.1) and it is the disruption of these unique interactions that brings about instability in a wetland and ultimately result in its degradation and destruction.

The socio-economic value of wetlands has been well documented, and this knowledge has been an important tool in building a strong momentum in developing strategic ways to protect these ecosystems. Driver *et al.* (2012: 81) state that "the solution to protecting a representative spread of wetland ecosystem types lies in a combination of measures for on-site protection, and measures implemented upstream and in the surrounding catchment to secure the quality, quantity and timing of water upon which the wetland's character and functioning depend".

2.1.4 Legislation and protection of wetlands in South Africa

As a consequence of a growing awareness of the importance of wetlands to society, and in response to their unchecked destruction, there has been considerable progress made over the last 20 years in the protection of wetland areas globally. The recognition of the value of wetlands has resulted in an emphasis on wetland protection in many laws and international agreements (Mitsch and Gosselink, 2000). The Ramsar Convention has been at the forefront of international intergovernmental cooperation on the protection of wetlands and/or wetland conservation. Developed and adopted in 1971 in the town of Ramsar, Iran, this global treaty provides the framework for the international protection of wetlands as habitats for migratory fauna that do not observe international borders and that benefit human populations dependent on wetlands (Mitsch and Gosselink, 2000). The mission of the convention is the conservation and wise use of wetlands through national action and international cooperation, with sustainable development being the overarching goal. Subsequent to its initial objectives for protecting waterfowl, the aims have been broadened to include: preventing the loss of wetlands to preserve their fundamental ecological functions as well as their economic, cultural, recreational, scientific and educational value (Ramsar, 2011).

Countries that have ratified the Ramsar Convention are obliged to formulate and implement their planning to promote the wise use of all wetlands and develop national wetland policies. Member countries are required to designate at least one wetland as a Ramsar site and to establish nature reserves at these and other wetlands. As of October 2012, 163 countries had joined the Ramsar Convention, and there were 2 059 Ramsar wetland sites, totalling an area of approximately 197 million ha worldwide[1]. South Africa has 20 Ramsar sites, totalling an area of approximately 553 000 hectares.

Alongside the Ramsar Convention, many countries have formulated their own legislation and regulations relating to the protection of wetlands against degradation and destruction. “Other mechanisms for wetland protection include acquisition, planning, mitigation, disincentives for conversion of wetlands to other land uses, technical assistance, education, and research” (Votteler and Muir, 1996:57).

South Africa’s strong environmental legislative framework and its status as a contracting party of the Ramsar Convention provide the ideal platform to comply with numerous international conservation and environmental agreements, creating a legal framework for enabling the conservation of important natural assets, both terrestrial and aquatic (Table 2.2). In spite of all the laws related to wetland management, it is important to note that there is no specific national wetland legislation. South Africa’s Department of Water Affairs and Forestry (DWAF) outlines how the legislation applicable to wetlands and wetland rehabilitation arises from various other legal instruments currently in place (DWAF, 2007).

The Ramsar Convention covers all aspects of wetland conservation and wise use, recognising wetlands as ecosystems that are extremely important for biodiversity conservation and for the well-being of human communities (SCBD, 2006:51). As a signatory to the Ramsar Convention, South Africa is obliged to comply with all aspects of the agreement.

[1] According to www.ramsar.org /cda/en/Ramsar-about-parties-parties/main (date accessed 09/10/2012)

Table 2.2: Conventions, legislation, policies and plans directly or indirectly related to the conservation and protection of wetlands in South Africa

International	Contribution
Ramsar Convention (1971)	- Inter-governmental co-operation on the protection and conservation of wetlands. - International cooperation in the wise use of wetlands. - Strong legislative framework to build policies and agreements in South Africa.
Convention on Biological Diversity (1994)	- Led to important South African legislation (National Spatial Biodiversity Assessment, National Biodiversity Strategy and Action Plan (2004), (Driver *et al.*, 2005). - Provides the platform for the rehabilitation of wetlands as stated in its agreement.
South Africa	
Constitution of the Republic of South Africa (Act No. 108 of 1996) (RSA, 1996)	- Creates a duty on the state to conserve and rehabilitate wetlands, because "everyone has the right to have the environment protected, for the benefit of present and future generations."
National Environmental Management Act (Act No. 107 of 1998) (RSA, 1998b)	- Produced the Biodiversity Act (2004) (see below) and the National Environmental Management: Protected Areas Act (Act 57 of 2003). - Important legislation leading to the conservation of terrestrial biodiversity and then freshwater biodiversity conservation. - Emphasised the avoidance and minimisation of disturbance of ecosystems and loss of biological diversity.
National Environmental Management: Biodiversity Act (Act No. 10 of 2004) (RSA, 2004)	- Provides the management and conservation of South Africa's biodiversity within the framework of the National Environmental Management Act 1 998; which protects species and ecosystems that warrant national protection
National Water Act (Act No. 36 of 1998) (RSA, 1998a)	- Legal framework for the effective and sustainable management of our water resources. It defines water resources as rivers, streams, **wetlands**, estuaries and groundwater (RSA, 1998a). - It contains rules about the way that the water resource is protected, used, developed, conserved, managed and controlled in an integrated manner. - Enforced the identification and prioritisation of catchments and wetlands. - Enforced the authorisation of any water use.
Conservation of Agricultural Resources Act (Act 43 of 1983) (RSA, 1983)	- It prevents the unauthorised cultivation of virgin soil, regulates the utilisation and protection of vleis, marshes, water sponges and water courses, and prevents the alteration of flow patterns of runoff water.
National Spatial Biodiversity Assessment (2004, 2011)	- Driving impetus to the endeavour of freshwater conservation through the assessment of South Africa's freshwater ecosystems.
National Biodiversity Strategy and Action Plan (2005)	- Preparation of national biodiversity strategy aimed at mainstreaming planning and activities that can impact (positively and negatively) on biodiversity i.e. wetlands (DEAT, 2006).

Each contracting party of the Convention on Biological Diversity agrees to "rehabilitate and restore degraded ecosystems", "regulate or manage biological resources important for the conservation of biological diversity" and "help develop and maintain necessary legislation ... for the protection of threatened species or populations". The Convention on Biological Diversity observes wetlands as important ecosystems, critical for human well-being, in terms of the goods and services they provide (SCBD, 2006).

Section 24 of the Constitution of the Republic of South Africa (RSA, 1996) states that everyone has the right to a healthy and protected environment. "Since wetlands are essential to ecological health which has, in turn, a direct bearing on human health, this provision imposes an implied mandate on all organs of State to take reasonable steps to ensure wetland health" (Winstanley, 2001: 5). Section 2 of the National Environmental Management Act (NEMA) (RSA, 1998b) states: "Sensitive, vulnerable, highly dynamic or stressed ecosystems, such as coastal shores, estuaries, wetlands, and similar systems require specific attention in management and planning procedures, especially where they are subject to specific human usage and development pressure". NEMA contains provisions that oblige developers and planners to conduct environmental impact assessments (EIAs) where the potential impacts are investigated, including the cumulative effects of the activity and possible alternatives (Winstanley, 2001).

The National Water Act (RSA, 1998a) makes provision for an 'ecological reserve', which is a particular water quality and quantity to be set aside to protect the ecological functioning of aquatic ecosystems, before other water uses can be authorised (RSA, 1998a). The Act provides the ideal framework to address wetland conservation. It identifies wetlands, along with rivers, estuaries, and groundwater, as components of the ecological reserve. The wetland component of the reserve relies on important initiatives such as the development of a national wetland classification system, a national wetland inventory, and freshwater conservation plans. Threats on freshwater ecosystems, and hence the ecological reserve, place pressure on provincial and national conservation authorities to develop important freshwater conservation plans in order to guide policy decisions.

The Conservation of Agricultural Resources Act (RSA, 1983) indirectly facilitates the conservation of wetland areas because of the positive role they play in agricultural activities. It is a regulating act that aims to conserve natural agricultural resources by preventing and combating erosion and the weakening or destruction of water sources (i.e. wetlands) (Winstanley, 2001).

The National Spatial Biodiversity Assessment (NSBA) spatially assesses the state of South Africa's biodiversity by incorporating four fundamental components: terrestrial, freshwater, estuarine and marine environments. This information is useful to policymakers, decision makers and practitioners. The 2011 NSBA was the first version to include the state of wetland ecosystems in South Africa, where the report stated that wetlands are the most threatened ecosystem in South Africa (Driver *et al.*, 2012). This threatened ecosystem status warrants the national protection of wetlands in terms of conservation and management, according to the National Environmental Management: Biodiversity Act (RSA, 2004). The 2011 NSBA will help revise and update key policies and strategies, including the National Biodiversity Strategy and Action Plan (NBSAP). The Department of Environmental Affairs and Tourism (DEAT), now known as the Department of Environmental Affairs (DEA) within the Department of Water and Environment (DWE), explains that "the NBSAP sets out a framework and a plan of action for the conservation and sustainable use of South Africa's biological diversity and the equitable sharing of benefits derived from this use" (DEAT, 2006: 6). The NBSAP is supposed to create a framework by means of which wetlands and other environmental components can be managed.

Globally, a greater focus on water issues came into effect with the 2000 Millennium Declaration and Agenda 21 (where the goal was to halve the proportion of people lacking safe drinking water and basic sanitation) (United Nations, 2010). In 2003 the United Nations General Assembly adopted a resolution proclaiming 2005–2015 as the International Decade for Action 'Water for Life' (United Nations, 2010). The lack of drinking water is a vitally important matter but should not obscure the fact that our freshwater systems are facing growing threats from human activity. DWAF aims to manage healthy wetlands that sustainably provide goods and services to society while maintaining their role in the hydrological cycle.

Other legislation pertaining indirectly to wetland conservation are laws against poor forestry and mining practices (the National Forests Act 84 of 1998 and the Mineral and Petroleum Resources Development Act 28 of 2002), and legislation protecting the environment against destruction caused by urbanisation and infrastructure (the Physical Planning Act 125 of 1991, the Development Facilitation Act 67 of 1995, the South African National Roads Agency Limited and National Roads Act 7 of 1998, and various Environmental Impact Assessment regulations) (Winstanley, 2001). However, government legislation, policies, and programmes are worthless in protecting wetlands if the regulations are not enforced. Perhaps a combination of educating the public on the benefits of wetlands and enforcement of wetland legislation/policies would be a better approach. If the public does not recognise the benefits of wetland preservation, wetlands will not be preserved (Votteler and Muir, 1996).

2.1.5 Conservation of wetlands

Conservation and biodiversity planning in South Africa are firmly embedded in policy and legislation. However, according to Rivers-Moore and Goodman (2010), conservation planning for freshwater systems (in which wetlands play an important role) has lagged behind terrestrial conservation by at least a decade, worldwide. The 2004 National Spatial Biodiversity Assessment estimated that over 50% of South African freshwater ecosystems associated with main rivers are critically endangered and therefore recommended that there be added impetus to the endeavours related to freshwater conservation (Driver *et al.*, 2005). More recently, with regard to wetlands, "a disturbing 65% of wetland ecosystem types in South Africa are threatened (48% critically endangered, 12% endangered and 5% vulnerable), making wetlands the most threatened of all ecosystems" (Driver *et al.*, 2012: 7). As mentioned before, it is impossible to map the historical extent of the wetland area over South Africa, but it is estimated that 50% of the original wetland area has been lost, with approximately 300 000 wetlands remaining, making up only 2.4% of South Africa's surface area (Driver *et al.*, 2012).

While wetland conservation in South Africa has mostly taken place in nature reserves and at Ramsar sites, most wetland areas are found outside these protected areas, hence there is little conservation of the majority of wetlands in South Africa. The 2011 National Biodiversity Assessment stated that only 11% of wetland ecosystem types are well protected (within national parks or nature reserves), and 71% are not protected at all (Driver *et al.*, 2012).

South Africa addresses the conservation of wetlands through the formulation of conservation plans, more specifically freshwater ecosystem conservation plans. Conservation planning is defined as a process where areas are conserved in order to protect biodiversity, including ecosystems, biological assemblages, species and populations (Margules and Pressey, 2000). Conservation planning takes into account the location of the reserve (area) in relation to the natural and biological patterns, and includes variables of reserve design such as size, connectivity, replication and the alignment of boundaries (e.g. watersheds) (Margules and Pressey, 2000). South Africa is now at the forefront of systematic biodiversity planning approaches and has (for more than a decade) been exploring the potential use of systematic biodiversity planning for freshwater ecosystems (Nel *et al.,* 2011a). Nel *et al.* (2011a: 7) stated that "given the connectivity of freshwater ecosystems, a focus on representing biodiversity in isolated areas, without regard for upstream, downstream or upland areas, is conceptually flawed".

There is a large spatial component involved in conservation planning because of the challenge in relating wetlands with upstream drainage networks, the surrounding riparian zones, downstream reaches, and the influence of multiple human stakeholders (Dudgeon *et al.,* 2006). There are thousands of wetlands in South Africa, many of which are in need of management intervention. Planning is essential in the management of wetlands, and mapping is a useful planning tool. Hence, GIS has facilitated great advances in not only freshwater ecosystem conservation planning but conservation planning in general. Conservation plans require the integration of a large amount of spatial information related to environmental processes and biodiversity spatial distributions as they occur in the landscape. The applicability and reliability of a conservation plan is therefore largely reliant on the quality and quantity of information used to derive the plan. Spatial analysis using GIS has facilitated the development of key spatial coverages, such as river eco-regions, geomorphic provinces, longitudinal zonations of rivers, estimated ecological condition of major river systems and national land cover maps. These products have been essential in the advancement of freshwater conservation planning, highlighting the importance of GIS methods in improving conservation efforts.

The National Freshwater Ecosystem Priority Areas (NFEPA) project has been an important product derived from GIS applications. GIS was used to graphically define strategic spatial priorities for conserving South Africa's freshwater ecosystems and supporting sustainable use of water resources (Nel *et al.,* 2011a). The NFEPA project has taken careful consideration of where wetlands occur in South Africa and has noted the importance of conserving them. The NFEPA project "provides guidance on how many rivers, wetlands and estuaries, and which ones, should remain in a natural condition. It supports the implementation of the National Water Act, the Biodiversity Act and the Protected Areas Act" (Nel *et al.,* 2011a: 1). The maps produced as part of the NFEPA project were developed using principles of systematic biodiversity planning, also known as systematic conservation planning (Margules and Pressey, 2000). All the maps produced as part of the NFEPA project form part of a comprehensive approach to the sustainable and equitable development of South Africa's scarce water resources. The NFEPA project has addressed the challenges of freshwater conservation by aiming to identify a national network of freshwater conservation areas, which can then inform management and conservation plans.

The NFEPA project identifies Freshwater Ecosystem Priority Areas (FEPAs) based on a range of criteria (wetland delineations, wetland clusters, wetland condition and type, river types, landforms, estuaries, fish sanctuaries, water yield areas, and groundwater recharge areas). There is considerable detail regarding how the range of criteria are technically integrated to identify FEPAs (Nel *et al.*, 2011b), but this review simply highlights the fact that the identification of wetlands in terms of locality, type and condition, forms a vital spatial component in identifying FEPAs. For example, the identification of wetland clusters (groups of wetlands embedded in a relatively natural landscape) is vital in identifying important groups that allow for ecological processes such as migrations of frogs and insects between wetlands. In many areas of South Africa, these clusters no longer exist due to human-induced landscape fragmentation (Nel *et al.*, 2011a). The identification of FEPAs is vitally important in the conservation of wetlands and the identification of wetlands is vitally important in the identification of FEPAs.

"The biggest challenge in wetland conservation is that South Africa lacks a comprehensive overview of the extent, diversity, distribution, status and relative importance of its wetlands" (Rountree *et al.*, 2009:17). Other challenges facing wetland conservation include: generating public awareness of the importance and value of wetlands; finding and training people to manage wetlands outside protected areas; and the lack of co-operation from non-governmental organisations (NGOs), government departments, land owners and the public (Nel, not dated). There is limited knowledge of where these wetlands occur, and therefore the locations for intervention for wetland conservation and rehabilitation are unknown. The degradation of wetlands cannot continue at the expense of the goods and services that wetlands can provide to society. Freshwater biodiversity conservation remains a conservation challenge because it is influenced by upstream drainage networks, the surrounding land, riparian zones, downstream reaches, and is subject to ongoing competition from multiple human stakeholders (Dudgeon *et al.*, 2006).

In the last decade progress has been made in the development of a South African National Wetland Inventory (NWI), a national wetland classification system, and a preliminary set of 791 wetland ecosystem types. The progress made in mapping wetlands for the NWI has produced valuable information with regard to the wetland ecosystem status, and this was reported on for the first time in the 2011 National Biodiversity Assessment (Driver *et al.*, 2012). The consultation process between several national government departments and national agencies has resulted in important national biodiversity targets being set and politically accepted (accepted by government as a new measure) in South Africa regarding freshwater ecosystems. Conservation targets are set on the known information regarding freshwater ecosystems, placing emphasis on the need to determine the full extent of our ecosystems before developing realistic targets. Based on these targets, South Africa has agreed to maintain at least 20% of each major freshwater ecosystem type in South Africa in a good condition (Nel *et al.*, 2011a). Again, emphasis is placed on aiming to improve our wetland inventory so that the stated conservation targets are set on all wetland ecosystems and not only those partially identified in an incomplete wetland inventory.

2.2 Mapping wetlands

The value in identifying the sizes, locations and types of wetlands in terms of conservation and protection of freshwater ecosystems lies in providing great input into the prioritisation, planning, and rehabilitation needed for the conservation of freshwater ecosystems globally. As seen in the NFEPA project, a wetland inventory forms the baseline spatial data and could be used for many purposes, including comprehensive resource management plans, environmental impact assessments, natural resource inventories, habitat surveys, and the trend analysis of trends in wetland status (Wilen *et al.*, 2002). Begg (1986: 48) stated that "one cannot hope to formulate a policy for the utilization (or management) of wetlands without knowing where these areas lie, or how much of each catchment comprises of wetlands".

There is an array of different techniques used to map wetlands. The choice of technique is dependent on the scale at which the wetlands are being mapped. The most accurate technique to map wetlands is the field survey technique, where the researcher literally walks the perimeter of the wetland with a GPS or scaled field map, plotting numerous points along the walk and then logging them into a geographic information system. This traditional technique is costly and time consuming, and is not feasible for mapping wetlands over large landscape areas. The introduction of aerial photographs and launching of satellites have resulted in rapid advancements in wetland mapping techniques over the past century. Wetlands exhibit distinct light-reflectance characteristics in the visible or infrared portions of the electromagnetic spectrum. Wetland soils and water have distinctive reflectance characteristics that can be used to identify the presence of a wetland (Lyon, 1993: 72). These characteristics are found in both aerial photographs and satellite imagery. With satellite imagery, large spatial domains can be rapidly mapped through various classification methods; therefore, the delineation of wetlands from satellite imagery is an internationally accepted method of creating regional inventories of wetlands (Thompson, 1994). This section reviews techniques used to map wetlands both internationally and in southern Africa.

2.2.1 International wetland mapping approaches

Techniques to map wetlands using satellite imagery are extensive and well reviewed (see Kulawardhana *et al.,* 2007; Ozesmi and Bauer, 2002). Currently, methods and approaches to map wetlands using satellite imagery range from supervised, to semi-automated, to unsupervised approaches in a range of temporal, spectral and spatial resolutions. Li and Chen (2005) contend that integrating spatial environmental data with satellite imagery has improved the accuracy of mapping wetlands. Using rule-based approaches where classification is based on spatial data themes (i.e. land cover, soil texture, terrain) in addition to satellite imagery, has raised the overall classification accuracy from 69% (traditional supervised classifier) to 83%. Ozesmi and Bauer (2002) summarised the literature on satellite remote sensing of wetlands and assessed the success of various methods in classifying wetlands and distinguishing them from other land cover classes. The main conclusions drawn from their review are as follows:

1. Classification of satellite imagery may not match the information gained through field surveys but can provide complementary information on wetlands over a large survey area. It can identify areas where change is occurring and more detailed information is needed.

2. Rule-based classification methods generally provide better results than conventional statistical classification methods, often because of their use of ancillary data.

3. Multi-temporal imagery and ancillary data allow for the highest accuracy in wetland identification and discrimination from other land cover types.

4. Wetlands should be separated from other land cover types prior to their classification using ancillary data.

There are a number of different automated and semi-automated approaches to mapping wetlands using satellite imagery, and Kulawardhana *et al.* (2007) evaluated these methods. The methods included the tasselled cap wetness index (TCWI), normalised difference water index (NDWI), multi-band vegetation indices (MBVI), two band vegetation indices (TBVIs), normalised difference vegetation index (NDVI), and data fusion involving Enhanced Thematic Mapper (ETM+) and SRTM (Shuttle Radar Topography Mission) data and then classifying the same. In this review Kulawardhana *et al.* (2007) concluded that many of the automated methods resulted in very low accuracies in the wetlands that were delineated and/or very high errors, and are therefore inappropriate in mapping wetlands at larger spatial scales. Automated approaches are unsupervised techniques that rely purely on spectrally pixel-based statistics and incorporate no prior knowledge of the characteristics of the themes being studied (Xie *et al.*, 2008). This study was done using single-date imagery, and multi-temporal imagery may have boosted the overall accuracy. However, the major limitation in the automated process was the spectral similarity between the vegetation canopies of wetlands and the vegetation or agricultural crop cover found outside wetlands. Semi-automated approaches on the other hand include the enhancement of satellite imagery to obtain a better contrast between wetland and non-wetland land cover types, digitising, and using ancillary data to supplement Enhanced Thematic imagery. Semi-automated approaches involve supervised classification methods where the established classification is learnt from a training dataset, which contains predictor variables measured in each sampling unit and assigns prior classes to the sampling units. Kulawardhana *et al.* (2007) stated that semi-automated approaches increased the accuracy by at least 30% in comparison with purely automated approaches when mapping wetlands.

Knight *et al.* (2009) used Landsat TM and Enhanced Thematic Mapper (ETM+) imagery over a 16-year period to develop comprehensive wetland maps for the state of Queensland, Australia. The approach was a classification method called the Standing Water Body (SWB) method. The SWB method separated the main spectral and land cover elements of wetlands (vegetation, standing water, and shadow cast by vegetation and topographic relief) and used rules to combine spectral classes to provide multi-temporal information for mapping of wetlands. They concluded that the SWB method requires enhancement through the inclusion of an NDWI and ancillary data such as vegetation mapping and drainage networks to improve the accuracy in mapping wetlands even further. Landmann *et al.* (2010) mapped wetlands over a wide area in semi-arid Africa using 250-metre Moderate-resolution Imaging Spectroradiometer (MODIS) metrics and topographic variables by using a multi-temporal approach. They used ancillary data to supplement the multi-temporal imagery, and this included data such as sinks and streamline areas. The ancillary data was used to mask potential wetland areas, minimising spectral confusion. This study method of mapping wetlands over a large spatial area was largely successful, producing the most detailed wide area wetland dataset available for West Africa. Landmann *et al.* (2010:1762) mention that this study "demonstrates the need for earth observation methods to not only base their assessments on satellite observations, but also utilise linkages to other higher resolution *in situ* point data sets". The resolution of MODIS imagery was problematic and resulted in inaccuracies in pixel variability and mapping errors, and it was challenging to assess even using high-resolution datasets.

The success of these wetland mapping techniques is a result of a combination of approaches that are simple, statistically significant, or highly complex. The combination of high spatial and temporal resolution imagery is essential in mapping wetland ecosystems, but such single satellite imagery is currently unavailable (Ryo *et al.,* 2012). To address this need, Ryo *et al.* (2012) researched the applicability of multiple end member spectral mixture analysis (MESMA) using single bi-sensor imagery with Landsat Thematic Mapper and MODIS data, across different orbiting periods. This complex research approach demonstrated how MESMA can be applied for multi-scale mapping of wetland ecosystems; however, spectral similarity between dark water and shade still affected the agreement in land cover classes (spectral confusion).

2.2.2 Southern African wetland mapping approaches

The climate over southern Africa is extremely variable, and therefore saturated soil is often an unreliable indicator of wetland conditions or boundaries, as during certain years wetlands may be much larger or wetter than during others. Using field surveys when mapping wetlands to confirm this variability would be ideal, but it is expensive and time consuming. It is clear that at a national scale the mapping of wetlands and the creation of an NWI is no simple task; however, these processes are of critical importance in the conservation and management of wetland systems. An NWI is important for identifying where wetlands are located, in order to prioritise sites according to the functions and values of each wetland site, including ecological, social and cultural values. The NWI provides a tool to inform the planning and management related to wetlands at all levels, and is a baseline for measuring future changes in a wetland's area, function and value (DEAT, 1997).

There is a wealth of site-specific reports on wetland location, extent and type across southern Africa. These are derived using a range of techniques, and need to be collated for strategic planning purposes. DEAT recognised the need for an inventory of South Africa's wetlands. The task of developing the South African NWI took place in the early 2000s, where the aim was to map the extent, location and characteristics of wetlands, partly adopting the approach used by Thompson *et al.* (2002). To date, four versions have been released. Their extensive use in numerous projects since the first version has led to the identification of the shortcomings of the NWI, and hence to rapid improvements in each version. Mbona (2011) explained that the first version of the National Wetland Map (NWM) was derived from the 2000 National Land Cover data, using areas described as 'water bodies' and 'wetlands'. In the second version, data on dams, lakes and rivers supplied by DWAF were added to the wetland map. The third version of the NWM (2009) included 1:50 000 inland water features, and then features in the layer were divided into wetlands, natural water bodies and artificial water bodies. Other wetland delineation products derived from various biodiversity planning initiatives were also included in the third version of the NWM. The final step was to classify the different wetland types within the NWM according to the National Wetland Classification System. The current NWI has information on the locality, type and extent of a wetland, as well as on various important attributes, including the wetland condition, the rank of importance and the Freshwater Ecosystem Priority Area (FEPA) status.

The NWI project is a 'living document' that continues to collect data and improve the mapping of wetlands and the attributes associated with each wetland (Mbona, 2011). NWIs are important in biodiversity planning initiatives and in freshwater conservation because they represent important information about individual wetland areas. A potential shortcoming of the South African NWI is that smaller wetland areas are not identified; therefore, the NWI is sometimes criticised for having gaps in the NWI layer and questions are raised about whether it is a valuable planning and management tool. Each province in South Africa strives to improve the quantity and quality of their wetland mapping by continuously adding to the NWI; however, some provinces have more resources than others, which aids them in continually updating their wetland inventory, and the result is a fragmentary, incomplete NWI. At South Africa's National Wetlands Indaba in 2012, it was stated that "the days of mapping wetlands at a national scale are over because of the enormity of the task and the general lack of resources to complete the aim and objectives of the NWI" (Dini, 2012 pers. comm.). It is difficult to map wetlands at a national scale, and the strategy being used to address this problem with the NWI is to encourage collaboration between individual wetland mapping initiatives at regional scales, and incorporation of the information into the NWI.

In the past, a number of wetland mapping techniques have been developed by a variety of local (southern African) organisations; there is, however, a general lack of published material in mainstream/peer reviewed publications on the methods and approaces used (Scott-Shaw and Escott, 2011; DLA-CDSM, 2006; Nel *et al.*, 2011a). Often wetland mapping approaches are incidental in nature, and are developed as part of completing an overarching research aim, with some researchers only interested in publishing the final result as opposed to working with or developing a valuable wetland mapping technique that could prove valuable to the GIS and the remote-sensing research community (Escott, 2012 pers. comm). Documenting and publishing these previous wetland mapping attempts and approaches could help to improve not only the current wetland maps in South Africa, but also the conservation initiatives that are informed by the knowledge of wetlands' current extent and location. Publishing this research on wetland mapping would also prevent researchers from 'reinventing the wheel,' as they would be aware of what has been done and what has been successful in the realm of wetland mapping.

A well-documented wetland mapping study compiled by Thompson *et al.* (2002) proposed a methodology for a South African NWI. This report was prepared for the South African Wetlands Conservation Programme, and described a methodological approach of mapping wetlands at a national scale. The aim of the study was to provide strategic wetland data on a repeatable, long-term and operationally sound basis, by means of a cost-benefit analysis that had a desired degree of accuracy. The wetland modelling and mapping done by Thompson *et al.* (2002) was completed at a national scale using a terrain-based hydrological model derived from a 20 m Digital Elevation Model (DEM) and multi-temporal satellite imagery. The terrain-based hydrological model was used to clip out areas that were highly likely to contain wetland areas, and those areas were then subjected to image classification. Using the reflectance NDVI values, the Winter NDVI values were subtracted from the Summer NDVI values, and the areas with the biggest difference were likely to be wetlands. Using the depression layer (DEM-derived) and the NDVI layer at different weightings, the overlaid areas would represent the wetland area in the final map.

The approach by Thompson *et al.* (2002) is inspiring in the way that a terrain-based hydrological model was used to limit the classification of satellite imagery only to areas where wetlands are highly likely to occur. However, a concern with this approach was that "non-wetland areas are defined as land cover types either within which wetlands would not occur (i.e. urban, mines, plantations, woodlots, and cultivated areas), or could not be identified, even if they existed, using remote imagery" (Thompson *et al.*, 2002: 10). According to DEAT's 2007 report, "South Africa Environment Outlook: A Report on the State of the Environment," cultivated and forested areas, urban areas and mines amount to 16.53 million hectares of land cover in South Africa, and a large percentage of that area could be excluded in the classification process of identifying wetlands (DEAT, 2007). However, the question is whether it is possible to avoid this exclusion when it comes to the classification of satellite imagery.

Some of the earliest local studies on mapping wetlands were conducted by Begg (1986; 1988; 1989), with the aim of starting a comprehensive wetland inventory of the wetlands in what was then Natal province (now KwaZulu-Natal) in South Africa. There were four phases to Begg's project:

1. Synthesise the available information on the distribution, structure, functioning, use and value of wetlands
2. Conduct an inventory of wetlands in the Mfolozi catchment
3. Describe the location, status and function of priority wetlands in Natal
4. Formulate a wetland policy statement for the management of wetlands.

Phases 2 and 3, where Begg (1988) describes the mapping approach used to identify wetlands in the Mfolozi catchment (phase 2), and locates priority wetlands in Natal (phase 3), are particularly relevant for this research. In phase 2, Begg (1988) used aerial photo interpretation, utilising the black-and-white aerial photographs available from the offices of the Surveyor-General, to map and identify wetlands. This approach involved searching aerial photographs for any signs (geology, vegetation, relief, position in landscape, soil types, etc.) that are characteristic of wetlands. Using a mirror stereoscope and interpretation experience, wetlands in the Mfolozi catchment were outlined on the aerial photographs based on areas that reveal the tone and relief characteristic of a wetland. Lastly the manual delineations were ground-truthed and transferred from the photographs to 1:50 000 maps, resulting in the first complete wetland inventory of a catchment in Natal. In phase 3, Begg (1989) used the same aerial photo interpretation method as in phase 2, however added wetland sites identified by the Surveyor-General on the standard 1:50 000 topocadastral maps, and general information about wetland areas. The aim of phase 3 was to locate and gather information on priority wetlands in the catchments of the major rivers influencing the Natal region. The aerial photo interpretation method is a tedious approach to locating and mapping wetlands, and therefore phase 3 excluded catchments smaller than 1 000 km^2, coastal wetlands, and wetlands smaller than 100 ha. The study by Begg formed the baseline for mapping the distribution of wetlands in KwaZulu-Natal, and is valuable in evaluating the conservation status, function and value of the wetlands described.

In a study by Dely *et al.* (1995), three wetland mapping techniques were evaluated in the process of preparing an inventory of wetlands in the Natal Drakensberg Park (KwaZulu-Natal, South Africa): aerial photograph and orthophotograph interpretation, satellite imagery interpretation, and topographical analysis. In the aerial and orthophotograph interpretation, possible wetlands were identified on the basis of observing colour, tonal and textual ranges, in a similar manner to Begg (1988, 1989). The disadvantage of this approach was that the irregular exposure and print quality of the photographs, which detracted from their application, and the general poor quality of the photograph, restricted the mapping of smaller wetlands. When compared to the use of aerial imagery, the ability to accurately identify wetlands using satellite imagery is significantly lower, because clouds, cloud shadow and terrain shadows cause confusion with the spectral signatures of other land cover types. In the topographical analysis the conclusion was drawn that permanent wetland areas are characterised by low slope angles of between 0˚ and 3˚, while seasonal and temporary wetland areas occur in areas that slope between 3˚ and 5˚. The results of this study were dependent on the imagery that was available prior to 1999; therefore, the advancements in satellite data resolution today will improve the mapping results drastically.

In another documented study on wetland mapping conducted by Pillay (2001), wetlands in the Midmar catchment in KwaZulu-Natal, South Africa were mapped using satellite imagery. The aim was to develop a methodology for the accurate and efficient delineation of wetland areas using satellite imagery. The methodology involved analysing summer versus winter Landsat ETM imagery, comparing supervised versus unsupervised versus level-slicing classification approaches (which were then tested against existing verified wetland datasets), and some ground-truthing. The finer classification techniques resulted in higher accuracy portrayals of the wetland layer. The study concluded that winter imagery produced spectral confusion (a common limitation in the classification of satellite imagery) caused by the presence of burnt areas, firebreaks and hill shade, and was therefore unsuccessful. There were three main recommendations from this study: that summer imagery taken during a high rainfall period be used to achieve greater accuracy in mapping wetlands; that the use of high-resolution satellite sensors would dramatically increase the accuracy of wetland mapping; and that wetland mapping studies and monitoring should constitute ‘living documents’ in which constant evaluation and improvements are sought. An important statement made by Pillay (2001) is that most of the inaccuracy found in the resultant wetland layer can be attributed to a change or modification in land cover, as there seems to be an overall loss of wetland areas. This emphasises the need to account for the inaccuracies in wetland identification and to develop a model to help detect probable wetland areas, irrespective of whether the land cover is now modified or not.

2.3 Addressing mapping limitations using probability modelling

Exploring new methods and approaches to identifying and mapping wetlands will result in an improved version of the NWI. Reviewing literature on the different wetland mapping techniques has made it clear that there is no straightforward method, approach or technique to map wetlands accurately. Multi-temporal classification is the popular choice for mapping wetlands using satellite imagery, but limitations in resolution, spectral similarity and spectral mixing seem to limit the success of these approaches. Kulawardhana *et al.* (2007) evaluated a number of wetland mapping approaches using satellite imagery, but it is evident that in many unsupervised classification approaches, spectral confusion results in unavoidable inaccuracies and errors. Although multi-temporal satellite imagery classification in wetland mapping is largely successful in many studies, there are notable errors. The use of ancillary data has been recognised (Ricchetti, 2000; Kulawardhana *et al.,* 2007) as a means of addressing these shortfalls.

Ancillary data can cover a range of approaches, including satellite classification, probability modelling, image analysis and populating metadata. This study explores the idea of using available data (ancillary) to create a probability model of where wetlands are likely to occur. There is general uncertainty as to where wetlands occur in the landscape because they can be found in different climatic, hydrologic, geomorphic, and physiochemical environments. Combining available ancillary data in a Bayesian statistical environment could possibly address this uncertainty. In the sections below, the use of probability models and Bayesian statistics will be discussed. A product that can identify areas highly likely to be wetlands can improve the classification of satellite imagery when mapping wetlands, minimising the inaccuracies that result due to misclassification of pixels, or spectral confusion, or minimising the potential area where wetlands are likely to occur.

2.3.1 Probability modelling

The method used in this study uses the idea of probability theory in its modelling approach, but more directly the Bayesian probability theory. It is therefore important to clarify the theory of probability, and then provide insight on Bayesian probabilities.

What is probability?

Probability is derived from the word ‘probable’, which is defined as “likely to be the case or to happen” or “likely but also plausible” (Oxford English Dictionary, 2002). A level of probability is therefore the answer to the following question: how likely is it that an event will happen? The higher the likelihood of something occurring, the higher the probability number. Usually we find that probability is expressed as a value between 0 and 1, where the probability of something happening is 1, and the probability of nothing happening is 0 (Box 1). In this study it can be said that the probability of definitely finding a wetland in a certain location is 1, and the probability of definitely not finding a wetland in a certain location is 0. So therefore in this case probability is the likelihood that a wetland will occur at a sample point. The uncertainty in the probability of wetland occurrence then arises with the environmental complexity in which wetlands are found. Wetlands occur in various environmental conditions; therefore, it is difficult to define a single probability related to the occurrence of a wetland. The traditional probability theory does not account for this uncertainty.

Box 1: Understanding the theory of probability

A popular example is the experiment of flipping a coin. There are two possible outcomes when flipping a coin: heads or tails. The sum of all probabilities must equal 1, and the probability of getting a head is equal to getting a tails; therefore, the probability of getting heads or tails equals 0.5.

Another example is rolling a die, with the probability related to the number of sides of the die. Since a die has six sides, the sum of probabilities of all sample points must equal 1. Therefore the probability of the die rolling onto a specific number on the die is 0.167 or ⅙.

Probability modelling examples

Probability models use the probability theorem to model occurrence and the likelihood of something or an event occurring given known frequency distributions. Probability modelling is commonly used in a variety of different fields, including environmental science, and has been used in the identification of wetland areas. In most cases, this is a prerequisite process or step in the method and approach of mapping wetlands. Thompson *et al.* (2002) used a probability model prior to satellite image classification to determine areas where wetlands are not likely to occur based on topographic characteristics. This then excluded all image-derived wetlands mistakenly identified in landscape areas not capable of hosting a wetland, which results from spectral confusion (fire breaks, hill shade and open water). Using differently weighted topographic components (slope, flow direction, flow accumulation, depressions), Thompson *et al.* (2002) then created a Landscape Wetness Potential (LWP) model that identified areas where water is likely to accumulate, indicating probable wetland areas. This weighted model spatially represented areas where there is a low to high potential of wetness. A final probability table was used to combine the LWP classes and the satellite image-derived wetland potential classes. The 'high' probability of wetland occurrence was associated with a 'high' image-derived wetland potential class and 'high' probable LWP class.

2.3.2 Bayesian probability

Thomas Bayes and Pierre-Simon de Laplace were responsible for formulating the Bayesian probability theory, a theoretical and practical framework developed in the eighteenth and nineteenth centuries. Bayesian probability borrows most of its theory from the traditional statistical probability theory, but is seen as a systematic way of approaching a problem in which one is confronted with incomplete knowledge about the problem (i.e. uncertainty) (Bruyninckx, 2002).

Bayesian probability theory differs from classic probability theory by updating probabilities as new information is observed (Box 2). Classic probability is based on frequency distributions, as is defined by the number of successful events out of trials observed. The competing view of probability is called 'subjective' and is often associated with the phrase 'degree of belief' (Gill, 2002). In Bayesian statistics the user objectifies the analysis. That means prior information in terms of a researcher's experience, intuition, and theoretical ideas are included in a Bayesian framework, and this information can be subjective; but this Bayesian prior information provides the reader with a specific, formalised statement of currently assumed knowledge in probabilistic terms (Gill, 2002). Bayesian probability makes use of evidence or existing knowledge to account for uncertainty. Adapted from Boone (2004), Bayes' theorem is usually expressed in this formula:

$$p(A \mid B) = p(B \mid A)\, p(A) / p(B)$$

where $p(A \mid B)$ is the probability of finding observation A, given that some piece of evidence B is present.

$p(B \mid A)$ is the probability of the evidence turning up, given that the outcome obtains. $p(A)$ is the probability of the outcome occurring, without knowledge of the new evidence, where $p(B)$ is the probability of the evidence arising without regard for the outcome. The presentation in the form of probabilities gives an explicit representation of uncertainty (Kragt, 2009).

Box 2: Understanding Bayesian probability

Imagine you are betting on a horse race (the example provided by Boone, 2004). Spike the racehorse has won 6 out of 10 races. Dreamer the racehorse has won 4 out of 10 races. Using traditional probability theory, you will bet on Spike because he has won 60% of his races and Dreamer only 40% of his races. Then you include another variable that may influence the outcome of the next race: the weather. You discover that Dreamer has won all his races in rainy weather, where Spike has won only 1 of his 6 races in rainy weather. The horse race on which you are betting is going to take place in the pouring rain; therefore, given the conditional probability theory, your bet would be safer with the horse Dreamer, who performs better in rainy weather than Spike, although Spike has won more previous races.

What if Spike won three out of six races in the rain? This makes the bet less certain. Using Bayes' theorem one can work out the probability that Spike will win, with the evidence that it will be raining, which is $p(A \mid B)$.

$p(A \mid B) = p(B \mid A)\ p(A) / p(B)$

$p(B \mid A)$ is the probability that it is raining given that Spike wins. Since it was raining three out of seven times, $p(B \mid A)$ is 3/7, or 0.43

p(A) is the number of times Spike won without the knowledge of any new evidence. Spike won 6 out of ten races and therefore p(A) is 6/10, or 0.60

p(B) is the probability of rain, without regard for which horse won the race. Since we know it rained on seven days, this value is 7/10, or 0.70

Therefore the probability that Spike wins the race given the evidence:

= 0.43 * 0.60 / 0.70

= 0.37 or a 37 % chance that Spike will win the race.

Therefore your bet is still safer with the horse Dreamer.

The general perception or belief is that wetlands are highly likely to occur on flat slopes, but what is the likelihood of wetland occurrence on flat slopes that vary in rainfall, evaporation rates, or differ in soil characteristics or geology? Bayesian probability theory accounts for this uncertainty by using existing or known information to generate conditional probabilities derived through a mathematical formula based on Bayes’ theorem. Conditional probabilities are numeric representatives given to reasoning or belief around the uncertainty.

Bayesian probability modelling has gained increasing recognition in the environmental sciences field. Lee *et al*. (2002) used probability modelling in landslide susceptibility analysis and verification in Korea, where a weight-of-evidence method (a Bayesian probability model) was used to evaluate the landslide susceptibility using GIS data. By using topographic factors such as soil, forest and land use, and other landslide-related factors derived from a dataset, the weight-of-evidence method calculated the probability of a landslide occurring in certain areas (landslide susceptibility index). In this study various combinations were performed with various factors to work out which combination was best in indicating the landslide susceptibility index after the validation process. Lee *et al*. (2002) found that the combination of slope, aspect curvature, soil drainage and wood type produced the most accurate landslide susceptibility index, with an accuracy greater than when all factors or variables were used. He *et al.* (2007) used Bayesian networks and GIS to map pre-European settlement vegetation, and found that the model could predict the historical vegetation distribution and is robust at multiple classification levels. In this particular study, the model provided a quantitative and spatial basis for the restoration and mapping of natural floodplain vegetation. Bayesian networks are probabilistic graphical models that represent a set of random variables and their conditional dependencies (Aguilera *et al.*, 2011) and have been used in numerous studies from medical research to air quality prediction. Bayesian networks are useful in situations of uncertainty where there are multiple conditional causes of an observed effect (e.g. a wetland within the landscape). Liu *et al.* (2008) used a Bayesian model for urban air quality prediction under uncertainty. Hooten *et al.* (2003) used Bayesian models in predicting the spatial distribution of ground flora on large domains. The study initially found that the occurrence of certain plants is correlated with several site-defining variables, and their aim was to create a robust methodology for modelling natural processes on a landscape. The study indicated that within the spatial structure, in addition to environmental effects, the distribution of species is influenced by uncharacterised spatial random effects. Bayesian models allow for the successful integration of these effects. Bayesian conditional probabilities could potentially account for the uncertainty of where the wetlands occur spatially.

Bayesian networks and probabilities

Bayesian networks are not a household name in the environmental research field yet. They are gaining popularity in the field, however, and are likely to establish their position as one of the standard methods of analysis, especially in problems dominated by uncertainty. Therefore, getting to know them at least on a general level will be beneficial to all environmental scientists (Uusitalo, 2007). Bayesian networks rely on Bayes' theorem of probability theory to propagate information between variables. Within a Bayesian network, the Bayesian theory provides an algorithm in which the likelihood of something happening is calculated in the face of some particular piece, or pieces of evidence. By providing evidence into the likelihood of something happening, a probability is derived that has accounted for the uncertainty by including known knowledge.

Bayesian networks are constructed as a diagram and careful consideration is required of how the system works, which is essential in the network-building process (Figure 2.2). The Bayesian network is an important tool that can help decision makers balance the desirability of an outcome against the chance that the option selected may fail to achieve it.

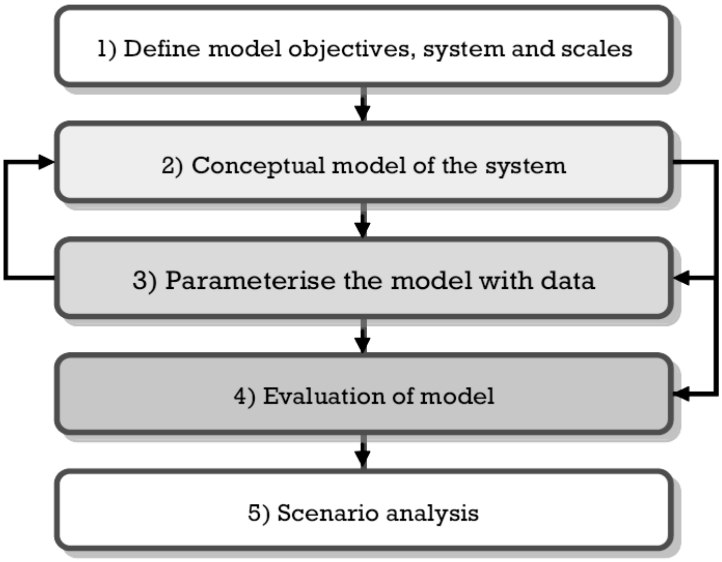


Figure 2.2: The major steps in developing a Bayesian network (adapted from Kragt, 2009).

A Bayesian network diagram (Figure 2.3) represents the variables through nodes, each with a finite set of mutually exclusive states. Cain (2001) explained that the links between these nodes represent their causal relationship, and each node has a set of probabilities, specifying the belief that a node will be in a particular state given the states of those nodes that affect it directly. It is important to note that each variable (node) has a finite set of mutually exclusive states (in Figure 2.3 it is states of high, medium and low).

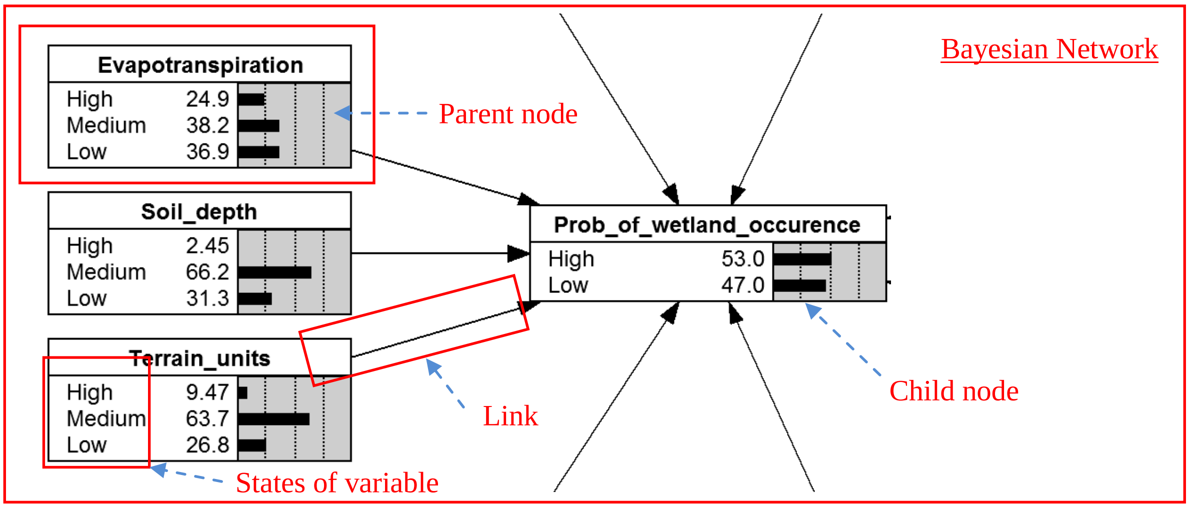


Figure 2.3: Basic components of a Bayesian network: parent nodes linked to a child node and the states found within a node.

Bayesian networks were originally developed to account for the uncertainty found in management systems in the decision-making process (Cain, 2001). In Bayesian statistics, the unknown parameters are treated as random variables, and their distributions are derived from known information. In simpler terms, a Bayesian model is a tool that can be used to build decision-support systems (DSS). Basically "the model can be developed simply to provide a mathematically optimal decision on the basis of the information provided to the model, or can be used in a way that promotes an improved understanding of the environmental system, leaving the decision makers to reach their own conclusions on the basis of that understanding" (Cain, 2001:06).

2.4 Conclusion

It is imperative that people realise the important role that wetlands play in society today, with regard to the goods and services they provide. The provisioning, regulating, supporting and cultural services wetlands provide are of key significance to water availability in the landscape. Currently the degradation of these systems has led to the limiting of ecological contributions and services that a wetland can provide. Preservation and conservation of these systems can only be achieved through a legislative framework and through international cooperation with the unified aim of addressing the loss of wetlands. International and local legislation, policies and plans form a framework to protect and conserve our wetlands; however, this framework is currently insufficient to endure and combat the current degradation of freshwater ecosystems. There is a lack of policy directly relating to the modification and degradation of wetland ecosystems, but there is a strong endeavour to focus attention on freshwater ecosystem conservation.

The current state of wetlands as the most threatened ecosystem in South Africa, according to the 2011 National Biodiversity Assessment, has provided a great impetus to freshwater ecosystem conservation. This has in turn highlighted the need for comprehensive wetland inventories to focus efforts on the conservation and management of these areas. Wetland inventories require constant updates and improvements in data on the location, extent and characteristics of wetlands in order to boost and improve wetland mapping efforts.

There are numerous studies that highlight the different approaches to mapping wetlands, and at the forefront of these approaches has been the use of satellite imagery. Classification of satellite imagery provides wetland data on a repeatable, long-term and operationally sound basis. Classification of satellite imagery includes an array of approaches from supervised, to semi-automated, to unsupervised classification in a range of temporal, spectral and spatial resolutions. However, certain limitations associated with satellite-image classification are identified and highlighted in the literature, and these include spectral confusion and misclassification of pixels. Emphasis has been placed on the need for and importance of ancillary data to increase the accuracy of identifying wetlands.

Based on this assessment of previous approaches to mapping wetlands, there is a clear need for improvement in the development of non-image-based wetland mapping. This calls for the development of a model that uses other defining variables that could improve the identification of wetlands in the landscape. The Landscape Wetness Potential model developed by Thompson *et al.* (2002) provided valuable insight for the final satellite-image classification of wetlands. Therefore, focus should be placed on the development of accurate and useful ancillary data that will help to define and identify wetland areas found in the landscape. This study explores this research gap by developing a model using Bayesian conditional probabilities of where wetlands are likely to occur in KwaZulu-Natal, South Africa. The end result could be used as ancillary data to earmark high wetland potential areas to further improve the wetland inventory of the province.

CHAPTER THREE

3 Methods

This chapter describes the research design, methods of data collection, identification of input variables, the analysis of variables and the method in building the wetland prediction model. The data used is quantitative in nature, as a strong emphasis has been placed on obtaining statistical frequencies and analytical differences in the frequency of wetlands found in each category in each variable.

3.1 Study area

The study area covered the entire province of KwaZulu-Natal (KZN), which is one of nine provinces in South Africa and is located in the eastern central part of the country (Figure 3.1). KZN covers an area in excess of 90 000 km² and lies between 26°50’ and 31°10’ South, and 28°50’ and 32°50’ East (Eeley *et al.*, 1999). The hydrological and administrative boundaries largely correspond with the rivers in KZN, which arise either from within the western escarpment zone, or internally within the province, and drain east into the Indian Ocean (Rivers-Moore *et al.*, 2010).

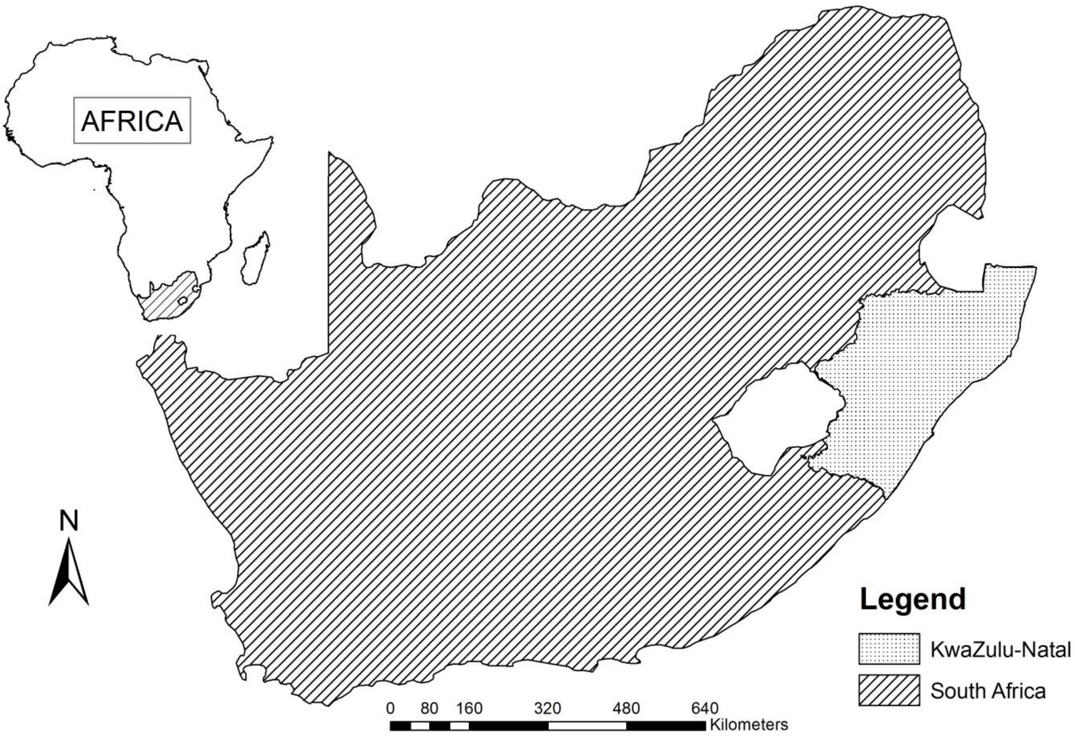


Figure 3.1: The study area of KwaZulu-Natal in South Africa.

The geology, topography and climate of the province are remarkably varied (King, 1978; Schulze, 1982). The western boundary of the province is marked by the Drakensberg mountain range that reaches over 3000 m AMSL in places and runs in a predominantly north-south direction 150 to 280 km from the coast. The land rises from the coastal plains over a series of plateaux to these signature mountains, with the rivers flowing in an eastward direction, cutting downward into the geology defining the topography seen today. The Drakensberg mountains and the warm Mozambique current account for much of the climate experienced in the province, which can be characterised by a large annual variation in temperature and rainfall (Eeley *et al.*, 1999). The entire province lies in a summer rainfall belt, and unlike many parts of Africa, KwaZulu-Natal experiences a relatively balanced water budget owing to favourable atmospheric circulations and the warm coastal current that supplies humidity to the overlying atmosphere (Jury, 1998). Approximately 43 cold fronts affect KZN annually, mainly in winter, and the summer months (November to March) account for 75% of the annual rainfall, which is produced by either large-scale line thunderstorms or orographically induced storms (Nel, 2008). KZN as a province contributes nearly twice as much runoff per unit rainfall than South Africa as a whole, and a quarter of South Africa's streamflow (Nel and Sumner, 2007). Average annual rainfall in KZN varies from region to region: 850–1 400 mm in the coastal lowlands located in the eastern and southeastern parts of KZN; 600–900 mm in the northern and northeastern parts of KZN; and 800– <1 500 mm in the uplands and highlands towards the western parts of KZN (Le Roux, not dated).

KZN has a substantial number of wetlands, covering a spatial area of at least 4 200 km^2 (Rivers-Moore and Cowden, 2012). The wetlands in KZN are particularly rich in three areas, namely the northeastern coastal plains, the northwestern headwaters of the Tugela Basin, and the Midlands and Drakensberg valley bottoms. The wetlands on the northeastern coastal plains owe their existence to the high rainfall averages and subtropical conditions, as well as to a series of marine regressions and transgressions that took place from 120 000 to 20 000 years BP (Patrick and Ellery, 2006). The geology in the coastal plains consists of mainly unconsolidated sediment, but the economic concentrations of minerals such as ilmenite, rutile and zircon make these areas susceptible to mining activities (UKZN, not dated). The wetlands in the northwestern parts of KZN formed as a result of a combination of factors, namely the rainfall, geology, soil characteristics and the gradual slope in the valley bottoms. The gradual slope in the area is a result of the eastward back-tilting associated with the upliftment of the old (interior) African land surface (Begg, 1989; Ellery *et al.*, 2008; Joubert, 2009). The geology in the area is dominated by the Ecca shale and sandstone group, which is known to weather easily and often present slope-stability problems; it is therefore because of the erosion-resistant basalt and dolerite outcrops (Drakensberg group) and the gradual slope that wetlands exist in this area (UKZN, not dated). The Midlands and Drakensberg are richly endowed with wetlands because these areas act as a sponge to any water running off the steeper mountain slopes, and this area is climatically and geologically conducive to wetland formation (high rainfall, gradual sediment-trapping slopes and Karoo dolerite key points). The Beaufort group (mudstone and sandstone geology) and Drakensberg group (rhyolite, basalt, and dolerite) form unique conditions of impervious soils, and horizontal sills and dykes act as new base levels that reduce the gradient of the landscape, creating conditions that allow wetlands to form (UKZN, not dated).

The province has an array of diverse natural resources, of which wetlands are one, and the favourable variation and combination of bio-resources lend themselves to varied agricultural production (mainly sugarcane, forestry and maize), mining activities and a variety of different domestic and industrial uses (KZNPPC, 2011). These activities lead to increasing pressure on and the decreasing ecological health of the province’s natural resources. Although KZN is the third smallest province in South Africa, its agricultural sector contributes to approximately one third of the national agricultural gross domestic product (GDP) and covers an area of approximately 6.5 million hectares. Although agriculture in KZN is a main contributor to the country's GDP, it has a significant effect on the province’s wetlands. The Mfolozi swamps located on the eastern coastal lowlands originally occupied

21 322 ha, but less than half of the system remains in a semi-natural condition due to its being modified into sugarcane cropland, which is the case for a large number of other wetland sites along the coastal lowlands and floodplains (Begg, 1989). Paddavlei is another wetland located in the heart of the Tugela catchment (central KZN) that bears the scars of KZN's socio-economic activities. The Ecca Shale geological structure, coupled with the increasing pressure of overgrazing, appears to be largely responsible for the eroded conditions of the central and bottomlands of the wetland (Begg, 1989). KZN is South Africa's only non-water-scarce province under the current climatic conditions (accounting for almost 40% of all the water within South Africa); however, changes in the river sediment budgets because of land cover change, catchment degradation, main-stem impoundment and inter-basin transfer schemes are particular threats to freshwater ecosystems as a whole (Rivers-Moore *et al.*, 2010).

3.2 Input variables in model

As reviewed in section 2.1.2 and drawing from the illustration in Figure 2.1, wetland areas form through the combination of climatic, geologic and topographic components all interacting at a range of spatio-temporal scales. This is reiterated by Mitsch and Gosselink (2000), who state that key components affecting the formation and function of a wetland are hydrology, the physiochemical environment, biota, climate and geomorphology. The components are not independent and there is significant feedback between them (Mitsch and Gosselink, 2000). The conceptual framework formulated by Batzer and Sharitz (2006) (Figure 2.1) formed the basis for selecting the maximal set of model variables, with topography, geology, climate, soil and hydrology being the key criteria in identifying spatial datasets necessary for wetland determination.

The model was also based on the use of existing data or information, and input variables were selected on the basis of what is currently available and whether they can be directly or indirectly related to wetland form and function. The idea was that by identifying key parameters, irrespective of land-cover characteristics, it would be possible to determine areas where water (and thus a wetland) is likely to accumulate (Thompson *et al.*, 2002).

The input data compiled for the study consisted of both vector and raster data, and was varied in terms of spatial resolution, extent and projection (Table 3.1). The following limitations and assumptions were recognised:

- More than 50% of the layers are model-derived such that errors in these datasets became additive in this model.
- Layers consisted of data in different projections, extent and format due to the input layers being sourced from different organisations and institutions. This was limiting because the layers needed to be standardised in terms of projection, extent and format to be directly comparable.
- The standardisation of all the input variables required the modification of all the layers into a standard projection, resolution and extent. This process may have compounded original errors and limitations because map algebra is used in these standardisation computations. This could not be avoided, but this limitation should be kept in mind.

- There was a lack of information or metadata on how some of the variable layers were derived; therefore, the strengths and weaknesses of the input variables are unknown.

Table 3.1: List of input variables used in the modelling process

	Input variable	Input units	Data type	Source	Supplied resolution (pixel cell size)
CLIMATIC VARIABLES	Solar radiation	$MJ.m^{-2}.day^{-1}$	Continuous Raster	Schulze, 1997	1 600 m
	Mean annual temperature	°C	Continuous Raster	Schulze, 1997	1 600 m
	Summer heat units	°days	Continuous Raster	Schulze, 1997	1 600 m
	Winter heat units	°days	Continuous Raster	Schulze, 1997	1 600 m
HYDROLOGIC VARIABLES	Mean annual precipitation	mm	Continuous Raster	Schulze, 1997	1 600 m
	Mean annual potential evaporation	mm	Continuous Raster	Schulze, 1997	1 600 m
	Mean annual evapotranspiration	mm	Continuous Raster	Schulze, 1997	1 600 m
	Groundwater depth	m	Continuous Raster	Colvin *et al.*, 2007	1 700 m
GEOLOGY, EDAPHIC AND TOPOGRAPHIC VARIABLES	Landform	n/a*	Ordinal Raster	Escott, 2011	80 m
	Clay content	n/a*	Ordinal Raster	Van der berg *et al*., 2009	28 m
	Soil associations	n/a*	Nominal Raster	Van der berg *et al.*, 2009	28 m
	Soil depth	n/a*	Ordinal Raster	Van der berg *et al.*, 2009	28 m
	Hydromorphic soil	n/a*	Nominal Raster	Van der berg *et al*., 2009	29 m
	Geology	n/a*	Nominal Vector	UKZN, 2011	Provincial scale polygon
	Soil moisture	n/a*	Ordinal Raster	Van der berg *et al.*, 2009	29 m
	Terrain units	n/a*	Ordinal Raster	Van der berg *et al.*, 2009	28 m
DIGITAL ELEVATION MODEL (DEM) DERIVED VARIABLES	Altitude	M	Continuous Raster	GISCOE, 2001	20 m
	Slope (degrees)	% rise	Continuous Raster	DEM-derived (2011)	20 m
	Aspect	degree	Continuous Raster	DEM-derived (2011)	20 m
	Flow accumulation	n/a	Continuous Raster	DEM-derived (2011)	20 m
	Flow direction	degree	Continuous Raster	DEM-derived (2011)	20 m
	Sinks ('Depressions')	n/a	Nominal Vector	DEM-derived (2011)	20 m

* See Appendix A, Table 6.1.1 for the legends of these modelled layers

3.2.1 Brief description of input variables

Climatic and hydrologic variables

Solar radiation, mean annual temperature and heat units are all related to the water availability in a wetland because they are related to atmospheric demand. Mean annual rainfall provides hydrologic inputs into a wetland through runoff and groundwater, while the variables related to atmospheric demand provide the energy that drives processes such as evaporation and evapotranspiration. Thus, evaporation and evapotranspiration relate to a wetland's water balance by being responsible for the loss of water in a wetland.

Solar radiation was estimated and derived from a single expression that includes the extraterrestrial radiation (i.e. radiation at the 'top' of the atmosphere), the maximum daily air temperature (as a surrogate for solar radiation) and the temperature range (as an index of humidity and cloudiness) (Schulze, 1997). *Mean annual temperature* was derived using data from a number of temperature stations, and then the mean annual temperature was calculated; however, there is a loss of effects such as diurnal, monthly and seasonal patterns of maximum and minimum temperatures. The *heat units* variable was calculated using the accumulation of mean temperatures above a certain threshold value (10°C) and below an upper limit, over a period of time (e.g. if the threshold temperature is 10°C and the mean temperature of a given day is 22°C, then 12 degree days, or heat units, are accumulated for that day to a previous total (Schulze, 1997)). This variable revolves around the development of a plant or organism and the total heat it is subjected to during its lifetime.

The *mean annual precipitation (MAP)* variable was derived using data from over 6 000 rainfall stations. Equations for the MAP were developed, from which the MAP gridded values were generated. The MAP is an important variable in terms of its effect on the hydrology of a wetland, and as mentioned by Mitsch and Gosselink (2000), hydrologic conditions are extremely important for the maintenance of a wetland's structure and function.

The *mean annual evaporation* variable was derived from A-pan equivalent evaporation values, which were summed at each of the 437 000 grid points covering southern Africa. The *mean annual evapotranspiration* is calculated through an adapted 'Penman-Monteith' method of calculating evapotranspiration using reference crop evapotranspiration, net radiation at crop surface, average screen-height air temperature and daily mean wind speed at a height of two metres (Schulze, 1997).

The *groundwater* variable lacked metadata or any information on how the input layer was derived (Colvin *et al.*, 2007).

Geology, edaphic and topographic variables

As referenced in chapter 2, section 2.1.2, geology, edaphic and geomorphology are key in the formation of wetlands. High clay content areas may result in clay-based wetlands, or at least signify areas that are impervious to water; therefore, if they are located in areas of gradual gradient or are surrounded by a geological control, such as a dolerite sill, water is likely to accumulate in that area that will allow a wetland to form. Soil moisture and hydromorphic soils are variables likely to define the location of a wetland area because they signify areas presumably well saturated under normal conditions. As mentioned by Ellery *et al.* (2008), the majority of wetlands in South Africa occur along rivers and on coastal plains. It is therefore important to have an idea of where valley bottoms occur, as well as where the gradually sloping lowlands occur, and therefore landform and terrain units are vital components in predicting where wetlands are likely to occur.

The *clay content, soil association* and *soil depth* variables were model-derived layers created using 42 450 field observations and 3 319 laboratory-analysed sites by the ARC–Institute for Soil, Climate and Water (ARC–ISCW). Through extracting raster values from soil colour, terrain units, slope, geology and soil moisture layers, unique combination tables were generated. This dataset was used to establish relationships between soil attributes and various spatial data layers, and then these relationships were used to establish scripting rules for integrating the various variables to create these model-derived layers (Van der Berg *et al.*, 2009). The *hydromorphic soils* variable was derived through generating the *soil association* layer. *The geology* variable was derived using soil surveys across KZN.

The *soil moisture* variable was computed and derived by another ARC–ISCW project (*Soil moisture estimation for KwaZulu-Natal*) by Metz (2008). The soil moisture estimation layer for KwaZulu-Natal was generated in order to assist in the data analysis for soil-form determination. Actual soil moisture is determined by recent precipitation and vegetation cover. The soil moisture layer is a seamless map with five moisture classes (estimates of the potential soil moisture). Although other methods to determine soil moisture exist, they have the limitations of being coarse in resolution and not correcting for vegetation as an influencing factor in soil moisture. Metz (2008) used an inverted Normalised Difference Wetness Index (NDWI) using Landsat imagery over 16 years of band 4 (identifying levels of chlorophyll) and band 5 (infrared; where low values indicate high water content). The formula for the normalised water content index was (255 − band5 − band4) / (255 − band5 + band4). The resultant soil moisture layer was not ground-truthed; however, a classification scheme was generated against other available data, predominantly geology and land cover. It readily distinguished different soils with similar vegetation that could then be classified into soil moisture classes. Metz (2008: 06) stated that “the soil moisture classes conform well against field observations, but actual soil moisture measurements are needed to assess the accuracy more systematically.”

The *terrain unit* variable was created using the Shuttle Radar Topography Mission (SRTM) data. Terrain units were created as defined by the National Land Type Survey. The five classes of terrain units were: crest, midslope convex, midslope concave, foot slope, and valley-bottoms. The *concave mid-slope* class is valuable in detecting potential hill-slope seepage areas.

The *landform* variable is a terrain-based analysis of the slope model derived from the 90 m SRTM DEM to generate topographical characteristics within KZN. Topographical characteristics included deeply incised streams, local ridges, midslope drainages, midslope ridges, mountain tops, open slopes, plains, U-shaped valleys, upland drainages, and upper slopes.

Digital Elevation Model (DEM) derived variables

The *DEM* was derived from vectorised 20 m contour data from the 1:50 000 topological maps provided by Surveys and Mapping, Chief Directorate: National Geo-spatial Information. It became apparent that while creating the DEM-derived layers, this layer was creating terrain errors with an excessive number of sinks and broken flow-accumulation lines. This is a noted limitation in the data that has to be accepted and could result in the over- or under-estimation of potential wetland areas. However, this could also prove to be negligible due to the number of other contributing input variables, such as the landform variable derived from accurate 90 m SRTM DEM. The 20 m DEM is also used as a separate input into the model.

The *flow accumulation* variable represents the flow of water across a surface by calculating the accumulated flow as the accumulated weight of all cells flowing into each downslope cell in the raster output. The *flow direction* is initially calculated, which indicates the direction in which the commodity flows through the network of cells, and this is then used to determine the accumulation of flow per cell. Both flow direction and accumulation are DEM-derived variables generated through the hydrology toolbox located in the spatial analyst toolbox of ArcMap 9.3 (ESRI, 2007). This modelling technique is usually used to delineate stream networks through the extraction of cells with a high flow concentration, or areas of concentrated flow. Therefore, areas that have medium to high accumulated flow and do not constitute a stream could be interpreted as possible areas of wetness, and could therefore indicate the possible presence of a wetland.

The *slope steepness* variable was derived using the slope tool in ArcMap 9.3 (ESRI, 2007) under the spatial analyst toolbox. The slope was expressed as degree rise and therefore low degree-rise areas (flat-surfaced areas) can be classified as higher potential wetland areas. The US Soil Conservation Service (SCS) has recognised that areas with a slope of less than 8° are potential wetland locations (Junhua and Wenjun, 2005).

The *sink or depression* variable was derived from the sink tool in the hydrology toolbox under the spatial analyst toolbox in ArcMap 9.3 (ESRI, 2007). This variable indicates areas that cannot be allocated a flow-direction value. In this study they will be assumed to be valid areas of internal drainage and therefore to have a high possibility of being a wetland area. Sinks can, however, also be errors in the DEM itself and therefore it is an overrated assumption that these are valid depressions in the landscape (Thompson *et al.*, 2002).

3.2.2 Current 2011 KZN wetland layer

The current KZN wetland layer was a key dataset for this study because it was used to extract environmental parameter statistics correlating with known wetland areas, and to assess and validate the final probability layer output. The current KZN wetland layer was compiled by Scott-Shaw and Escott (2011) as a product for Ezemvelo KZN Wildlife's (EKZNW) Biodiversity Conservation Planning Division. It is the most comprehensive coverage map for KZN and represents the best available wetland dataset for the province. The provincial wetland layer is a compilation product that has been ongoing over the past decade, having started circa 2000 (Escott pers. comm). The coverage draws from multiple sources, including the 1:250 000 geological map, the 1:50 000 Tugela Soils maps, the priority wetlands coverage identified by Begg (1989), manual mapping using a wide variety of aerial and satellite imagery, and source information compiled and submitted to EKZNW by external NGOs and private companies (e.g. Mondi, SAPPI, etc.), as well as submissions of coverages from local and district municipalities. Basically, the Biodiversity Conservation Planning Division sourced wetland data from wherever they could.

Scott-Shaw and Escott (2011) compiled all the available wetland data into a single dataset, producing the current KZN wetland layer. Initially, all the various datasets were simply compiled into a single coverage, with no consideration for the differences between original capture scales and/or temporal differences (i.e. summer or winter capture extents). To try to account for this, an initial 'standardisation' cleaning process was initiated in 2008/2009, during which the entire dataset was compared against the SPOT 5 2008 coverage (Escott pers. comm). The edits made to the wetland layer were an attempt to standardise the mapping of wetlands at a consistent scale, using imagery from a consistent time period, (albeit based on obvious vegetative changes associated with the wetlands as opposed to the true wetland extent as reflected by the underlying soil structure).

Boyd Escott (pers. comm.), the manager of the Biodiversity Spatial Planning and Information sector at EKZNW, elaborates that the weaknesses in the KZN wetland layer can be attributed to a few key sources of error:

1. The wetlands that have been mapped only reflect the visual vegetation changes associated with a wetland and not its true extent as defined by soil type.
2. The mapping was accomplished by manually digitising the extents using relatively low-resolution imagery. The manual process was and is very time consuming, and errors do occur where smaller wetlands (and sometimes larger ones) aren't mapped, either because they were missed or because they don't reflect obvious differences relative to the surrounding vegetation.
3. The wetlands were originally sourced from information captured at different resolutions. EKZNW did try to standardise this using a single SPOT 5 image reference and a set scale, but some were inevitably missed.
4. A wetland's digitised extent relates to the visual extent of the wetland vegetation associated with a wetland. While potential seasonal changes were addressed (and standardised) through the use of a single-reference imagery set (captured over a one-month period), changes in extent reflecting available surface water could not be accounted for, i.e. in a dry season, the

vegetative extent associated with the wetland will contract relative to that noted in a wet period.

5. Only the remaining extent of wetlands (i.e. unmodified extent) could easily be mapped. Historic wetlands that are now lying under dams, cultivated lands and modified lands cannot be mapped. There may also be value in mentioning that in the current KZN wetland layer, what may appear to be separate wetlands may in fact be part of larger systems that have been fragmented over time.
6. The classification of the wetlands has been done primarily through a desktop exercise, although expert input was also incorporated in specific areas.

There are certain ways in which the current wetland layer for KZN could be improved (Escott pers. comm.):

1. Ground-truthing of the wetlands is required. When field visits are conducted and/or wetland extents are submitted by consultants, the affected wetlands are corrected (both their attributes and, if possible, extents). There are fields within the spatial map that then capture the verification date and the name of the verifier. However, such ground-truthing takes place only sporadically at the moment.
2. Field studies are desperately needed. Boyd Escott would like to see a study that examines a single wetland category, and that involves a series of ground verification visits to obtain a measure of accuracy associated with the mapped wetland's class.

It must be reiterated that at the time of this study, no quantified accuracy assessment of this KZN wetland layer has been publicly documented or reported on (Escott pers. comm.). It was an assumption that the current KZN wetland layer was accurate in terms of the current visible wetland extent area and location, which required quantification for this study. To test this confidence in the KZN wetland layer, a rough accuracy assessment was completed before the Bayesian modelling began. Using 239 referenced aerial photos of wetland sites (Jewitt, 2011), the current KZN wetland layer was visually assessed to determine if the coverage had captured the location and extent of the referenced wetland sites. The 239 sites were visually assessed using low-flown aerial photographs with co-ordinates, clearly identifying different wetland systems spread broadly across the entire province. Based on the 239 sites, the current KZN wetland layer identified 196 of the 239 sites correctly, providing an accuracy estimate of 82%. Assessing the accuracy of the current KZN wetland layer in terms of completion, extent and location was important in providing a confidence estimate of the wetland layer used in the development of this model.

From the assessment, there was no conclusive evidence that suggested that wetlands were missed because they were concentrated in a certain bio-region, or because of their size, or could be attributed to certain land-use developments. Although questions regarding the shortfalls of the wetland layer are relevant and should be answered through further research, it was beyond the scope of this study to do so, and the objective of this exercise was only to gain a rough idea of the accuracy of the KZN wetland layer.

3.3 Modelling approach

The broad method and approach is summarised in a flow chart that represents the method used to create the model by breaking the process down into a number of steps (Figure 3.2). Each step is described as per the numbers in the hashed boxes.

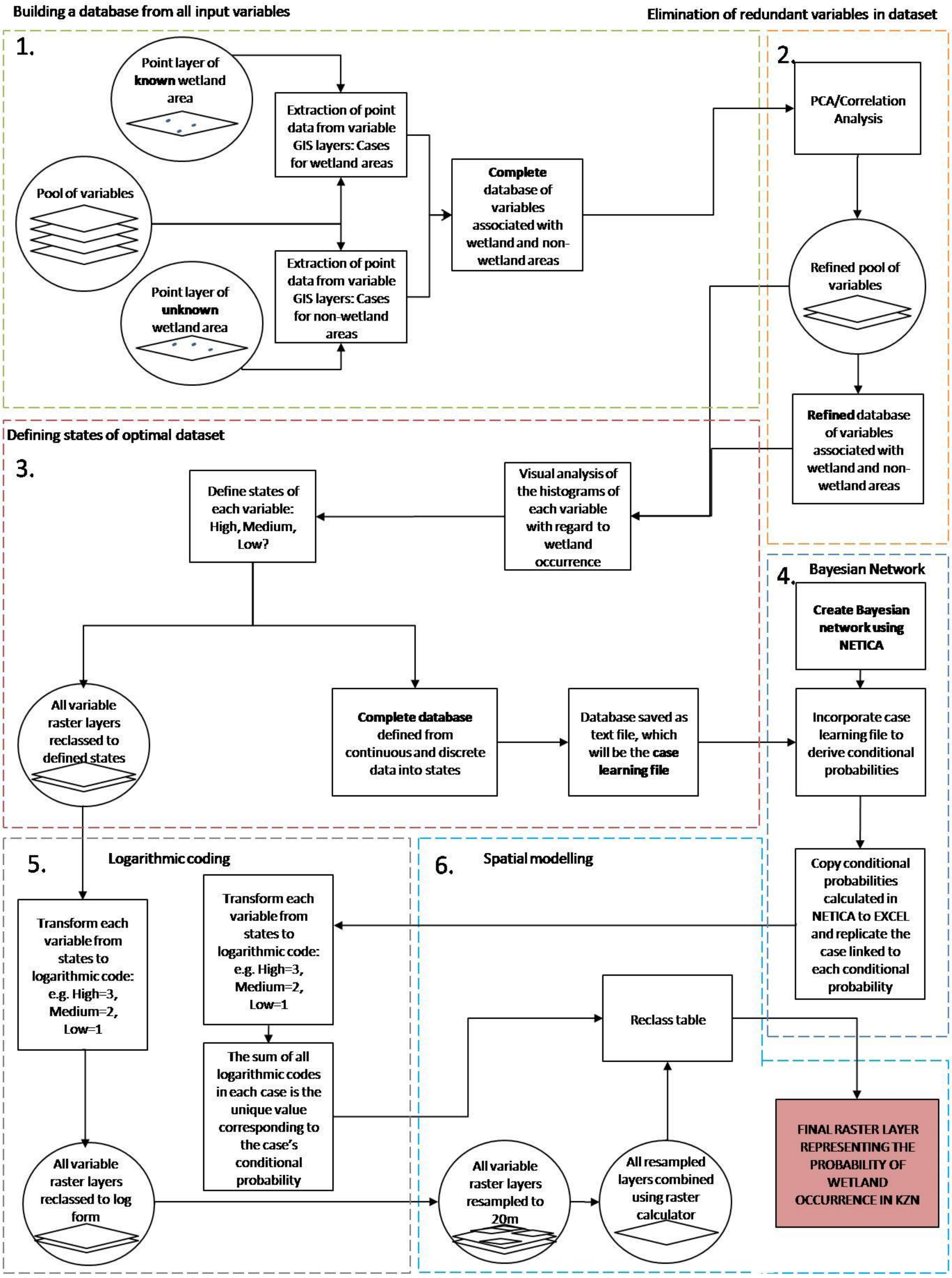


Figure 3.2: Flow chart illustrating an example of the method and approach used in generating the final probability layer of wetland occurrence in KZN.

1. Building a dataset from input variables

The aim of the dataset is to generate numerical values for each input variable and obtain analytical differences in known wetland areas and unknown wetland areas using ArcGIS 9.3 (ESRI, 2007). The dataset is a compilation of extracted statistical values from each input variable in both known wetland areas and non-wetland areas. The current KZN wetland layer was randomly split (Hawth tools (Beyer, 2004)) into a *training* and *test* dataset to avoid over-fitting of the model (Aguilera *et al*., 2011). The training dataset was used to build the dataset by extracting statistical information in known wetland areas. The test dataset formed a control dataset of known wetland areas used to assess and validate the final model output. Randomly generated points (created using Hawth tools (Beyer, 2004)) were used to extract statistical information in non-wetland areas. The ‘pool of variables’ refers to all the input variables that were used to build the dataset.

A total of 45 000 points were used to extract statistical information for the dataset, of which 25 000 were known wetland areas, and the remainder were non-wetland areas. The complete dataset contained information on each input variable in known wetland areas, and the non-wetland dataset information formed the control for comparison and statistical analysis. The dataset was exported to a spreadsheet for final editing and compilation. The dataset formed the quantitative basis for all the steps to follow in this method and approach. (Descriptive statistics of all input variables can be found in Appendix B, Table 6.2.1).

2. Elimination of redundant variables

There were 22 input variables preliminarily selected for this model. The objective of this step was to investigate whether there were high levels of correlation between the input predictor variables, and, from the analysis, to derive an optimal dataset by eliminating redundant variables. Principle Component Analysis (PCA) was used to refine the set of input variables, to minimise redundancy in the input layers used in the study, and to provide the greatest predictive power with regard to where wetlands are likely to occur. The program used to run the PCA was the MultiVariate Statistical Package (MVSP 3.2).

PCA is a robust multivariate technique for assessing the degree of correlation between several quantitative variables (Abdi and Williams, 2010): “PCA’s are used for: extracting the most important information from the data table, compressing the size of the data set, and analysing the structure of the observations and the variables” (Abdi and Williams 2010: 3).

This approach of eliminating variables using PCA could only be processed using ordinal data. This step therefore temporarily excluded nominal variables, namely *soil association*, *sinks*, *geology* and *hydromorphic soils*. The process of elimination involved the step-wise analysis of the biplots and variable loadings of each input variable. The vector length of a variable measures the magnitude of its effect on the presence and absence of a wetland. The cosine of the angle between the vectors of two variables measures the similarity between them relative to the determination of wetland occurrence (Yan and Tinker, 2004). If the angle between two variables was small or non-existent, the variable with the lower variable loading would be eliminated. Preference was given to variables with a higher spatial resolution because of their ability to distinguish value ranges at a finer scale. Correlation between variables was determined through the ‘similarity matrix’ calculated by running the PCA in MVSP 3.2.

The PCA was rerun following each variable elimination. The co-linearity of the dataset was tested following each rerun, until it was below the critical value of 10 (Fry, 1993). The result of the PCA process is a refined list of input variables. At this stage, following the process of elimination, the excluded nominal variables were reintroduced into the variable pool for inclusion in step 3.

3. Defining variable states

Calculating conditional probabilities through Bayesian networks required the data to be reduced to a set of mutually exclusive qualitative states (e.g. high, medium and low). The approach taken to calculate conditional probabilities was adopted from Cain (2001).

The refined pool of input variables was translated from continuous values or qualitative states to discretised states. Both dataset and spatial layers were translated to states. States were defined by analysing the data distribution (through a histogram) of each variable, in ArcGIS 9.3. Two approaches were used in discretising the data:

1. Continuous data was reclassified into states using the Jenks Natural Break algorithm (see Jenks and Caspall, 1971). Natural break classes are based on natural groupings inherent in the data. Using the Jenks Natural Break algorithm, class breaks are identified that best group similar values and that maximise the differences between classes.

2. Data in qualitative states were defined into states that best represented the respective variable characteristics. This was done for both the dataset (Microsoft Excel 2007) and the spatial layers (ArcGIS 9.3). All spatial layers were reclassified into states defined by the intervals used in the dataset. It is important to note that the states of high, medium and low had no meaning in terms of the probability of wetland occurrence, and were only used to define the variables into qualitative classes. Nominal variables that could not be quantified into three single qualitative states as required for the Bayesian network were eliminated from the model.

4. Constructing the Bayesian network

The dataset defined into states was saved as a text file (Appendix B, Table 6.2.2). This became a *case learning file* for the Bayesian software NETICA (Norsys Software Corporation 1995–2011), described by Aguilera *et al.* (2011) as the most popular software used in environmental modelling. The next step was building a BN that represented the remaining variables and defined the model's objective. The remaining variables represented their conditional dependencies, from which the outcome was the occurrence of a wetland. The conditional probability of wetland occurrence is dependent on the conditional dependencies of the variables.

The *case learning file* formed the central input for calculating the prior probabilities of all the parent node variables in the BN, as well as the final conditional probabilities. Using a counting learning algorithm, the NETICA software created a conditional probability table (CPT). The counting learning algorithm is a Bayesian learning technique, where prior probabilities are determined for each node through counting how many cases, of the defined dataset, are high, medium or low, irrespective of whether they are wetlands or not, which is then presented as a percentage. The CPT is a table of all possible variable-state combinations with a representative conditional probability, where the conditional probability is based on the prior probabilities found in the parent nodes (for clarity on parent nodes see Figure 2.2, section 2.4.2). Each variable-state combination is called a scenario. The number of scenarios in a CPT is dependent on the number of possible variable-state combinations.

The final CPT contained the conditional probabilities of wetland occurrence and formed the basis for creating the final probability layer. The CPT was replicated in a spreadsheet (see Appendix B, Table 6.2.3).

5. Coding variables for the spatial output

Bayesian networks are not spatially explicit (Grêt-Regamey and Straub, 2006) and in order to generate the final wetland probability layer, the conditional probabilities had to be spatially represented. NETICA did not support the replication of the conditional probabilities in spatial format in ArcGIS 9.3 by script or algorithm exportation.

This limitation was averted by logarithmically coding both the CPT and the spatial layers, so that when the spatial layers were combined to create the single modelled layer, each pixel in the single combined layer could be linked to scenarios and subsequent conditional probability in the CPT. Combining the coded spatial layers, through simple addition, into a single layer created a single layer with pixels containing unique logarithmic values; this matched unique logarithmic values and their conditional probabilities in the CPT.

To clarify, GIS layers and the CPT were modified from their defined states into logarithmic code. States that were 'HIGH' became log forms of 3, 'MEDIUM' became log forms of 2, and 'LOW' became log forms of 1 (Figure 3.3). Combining the logarithmic translated scenarios in the conditional probability table produced a unique value, with a corresponding conditional probability (see Appendix B, Table 6.2.4). Combining all the logarithmically modified variable spatial layers produced pixels with unique values that could be linked to the CPT. The logarithmically coded Bayesian CPT formed a *reclassification* text file (see Appendix B, Table 6.2.5). The reclassification file was used to reclassify pixel values in the spatial layer to conditional probabilities.

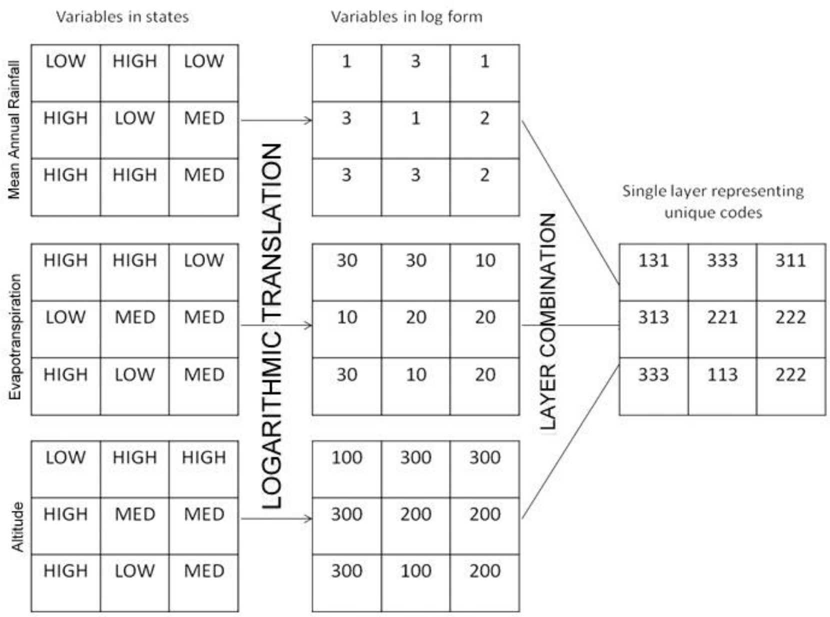


Figure 3.3: A diagram indicating the logarithmic translation of the spatial layers from discrete states to logarithmic code at pixel scale, and the final combination of the layers to form the unique values.

6. Generating the final probability model layer

The final step following the logarithmic modifications was combining the GIS layers into a single raster layer of unique codes (Figure 3.3) and reclassifying the unique codes in the single raster layer to correspond to the conditional probabilities in the table.

Before combining all spatial layers, all nine variable raster layers were standardised. Standardisation involved transforming all layers to a projection of 'WGS 84' and a standard extent used in the model, This was important so that spatial analysis occurred on a 'pixel on pixel' basis, preventing the misalignment of pixels when creating the unique values.

Spatial resolution was standardised to 20 m. Various raster layers were interpolated from a coarse resolution (+/- 30 m and 100 m) to a 20 m resolution using a cubic or nearest neighbour resampling technique. The cubic resampling technique was used for the continuous data, which performed a cubic convolution, determining the new value of a cell based on fitting a smooth curve through the 16 nearest input cell centres (ArcMap 9.3 Help). The nearest neighbour resampling technique was used for the categorical data, where the new cell value is determined by the nearest neighbour. This is important because it will not alter the cell values in the categories.

Following the standardisation procedure, the raster input layers of variables were combined into a single layer in ArcGIS 9.3, producing a single raster layer with unique values (as seen in Figure 3.3). The unique values were reclassified to conditional probabilities using the *reclass text file*. This step created the final probability raster layer, spatially indicating the probability of wetland occurrence in KZN.

3.4 Model verification and assessment

Once the final probability raster layer was created, it was important to verify the model in terms of its accuracy in predicting the likelihood of wetland occurrence. The assessments in this study aimed to address three questions:

1. How has the modelled probability layer predicted the distribution of wetland areas across KZN?
2. In comparison to wetland areas already identified, does the modelled probability layer correlate well with these areas?
3. In terms of applicability in determining new wetland areas, what probability values should be used to predict the occurrence of a wetland and its extent?

Many BN studies do not verify their model, and methods of assessment largely depend on the original aim of the model (Aguilera *et al.*, 2011). In this study, 'verifying the model' is defined as confirming that the model correctly predicts wetland areas. This was done by assessing the final layer and addressing the three questions in the following ways:

1. **Verification of final probability layer**

   Probabilities from the BN model were assessed in approximately 8 600 test wetland sites (derived from the test wetland dataset (section 3.3., Step 1, line 4)), and 8 600 non-wetland sites (derived randomly using Hawth tools). The c. 16 000 wetland and non-wetland sites formed the binary data necessary for the verification of the modelled probability layer.

A Receiver Operating Characteristic (ROC) curve and Area Under Curve (AUC) were chosen to verify the model using MedCalc for Windows, version 12.5 (MedCalc Software, Ostend, Belgium). ROC analysis is a useful technique in visualising the performance of a binary classifier system, as its discrimination threshold is varied. This provides a richer measure of classification performance than scalar measures such as accuracy, error rate and error cost (Fawcett, 2006). In this instance, key to this analysis was determining the sensitivity and specificity of the final probability layer at a criterion threshold as well as the AUC. The sensitivity defined how many positive results occur among the 8 600 test wetland sites (Equation 2), and the specificity defines how many correct negative results occur among the 8 600 non-wetland test sites (Equation 3).

$$\text{Sensitivity} = TP/P = TP/(TP + FN) \qquad [2]$$

$$\text{Specificity} = TN/N = TN/(FP+TN) = 1\text{-Sensitivity} \qquad [3]$$

where P is positive and N is negative, TP is true positive and FP is false positive, FN is false negative and TN is true negative.

2. **Assessment of the final probability layer**

i. **Visual assessment of the final probability layer**: Visual desktop assessments were performed on five wetland areas across KZN (Figure 3.4). The process involved visual interpretation of what was seen at the five sites at a scale of 1:200 000 and comparing it to the current KZN wetland layer. This was a descriptive assessment of the performance of the final modelled wetland probability layer.

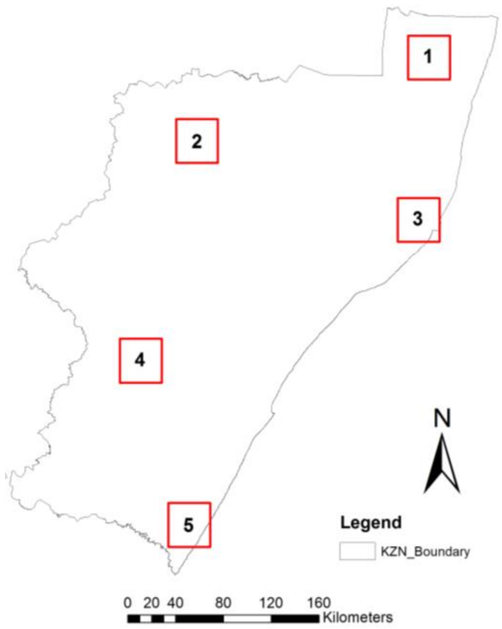


Figure 3.4: The five case study sites (red boxes) that were chosen for the visual assessment against the current KZN wetland layer.

ii. **The final probability layer within historical wetland boundaries:** Another three well-documented individual wetland systems across the province were chosen for closer visual assessment. The wetland systems chosen were the Pongola floodplain (27°3'01"S 32°15'47"E), the highly modified Mfolozi floodplain (28°28'13"S 32°17'46"E), and the Mgeni Sponge (29°30'03"S 29°51'15"E). These systems were chosen on the basis that they best represent the true historical boundary of the wetland and have been derived from various field-verified assessments (see Begg, 1989; MacFarlane *et al.*, 2011). This was intended to quantify the degree to which the model's output corresponds with comprehensively field-verified historical wetland maps.

3. **Assessing the accuracy of the final probability layer in predicting wetland areas**: The final probability layer was assessed and verified in terms of its accuracy in correctly predicting wetland occurrence (presence or absence of a wetland) and the wetland extent (area covered by wetland compared to the 2011 KZN wetland test layer). There were three steps to this assessment:

i. *Assessing the model's accuracy in correctly predicting wetland occurrence (presence or absence of a wetland).* 1 000 random points were generated across 20 classes of probability (i.e. 5–10%, 10–15%, 15–20% etc.) using Hawth tools (Beyer, 2004). Through a desktop exercise, each point was assessed using 2009 SPOT 5 imagery and Google Earth. The researcher drew on past experience with desktop wetland delineation to assess whether each point occurred in a wetland or not. If the researcher could not, using the imagery, conclusively substantiate that the point assessed was a wetland, then the point was recorded as having no wetland present.

ii. *Assessing the model's ability to predict the mean wetland extent area.* The 2011 KZN wetland test polygons were used to extract the area of wetland extent covered with the increase in probability of the model. The mean wetland extent area was calculated by dividing the total wetland area inside the wetland test polygons by the number of wetland test polygons.

iii. Wetlands are often classified into wetland types using the hydrogeomorphic (HGM) classification system and therefore the model was assessed to see if it predicted some HGM wetland types better than others. HGM is a wetland classification system that places an emphasis on geomorphic and hydrologic attributes, as opposed to classification systems that are limited to biotic characteristics (USDA, 2008). The NFEPA wetland layer is a national coverage that has been classified using HGM types. It was therefore used to extract probability values within six different wetland types: unchannelled valley bottoms, channelled valley bottoms, depressions, seeps, flats and floodplains (refer to Appendix A, section 6.1.2 for more information on each HGM type). This provided insight into which HGM types the model predicts best.

CHAPTER FOUR

4 Results

4.1.1 Identifying independent variables for inclusion in the Bayesian probability model

The original dataset of 18 candidate **ordinal** wetland predictor variables was reduced to eight variables through the PCA elimination process. *Nominal variables*, *soil association*, *sinks*, *geology*, and *hydromorphic soils* were excluded from the PCA elimination process since the process was restricted to quantitative data only. It is important to note that the *nominal variables* were only excluded during the PCA process and were again added to the remaining variables following the PCA process. The maximal PCA accounted for 53.66% of the cumulative variance in axes 1, 2 and 3, and showed high levels of correlation (i.e. redundancy) between a number of variables (Figure 4.1). The co-linearity condition number of the first PCA iteration was 37.864, far exceeding the critical value of 10 (Fry, 1993). An example of this redundancy was the high correlation between the *clay* variable and the *altitude* variable (correlation < - 0.9). The *clay* variable was eliminated from the PCA due to this correlation, and due to the *altitude* variable having a higher spatial resolution and accounting for more variation. Following the first iteration of the PCA, *flow direction*, *flow accumulation*, *soil moisture* and *aspect* (modified to ordinal variables) were also eliminated because of their short vector length signifying only a small magnitude of their effect on the presence or absence of a wetland. In the second PCA iteration, the temperature variables (*winter heat units*, *summer heat units*, and *mean annual temperature*) were highly negatively correlated with the *altitude* variable vectors indicated by the +/- 180˚ angle separating the vectors (correlation < -0.9). Due to *altitude* accounting for more variation and having a higher spatial resolution, the temperature variables were eliminated. In the third PCA iteration, the biplot vectors of the *groundwater* variable and *slope* (DEM-derived) variable displayed virtually no angle between them (i.e. correlated variables), resulting in the elimination of *groundwater* because it is the variable with the shortest vector and smaller variable PCA loading. In the final PCA iteration, *evaporation* was highly correlated (0.895) with the *evapotranspiration* variable. *Evaporation* was eliminated because it had a lower variable PCA loading between the two. Following the elimination of the *evaporation* variable, the dataset's co-linearity condition number fell below the critical value of 10 (Fry, 1993). After a number of iterations, the maximal dataset was reduced to eight ordinal variables, with a resultant co-linearity condition number of 5.345. The final iteration of the PCA accounted for 69.15% of the cumulative variance in axes 1, 2 and 3 (Figure 4.2).

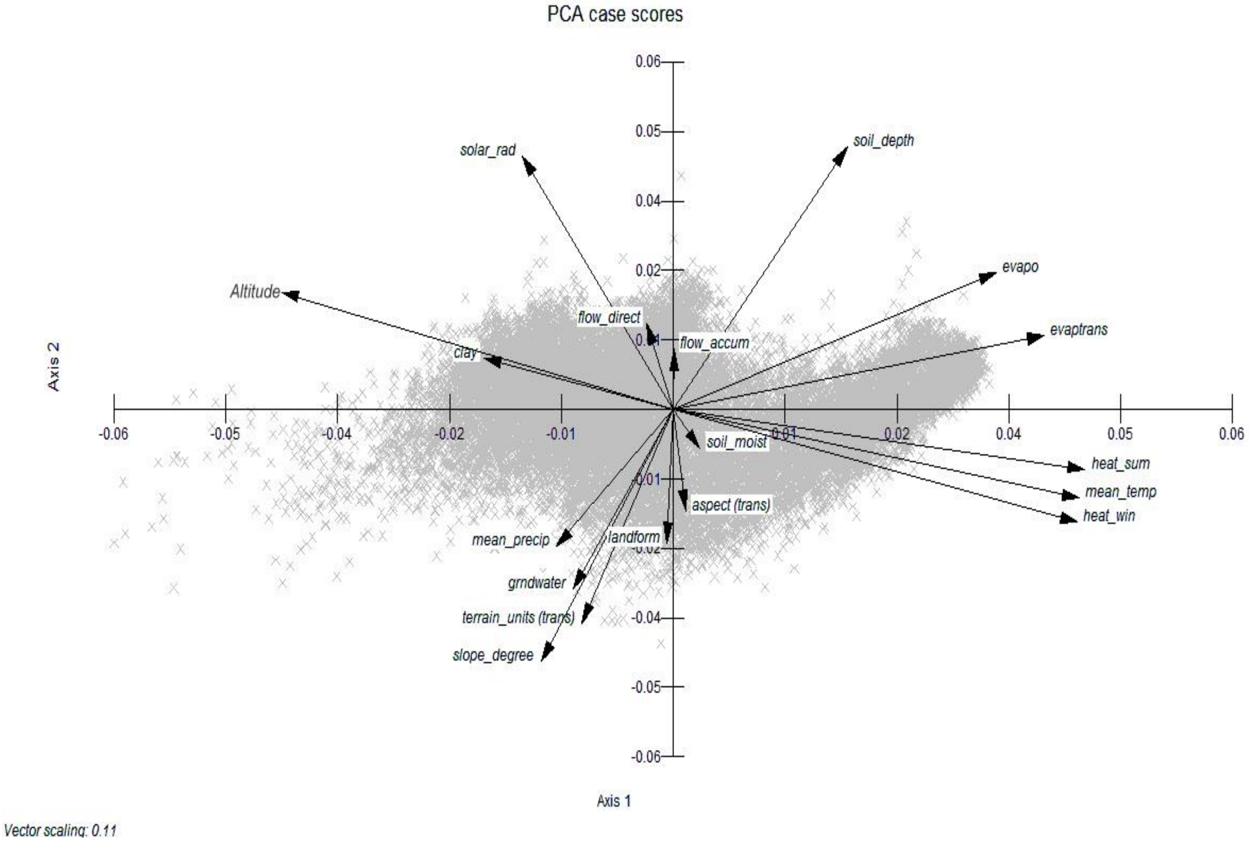


Figure 4.1: A biplot of the initial Principle Component Analysis involving all variables.

Following the final PCA iteration, both *landform* and *terrain units* could have been eliminated because of their close correlation with *mean annual precipitation* and *slope* (in degrees), respectively (Figure 4.2). *Soil depth* was also negatively correlated with *slope* (in degrees) and *terrain units*, which meant that soil depth decreased with the increase in slope angle. In this case, these variables were retained because the process of elimination was stopped when the co-linearity condition number fell below the critical value of 10. This meant that the redundancy in the dataset was within acceptable limits and presented an optimal dataset for further analysis.

According to the eigenvector scores, the variables that accounted for the highest amount of variance in Axis 1 in explaining wetland presence were *soil depth*, *evapotranspiration* and *altitude* (eigenvector scores of = 0.483, 0.447 and -0.436 respectively; Table 4.1). The variables with the highest eigenvector scores for Axis 2 were *solar radiation* (eigenvector value = - 0.650) and *altitude* (eigenvector value = - 0.509) (Table 4.1).

Table 4.1: Eigenvector scores for the remaining ordinal input variables for Axis 1 and Axis 2

PCA variable loadings	AXIS 1	AXIS 2
Soil Depth	0.483	-0.234
Terrain Units	-0.369	0.222
Solar Radiation	-0.148	-0.650
Mean Precipitation	-0.181	0.271
Evapotranspiration	0.447	0.154
Altitude	-0.436	-0.509
Slope (degree)	-0.398	0.271
Landform	-0.167	0.212

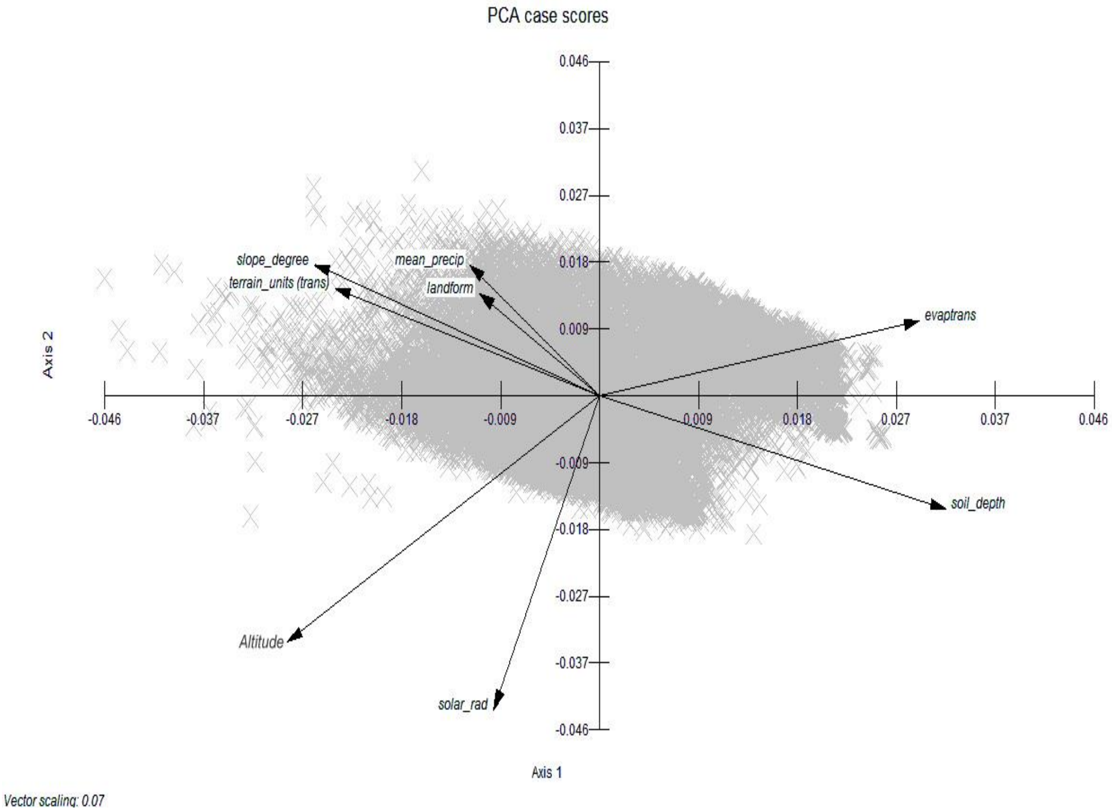


Figure 4.2: Biplot of the final PCA showing eight variables and with the co-linearity coefficient reduced to 5.345.

The remaining ordinal variables were *mean annual rainfall*, *slope in degrees*, *altitude*, *mean annual solar radiation*, *soil depth*, *evapotranspiration*, *terrain units* and *landform*. These eight input variables were the surrogate predictors in determining the likelihood of wetland occurrence across KZN.

4.1.2 Defining variable states

Following the PCA elimination process, *hydromorphic soils* was the only nominal variable added post PCA and the reason for this is as follows: in preparation for the BN, the remaining variables had to be defined into states (a prerequisite for the Bayesian probability modelling) and this required careful consideration as to how the data of each variable was discretised. Once again, nominal variables posed a problem. *Soil association* and *geology* were made up of a number of nominal classes, which could not be discretised into states (high, medium or low) because they could not be quantified in that sense. The discretising of variables into states with many nominal classes is a limitation of this method and approach, and is a challenge of BNs in general (Uusitalo, 2007). The *sinks* (DEM-derived) variable had a high error possibility. Sinks can be errors in the DEM itself and therefore it was an overrated assumption that these are valid depressions in the landscape (Thompson *et al.*, 2002). Therefore, the variables *soil association*, *geology*, and *sinks* (DEM-derived) were eliminated from the model. *Hydromorphic soil* was thus the only addition post PCA because hydric soils could potentially be a powerful predictor of wetland areas. The refined pool of variables was a total of **nine variables**. Table 4.2 illustrates how the remaining input variables were discretised into states. They were defined either by 1) reclassifying the variable into states using the Jenks Natural Break algorithm, or 2) defining variables into states that best represented the respective variable characteristics (e.g. see *soil depth* variable in Table 4.2).

Table 4.2: Discretising input variables into states

Variable	Values	New State	Defining method*
Mean annual rainfall (mm)	0-804	Low	1
	804-1028	Medium	
	1028-1977	High	
Slope in degrees	0-8.2	Low	1
	8.2-21.5	Medium	
	21.5-87.4	High	
Altitude (DEM) (m)	0-632.6	Low	1
	632.6-1333.4	Medium	
	1333.4-3449.1	High	
Solar radiation (MJ.$m^{-2}.day^{-1}$)	0-21.3	Low	1
	21.3-23.0	Medium	
	23.0-27.3	High	
Soil depth	Shallow soils dominant (1)	Low	2
	Medium deep soils dominant (2) and Deep soils dominant (3)	Medium	
	Water (4)	High	
Evapotranspiration (mm) (annual)	0-100	Low	1
	100-112	Medium	
	112-132	High	
Terrain units	Foot slope (4), Valley bottom (5)	Low	2
	Midslope convex (2), Midslope concave (3)	Medium	
	Crest (1)	High	
Landform	Plains (5)	Low	2
	Canyons (1), Shallow valleys (2), U-shaped valleys (4), Open slopes (6)	Medium	
	Upland drainage (3), Upper slopes (7), Local ridges (8), Midslope ridges (9), Mountain tops (10)	High	
Hydromorphic soil	Yes (1)	High	2
	No (0)	Low	

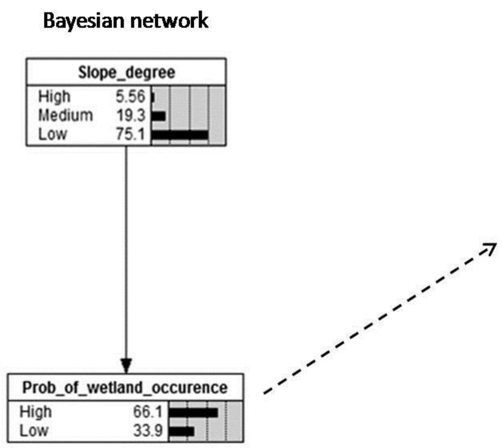
* 1)Variable have been reclassified using Jenks Natural Break algorithm 2) Variables have been reclassified arithmetically

4.1.3 Deriving the Bayesian network to predict wetland occurrence

A BN was compiled to derive conditional probabilities of wetland occurrence based on the nine predictor (parent node) variables. The conditional probabilities derived in the BN were dependent on the known information of the parent nodes; the addition of a variable (parent node) to the network altered the conditional probability of the child node (Figures 4.3 and 4.4). Their causal relationship is indicated by the link arrow between the two nodes. The variables were presented in their defined states (high, medium, and low) in the BN. The final BN was a simple network of the selected input variables surrounding the model's objective of modelling the probability of wetland occurrence (child node) (Figure 4.5).

Using Bayes' theorem, the prior probabilities were used to calculate the conditional probability table (CPT) for the child node, which estimated the probability of wetland occurrence based on the input nodes. The prior probabilities were determined during the case file learning phase, where the defined dataset (following section 4.1.2) was incorporated into the BN.

The first step in building the BN was adding a single parent node (*slope (degrees)*) to the BN (Figure 4.3). The prior probability for the state's high, medium and low, for the *slope* variable, were 5.56, 19.3 and 75.1 respectively, which indicated that 75.1 % of the cases in the dataset (wetland and non-wetland) occurred in low sloping areas. The CPT is populated with conditional probabilities derived through the Bayesian algorithm and represents the probability of wetland occurrence given the state of *slope (degrees)*. According to the CPT the *low* state of slope had the highest probability of wetland occurrence (78.19%), and the *high* state of slope had the lowest probability of wetland occurrence (9.99%). This meant that wetlands are more likely to occur on low/gradual sloping areas than in steep sloping areas.



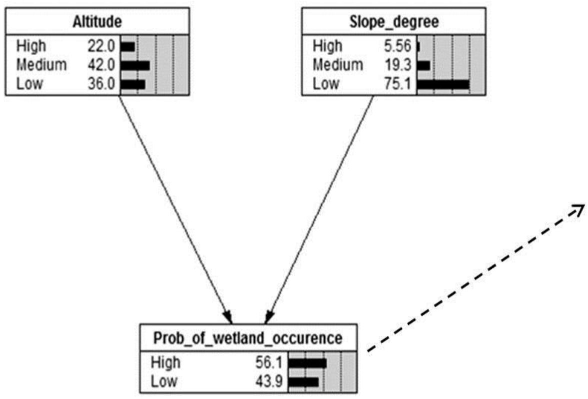
Conditional Probability Table

	Probability of wetland occurrence	
Slope_degree	**High**	**Low**
High	**9.9861**	**90.0139**
Medium	**35.3586**	**64.6414**
Low	**78.1936**	**21.8064**

Figure 4.3: Diagram illustrating the building of the Bayesian network used in calculating conditional probabilities.

The next step in building the BN was adding more parent nodes to the BN. The *altitude* variable was added to the network, and the addition of new information altered the CPT. The probability of wetland occurrence was dependent on the combination of states of both parent nodes. The CPT therefore grows exponentially (to the power of 'the number of states present in the new node'), representing a table of all possible scenarios between *altitude* and *slope,* with a corresponding conditional probability (Figure 4.4). The combination of variable states represents a single scenario, and the CPT in Figure 4.4 has nine possible scenarios. Therefore, if the *altitude* variable had a *medium* state and the *slope* variable had a *low* state, then the probability of wetland occurrence would be 61.44%. The highest conditional probability of wetland occurrence (78.19%) occurred when the *altitude* variable state was *high* and the *slope* variable state was *low*.

Bayesian network



Conditional Probability Table

		Probability of wetland occurrence	
Altitude	Slope_degree	High	Low
High	High	9.98613	90.0139
High	Medium	35.3586	64.6414
High	Low	78.1936	21.8064
Medium	High	13.3858	86.6142
Medium	Medium	36.4599	63.5401
Medium	Low	61.4383	38.5617
Low	High	5.66251	94.3375
Low	Medium	15.8913	84.1087
Low	Low	65.3203	34.6797

Figure 4.4: A Bayesian network illustrating two parent nodes and one child node.

In the final BN, all the variables were added as parent nodes in the network (Figure 4.5). This meant the probability of wetland occurrence was dependent on all nine variables and their different states. This resulted in a CPT with 13 122 different possible variable state combinations or scenarios, and each scenario had a conditional probability (see Appendix B, Table 6.2.3 for an illustrative example). The final modelled probability layer was generated by making these conditional probabilities in the CPT spatially explicit.

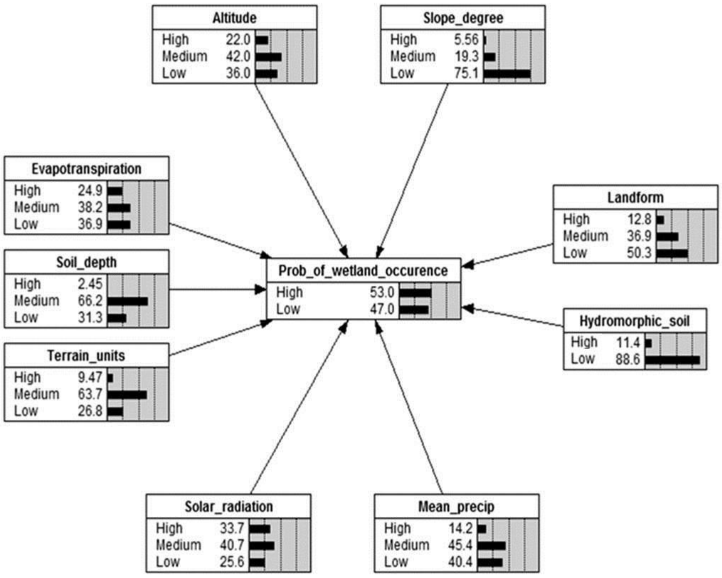


Figure 4.5: Final Bayesian network illustrating nine parent nodes and one child node.

4.2 The final modelled probability layer

The final output of the method was a raster layer at a provincial scale, with the pixel values representing conditional probabilities of being a wetland, at a spatial resolution of 20 m. Probability values in the results were presented as a percentage and not as a value between 0 and 1. Probability values ranged from 2% to 98% (Figure 4.6), where the higher the probability values of a pixel, the greater the likelihood of wetland occurrence.

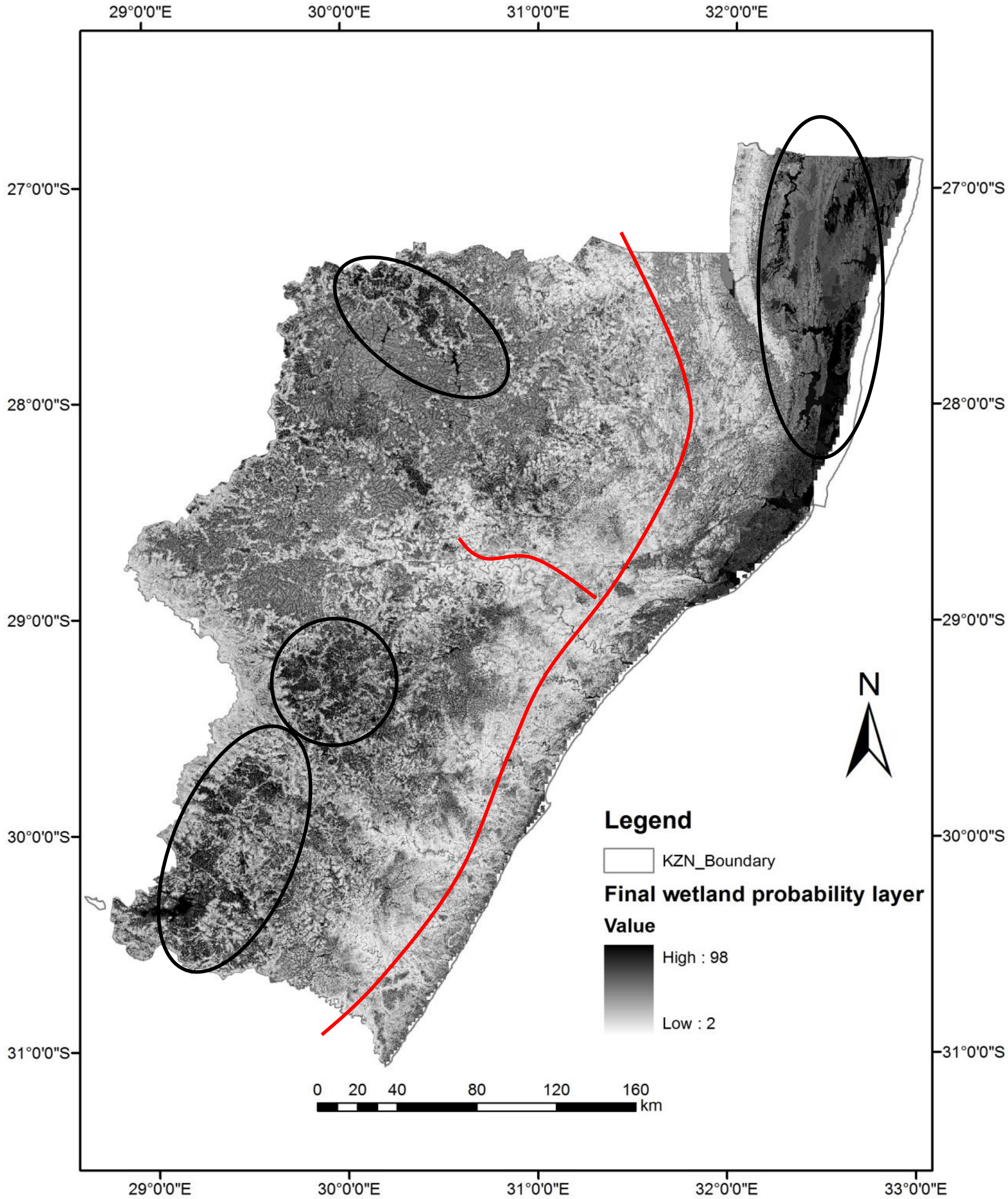


Figure 4.6: Final modelled wetland layer representing the probability of wetland occurrence in KZN. Black circles represent wetland-rich areas and the red lines represent wetland-poorer areas.

The final probability layer was a mosaic of different probability values, and it was evident that the model had identified multiple wetland-rich areas in KZN, represented by the darker shades of grey (Figure 4.6). Wetland-rich areas in KZN were easily identifiable from the modelled surface. These areas were (black circles in Figure 4.6): the northeastern part of KZN known as the Zululand coastal plains and Makatini flats (27°5'S 32°3'E); the northwestern part of KZN covering the upper reaches of the Tugela Basin; and the western and southwestern areas known as the Midlands (29°29'S 29°50'E) and Drakensberg foot slopes (30°19'S 29°7'E to 29°12'S 29°45'E). Evident from Figure 4.6 is that there are areas in KZN that are not so well endowed with wetlands. The red line in Figure 4.6 runs from 29°54'E 30°46'S in the south of the province to 31°40E 27°35S in the north of the province, and branches slightly into the west; this highlights the area with a lower overall probability of wetland occurrence.

In Figure 4.7, the modelled probability layer was reclassified to highlight areas with probability values above 60%, above 80%, and above 90%; this helped to identify high wetland probability areas across KZN. The distribution of high probability values was not evenly spread across KwaZulu-Natal province, and it was only when the modelled layer was reclassified to these probability classes that these patterns of wetland distribution across the province emerged. The modelled probability values below 60% accounted for over 78% of the total area of KZN (Table 4.3). The probability values above 80% accounted for only 6% of the total area of KZN (Table 4.3). This meant that a 5 520 km^2 area had a probability of wetland occurrence above 80%. The trends in Figure 4.7 reveal that with the increase in probability value there was a decrease in the area predicted. Table 4.3 and Figure 4.7 provided a good understanding of the distribution of probability values across the province of KZN.

Table 4.3: Percentage area of KZN covered at increasing probability thresholds

Probability class (%)	Area (%)
2-98	100
5-98	98.58
10-98	91.83
15-98	83.01
20-98	74.20
25-98	67.96
30-98	61.57
35-98	54.69
40-98	48.53
45-98	41.25
50-98	36.08
55-98	28.72
60-98	21.58
65-98	16.93
70-98	12.51
75-98	8.84
80-98	6.01
85-98	3.75
90-98	2.07
95-98	0.67

The modelled probability layer that was reclassified to show probability values greater than 60% illustrated a possible overestimation of probable wetland areas across the province (Figure 4.8, A). Probability values above 60% accounted for approximately 20 000 km^2 of the total provincial area, far exceeding the current KZN wetland area estimate of 4 200 km^2 (Rivers-Moore and Cowden, 2012). The modelled probability layer that was reclassified to illustrate values greater than 90% (Figure 4.8, C) represented areas with a very high probability of wetland occurrence, and major wetland systems clearly stood out (i.e. the Pongola floodplain, the Muzi swamps, and the Great Mkhuze swamp found on the northeastern coastal plains of KZN).

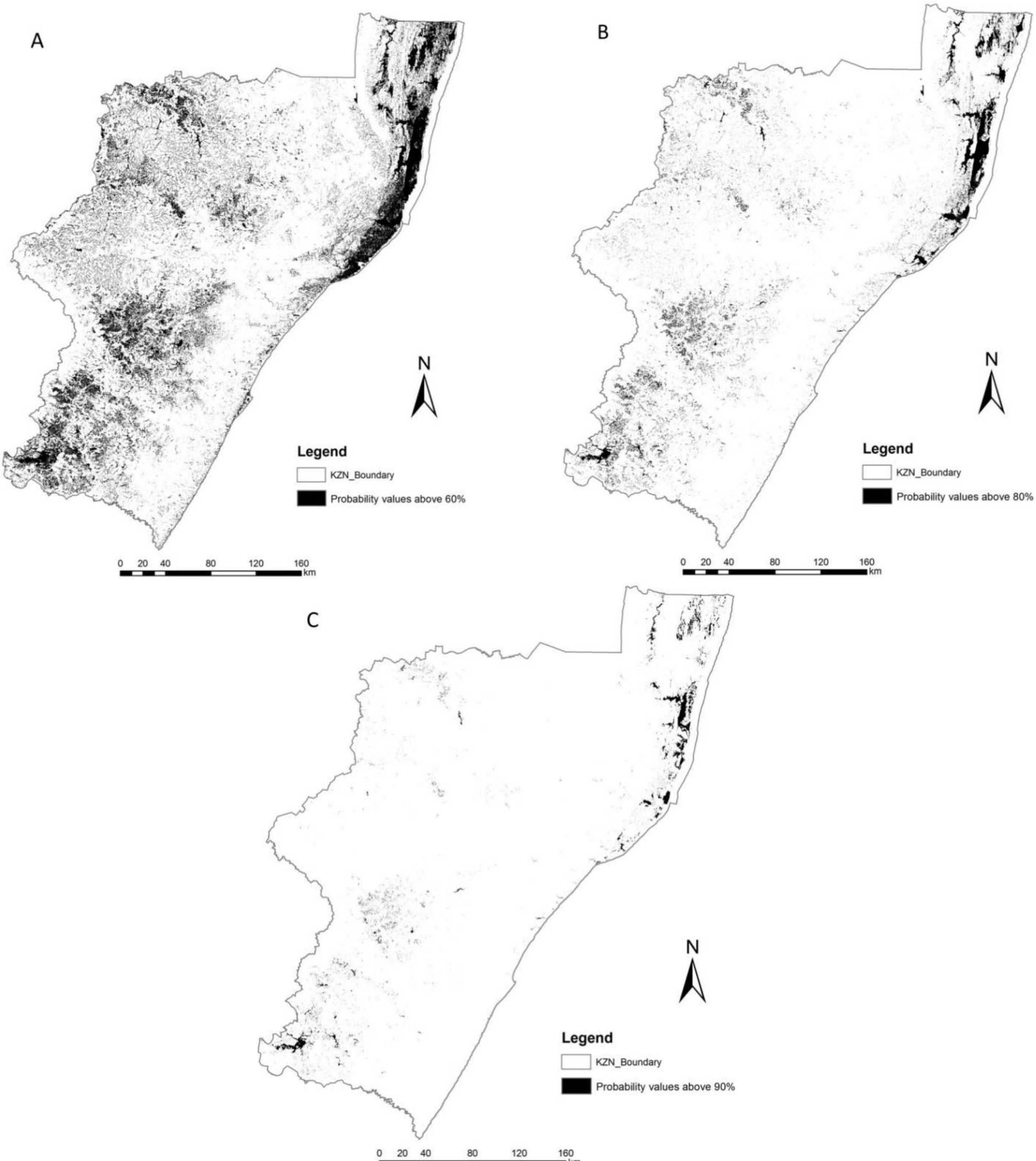


Figure 4.7: Final modelled probability layer reclassified to show probability pixel values above 60 % (A), above 80% (B), and above 90% (C).

4.3 Model Verification

Receiver Operating Curve (ROC) analyses performed using the test wetland dataset indicated that the AUC for the predicted probabilities of the BN model were 0.853 (Std. Error= 0.00287; Cl95% 0.847 to 0.858). An AUC of 1 would indicate perfect prediction, whereas an AUC of 0.5 would indicate completely random binary prediction; therefore, an interpretation of this result is that the BN modelled probability layer predicts wetland areas very well (Fig. 4.8) and indicates the average value of sensitivity for all possible values of specificity (Zhou *et al.*, 2002). The ROC analysis determines that the criterion model probability threshold (at which probability both the sensitivity (71.6) and specificity (81.2) are the highest as a pair) was greater than 0.60 (i.e. >60%). This meant that if the probability values were split into binary classes of wetland and non-wetland areas, then the 0.60 (or 60% probability value) would be the ideal split to maintain good predictability of wetland and non-wetland areas.

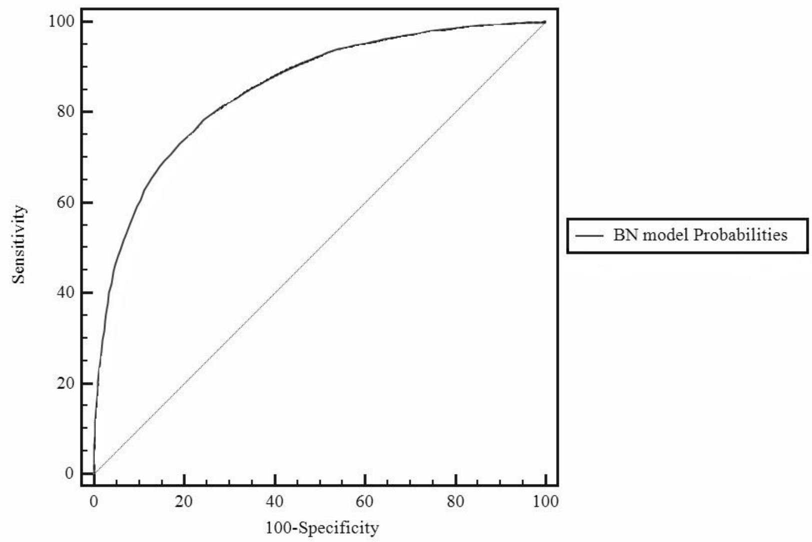


Fig. 4.8: The ROC curve illustrating the prediction accuracy of the BN.

4.4 Model Assessment

4.4.1 Visual assessment of the final modelled probability layer

The locations of the study sites reflect relevant and important areas in KZN (Table 4.4). The 1:200 000 figures of the study sites illustrated the modelled wetland layer with the delineations of the current KZN wetland layer for comparison purposes. (Appendix C, Figures 7.1–7.5 provides more detail on the visual assessment results).

Table 4.4: A summary of key findings in the five study sites chosen for visual assessment

Study Area	Geographic Area	Key findings
Site 1	Makatini flats and Zululand Coastal plains in Northern KZN	• The model clearly identified the Muzi swamp and the Pongola floodplain. • Good visual agreement with the 2011 KZN wetland layer, but significantly more wetland area predicted in the south of the Pongola system.
Site 2	Upper reaches of the Buffalo River in the Tugela basin, northwest KZN	• The model identified the major wetlands in the study site. • The model identified drainage lines (riparian areas) as highly probable wetland areas. Potentially a good indicator of historic connectivity between freshwater ecosystems.
Site 3	Mfolozi swamp, southern tip of Greater Mkuze Swamp System which is St. Lucia, north coast of KZN	• It is difficult to measure the accuracy of wetlands predicted in the study site due to the patchy mosaic of wetland systems found in coastal plains. • However, the model had predicted wetland area well in a highly modified agricultural environment and complex coastal system. There was good visual correlation with the 2011 KZN wetland layer. • A major limitation was the pixelated and undefined coastal boundary.
Site 4	The Midlands of KZN	• The model correlated very well visually with the wetlands identified in the 2011 KZN wetland layer in this area. • The 2011 KZN wetland layer gives a fragmented overview of wetlands in the area, where the model has predicted greater connectivity between the wetland systems • Clearly identified the Mgeni Sponge system found on the Impendle block.
Site 5	Lower Mzimkulu catchment near Port Shepstone, south coast of KZN	• The majority of probability values in this area were <40%, as a result of the undulating relief of the area, which is non-conducive to wetland formation. • The model had predicted significant potential wetland areas in an agriculturally modified landscape.

4.4.2 The final probability layer within historical wetland boundaries

A closer assessment of the modelled wetland probability layer in three historically delineated wetland systems revealed that the modelled wetland layer has the potential to predict historical wetland extent (Table 4.5). When the probability values were reclassified to illustrate probability values of 60% and greater through five classes, it was evident that the higher probability values (90–98%) are found in the heart of the wetland systems (Figures 4.9–4.11). In reality, wetland extents have different wetness zones (i.e. temporarily, seasonally, and permanently wet). A probable explanation for the results seen in this assessment are that the modelled probability layer has portrayed these wetness zones through the compilation of probability values within the historical extent of a wetland (i.e. lower probability values represent temporarily wet zones of a wetland, while higher probability values represent permanently wet zones of a wetland).

The Pongola floodplain is approximately 15 000 ha and is characterised by low-lying adjacent areas with almost 100 off-channel depressions (Begg, 1989). In Figure 4.9, it is evident that probability values above 80% predict the majority area of wetland system (9 050 ha), while probability values between 60 and 80% predict the adjacent areas (2 290 ha). The adjacent areas of the floodplain retain floodwater for varying lengths of time, depending on natural seasonal events and artificial releases from the dam, which could justify the difference in probability values from the main stem area to the adjacent lower lying areas. The average probability percentage predicted in the Pongola floodplain was 76%.

Table 4.5: Analysis of areas within the mapped historical wetland boundary in relation to modelled probabilities for three case study wetlands with mapped wetland boundaries

Probability Class	Pongola floodplain (ha) (Begg, 1989)	Mfolozi floodplain (ha) (Begg, 1989)	Mgeni vlei (ha) (Begg, 1989)
0-60%	3 850	1 520	200
60-70%	470	1 970	140
70-80%	1 820	2 440	270
80-90%	3 350	7 560	390
90-98%	5 700	6 670	1 070
Total ha:	15 190	20 160	2 070
Average Prob. Value	76%	82%	84%

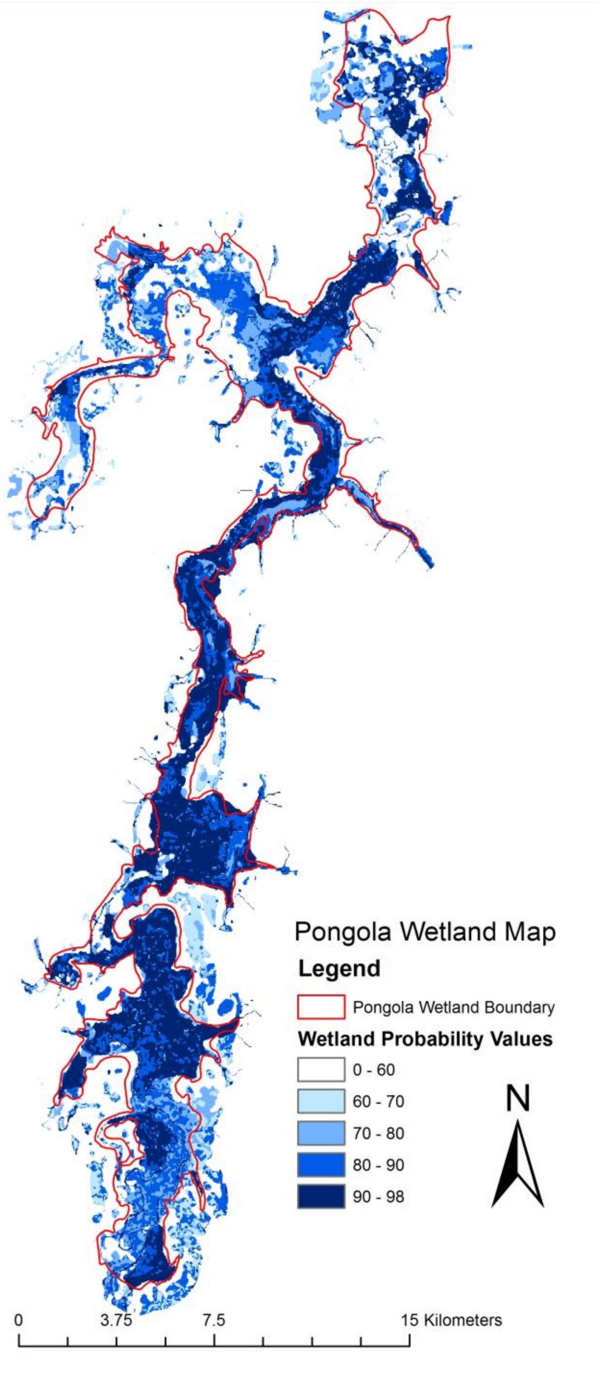


Figure 4.9: The modelled probability layer within the historical boundary of the Pongola floodplain. The red line shows the wetland boundary mapped independently of this study (Begg, 1989).

The Mfolozi floodplain is a highly modified area, with the main modification being from natural vegetation to sugarcane agriculture, but this floodplain wetland occupies over 20 000 ha and still constitutes one of the most important wetland systems in KZN (Begg, 1989). The modelled probability layer in Figure 4.10 generally resembles the shape of the historically delineated wetland, with values of 80% and greater accounting for 70% of the entire wetland (Table 4.5). The adjacent channels running into the main body of the wetland are represented by lower probability values (i.e. 60-90%).

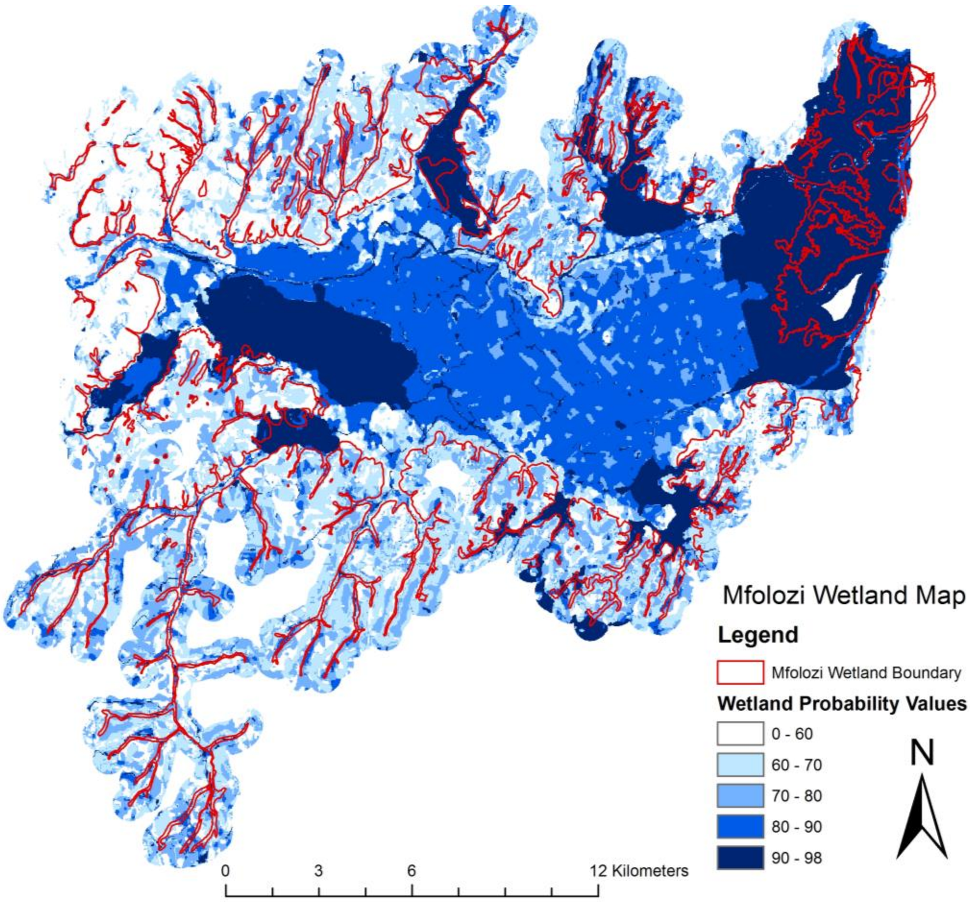


Figure 4.10: The modelled probability layer within the historical boundary of the Mfolozi floodplain. The red line shows the wetland boundary mapped independently of this study (Begg, 1989).

The Mgeni vlei is situated at the source of the Mgeni River, which is recognised as one of the most important river systems in South Africa, and certainly as the most important river in KZN, as its catchment produces approximately 20% of the nation's GDP (Jewitt and Kotze, 2000). The modelled wetland layer has predicted this area well with high probability values (>90%) accounting for more than 50% of the area within the delineated historical boundary. The extent of the wetland predicted by the model (>60%) is greater than the extent of the delineated historical boundary (Figure 4.11). This may account for numerous other wetlands varying in size that surround the Mgeni vlei, as mentioned by Begg (1989). The wetlands in this area owe their existence to the underlying geology, namely Karoo dolerite, which not only exercises strong structural control over drainage in the area, but also contributes greatly to the retention and storage of water in the area (Begg, 1989).

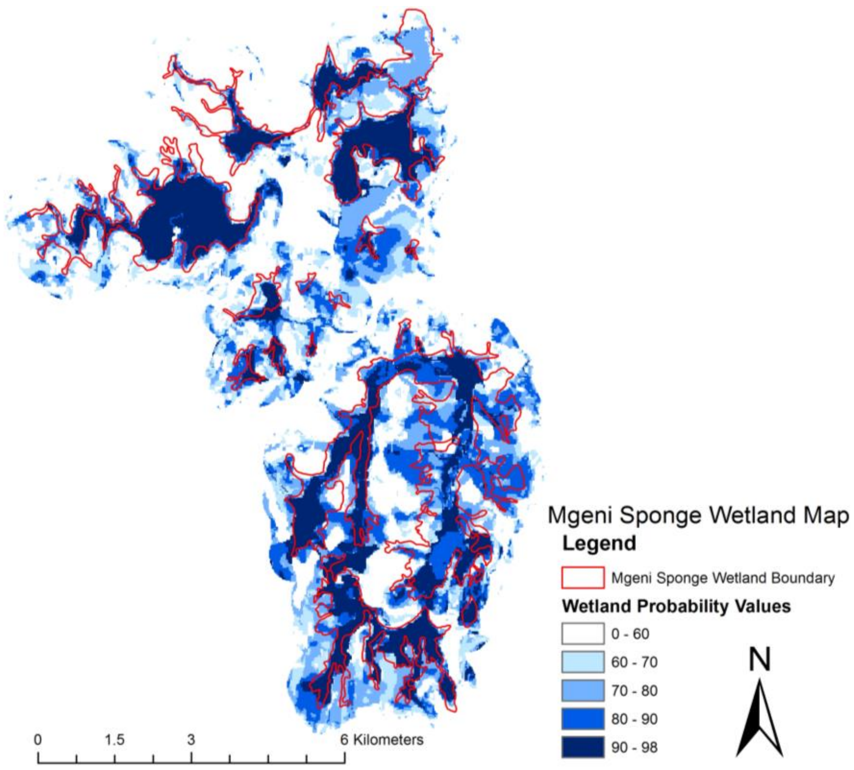


Figure 4.11: The modelled probability layer within the historical boundary of the Mgeni vlei. The red line shows the wetland boundary mapped independently of this study (Begg, 1989)

4.4.3 Assessment of extent and occurrence of wetlands

Layers used in all assessments: *Modelled probability layer* **Layer used in assessment of wetland extent**: *TEST 2011 KZN wetland dataset (Scott-Shaw and Escott, 2011)* **Layer used in assessment of wetland occurrence**: *1 000 randomly generated points across 20 probability classes, SPOT imagery, Google Earth* **Layer used in assessment of wetland type**: *NFEPA wetland layer (Nel et al., 2011b)*

The accuracy in predicting wetland occurrence (the presence or absence of a wetland) increased with the increase in modelled probability value (dotted black line, solid red trend line, Figure 4.12). Therefore the higher the probability value, the higher the likelihood of correctly identifying wetland occurrence. In Figure 4.12, probability values of between 2% and 40% predicted no wetland, as expected with low probability values, but probability values between 90% and 98% predicted over 90% of the points as wetlands correctly.

In contrast with the increase in accuracy in predicting wetland occurrence, there was a decreasing trend in accurately defining the mean area of wetland extent covered, with the increase in modelled probability value (hashed dark grey line) (Figure 4.12). The accuracy in defining the true wetland extent area decreased with the increase in probability value intervals. Probability value intervals of 2‒98% correctly predicted 100% of the wetland extent as expected, because the probability interval contained all probability values. Probability intervals of 50‒98% correctly predicted approximately 90% of the wetland extent, while probability intervals of 90‒98% only correctly predicted approximately 10‒30% of the wetland extent (Figure 4.12). This trend was observed because, as seen in Figures 4.9–4.11, the average wetland extent contains a range of probability values and therefore a restriction of a higher probability value results in a decrease in the average wetland extent area predicted, because low- to mid-range probability values (that still make up a certain percentage of the wetland extent) are excluded.

Evident from this assessment was that a higher range of probability values in the modelled wetland layer would result in a higher accuracy in predicting wetland locality, but not necessarily the full wetland extent. By illustrating both assessments on a single graph against the increase in probability value (Figure 4.12), it was evident that there was a trade-off between the accuracy of predicting wetland extent and the accuracy of predicting wetland occurrence. This trade-off is important in terms of the applicability of such a modelled layer (i.e. using the modelled layer to create a single wetland map) (Figure 4.12). Creating a single wetland map from the modelled wetland layer would require the user to objectively determine what probability values would constitute a wetland map. From Figure 4.12 it is possible to obtain an accuracy of extent and occurrence at any probability value. For example, at a probability value of 90‒98%, the maximum accuracy of wetland occurrence will be 88%, but the maximum accuracy in wetland extent will only be 33%. As an optimised trade-off, the point "A" in Figure 4.12 provided the optimal cut-off probability (D = >79%), at which the accuracy in wetland extent (B = 63%) intersects with the accuracy in wetland occurrence (C = 63%).

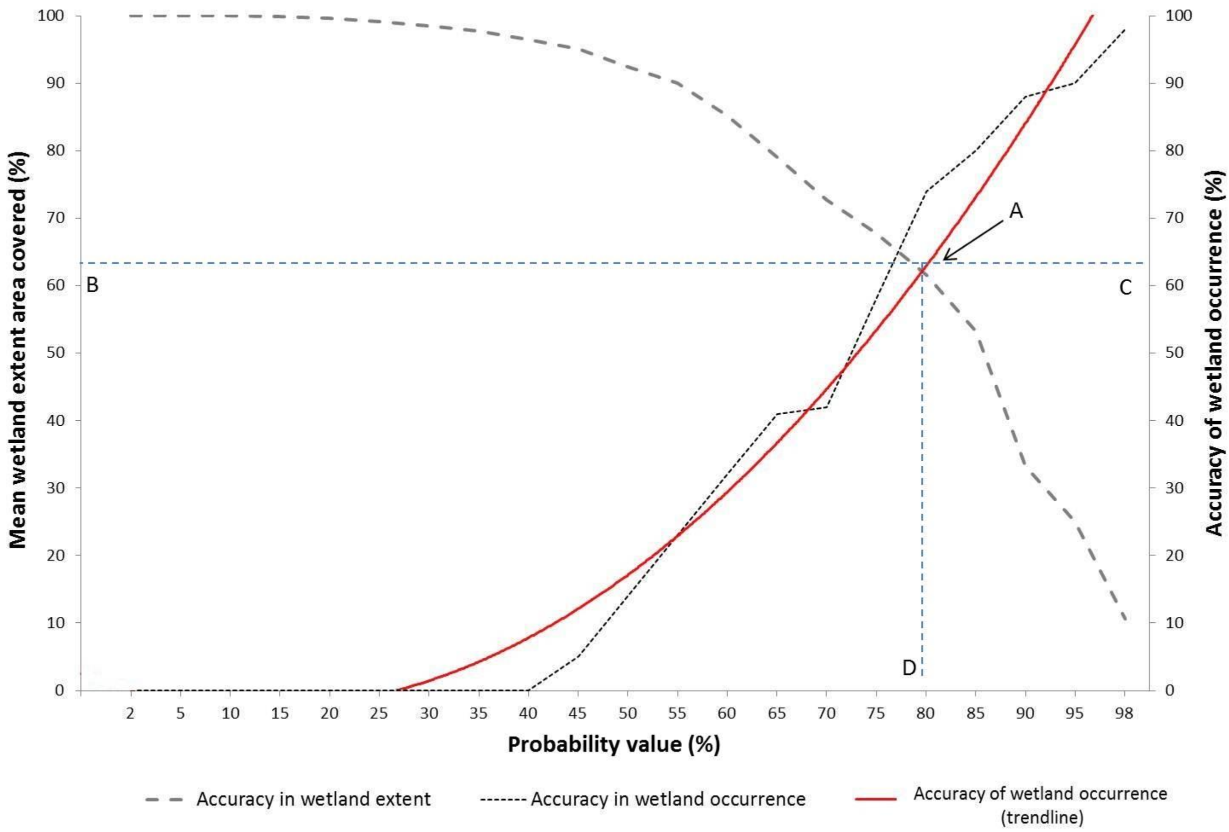


Figure 4.12: Modelled wetland layer accuracy in terms of wetland extent and wetland occurrence with the increase in probability percentage.

The same assessment of the model's ability to predict wetland extent was performed across six different HGM wetland types using the NFEPA wetland layer. This assessment revealed that the model's accuracy varied across the seven different wetland HGM types (Figure 4.13).The point at which the accuracy in predicting wetland extent intersects the accuracy in predicting wetland occurrence (e.g. "A" in Figure 4.13) represents the *optimised trade-off point* between the two accuracies for each wetland type. It was therefore assumed that the intersection point that occurred at the highest point along the red line (accuracy of wetland occurrence) was the wetland type that had been predicted with greatest accuracy by the modelled probability layer. Therefore, the wetland type 'flats' (solid blue line) and 'channelled valley bottoms' (solid yellow line) were the worst predicted type, where the optimal trade-off point resulted in a predictive accuracy below 50%. Depressions (solid green line) were the best predicted wetland type, where the optimal trade-off point resulted in a predictive accuracy greater than 70%.

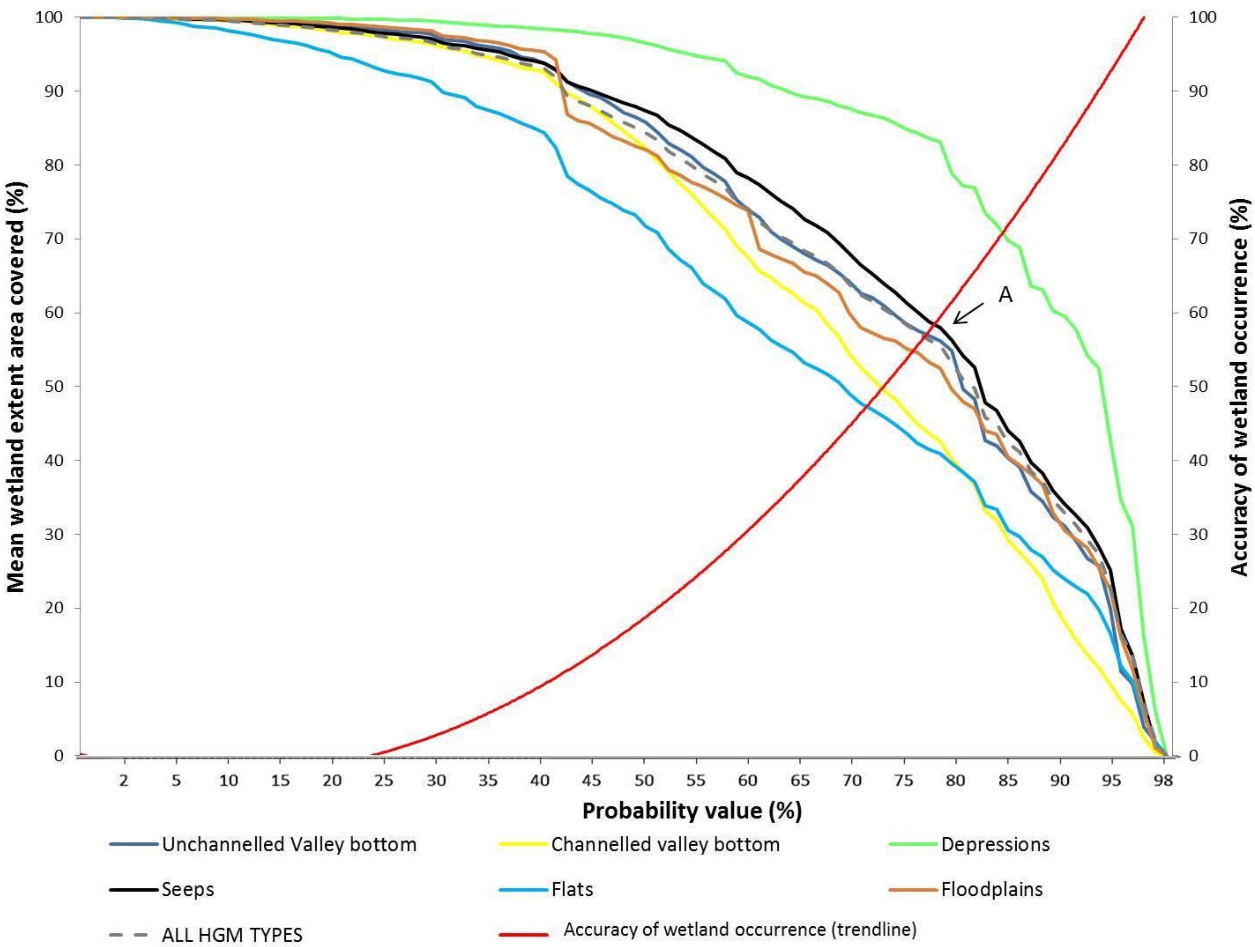


Figure 4.13: Model accuracy in terms of wetland extent and wetland occurrence with the increase in probability percentage for each wetland HGM type.

CHAPTER FIVE

5 Discussion

The formation of wetlands is typically constrained by the complex interactions of various environmental components, allowing the time and space for the wetland to form and function (Batzer and Sharitz, 2006: Ellery *et al.*, 2008; Mitsch and Gosselink, 2000). Defining and understanding the interactions and relationships of these environmental components is difficult. Although the literature review highlights the key factors involved in the formation of hydrologic conditions found in wetlands, there is a high degree of complexity in determining what combination of factors results in a setting conducive to wetland formation (sections 2.1.2–2.1.3). This study has provided a method and approach that uses existing environmental information in conjunction with Bayesian probability modelling to predict the probability of wetland occurrence at a provincial scale. Assessments of the final modelled probability layer provided support for the possible applicability of such an output. The final result of this study was a raster layer of probability values predicting the likelihood of wetland occurrence across KZN at a resolution of 20 m, thus satisfying the overall aim of the study, which was to model the likelihood of wetland occurrence in KZN using conditional probabilities derived through a BN.

5.1 The distribution of wetlands across KZN

In this study, an examination of the distribution of the predicted wetlands across KZN concluded that high probability wetland pixels occur in a greater density in some areas as opposed to others (Figure 4.7). The model predicted a higher density of wetland occurrence in three major zones around KZN: in the northeastern part of KZN; in the northwestern part of KZN; and in the broad valley bottoms in the foothills of the Drakensberg and into the midlands of KZN.

The northeastern area of KZN is known variously as the Makatini flats, 'Palm zone' Maputaland, and the Zululand coastal plains. This region of wetlands was identified and examined by Begg (1989), who discussed the mosaic of important wetland systems found in these areas, namely the Pongola floodplain, the Muzi swamp, and the Great Mkuze Swamp System. These systems were easily identifiable in the final modelled probability layer (Appendix C, 6.1). The Pongola floodplain formed as a result of the flat topography through which it meanders, and the wetlands formed on the low-lying adjacent areas characterised by pans and off-channel depressions, which retain floodwater for varying lengths of time (Begg, 1989). Muzi swamp occupies an area of approximately 15 000 ha and formed across a sandy plain which has a low relief (<40 m AMSL), subtropical conditions and a high water table. The formation of the Great Mkuze wetland system can be attributed to the formation of dune cordons (formed as a result of a series of marine regressions and transgressions) that represented former shorelines (Ellery *et al.*, 2008; Patrick and Ellery 2006). These cordons impeded the flow of the Mkuze system, shifting it from a seaward-bound easterly orientation to a southerly direction. The subtropical conditions resulting in hot and wet summer months recharge the groundwater, and the dunes' elevated water table relative to the coastal plain causes the inland flow of subsurface water into interdunal depressions, which recharges lakes and wetland swamps (McCarthy and Hancox, 2000). This provides some justification for why the model has predicted this high density of wetlands in the north-eastern region of KZN.

In the Pongola floodplain, the modelled probability layer predicted the historical wetland area well. Begg (1989) emphasised that the Pongolapoort Dam (also known as Jozini Dam), which was constructed in the 1960s, had profound effects on the floodplain. The floodplain is now dependent on artificial releases from the dam as opposed to natural seasonal events. Mismanagement of these artificial releases over a period could potentially dry out certain areas of the floodplain. Besides *hydromorphic soil*, the model used eight surrogate predictors which were all completely unrelated to the development of the Jozini Dam but were important factors in the formation of wetlands. Therefore, the model predicted wetland area regardless of artificial influences, and based only on topographic and climatic variables in the area. The advantage was that the model predicted the wetland area in the Pongola floodplain regardless of whether there is a dam or not, and has therefore potentially predicted wetland area prior to construction of the Pongolapoort dam, hence predicting its historical extent (Figure 4.9).

The northwestern region of KZN forms the uppermost reaches of the Tugela catchment and is home to important wetland systems such as the Wakkerstroom Vlei, Groen Vlei, Boschoffs Vlei, Blood River Vlei, and Stilwater Vlei. The model identified a dense distribution of wetlands in this area. The formation of these wetlands in this area is as a result of a combination of components, namely the slope and width of the valley bottoms, the geology and distribution of soils within the landscape, and the rainfall in the area. As mentioned in the literature review and description of the study area (sections 2.1.3 and 3.1), the gradual slope of the valley bottoms in this area is a result of the eastward back-tilting associated with the upliftment of the old (interior) African land surface approximately 20 million and five million years ago (Begg, 1989; Ellery *et al.*, 2008; Joubert, 2009). The high annual rainfall in the upper reaches of the Tugela catchment creates large volumes of run-off into these flat wetland areas. Wetlands in this area are a result of erosional and depositional processes, and from a geological perspective the landscape is dominated by Karoo dolerite outcrops, which play a fundamental role in the formation of wetlands. The visual assessment in that area indicated that the model had successfully identified wetland systems in this area, but also highlighted drainage lines as probable wetland areas (Appendix C, Figure 6.2). These drainage lines could signify important connectivity between freshwater ecosystems (this is discussed later).

Another high probability wetland area found in the modelled layer spread from the broad valley bottoms below the Drakensberg mountains into the Midlands of KZN. This area is known as the 'East Griqualand plains' and includes the Swartberg and Impendle block, moving eastward to the headwaters of the Mvoti catchment (Begg, 1989). The Drakensberg escarpment is considered to be an important catchment area for three major rivers in KZN, namely the Tugela, Mkhomazi and Mzimkulu rivers (Dely *et al.*, 1995). The Impendle block is home to the Mgeni Vlei, which is part of the most important and recognised river system in KZN, the Mgeni River System. The valley bottoms below the Drakensberg and in the Impendle areas are richly endowed with wetlands because water running off prominent surrounding peaks onto the gradual sloping valleys saturates the uppermost part of the sediment surface for prolonged periods (Begg, 1989). Sedimentological connectivity within drainage systems (creating favourable conditions for wetlands) may be controlled spatially by 'pockets' of intact valley fill, alluvial fans impinging laterally on mainstem rivers, floodouts impinging longitudinally on valley floors, and downstream resistant rock bands and their effect on valley width (Grenfell *et al.*, 2009). These deposition environments found in and around the Drakensberg, together with the climatic conditions, make the area highly conducive to wetland formation: high rainfall, gradual sediment-trapping slopes and Karoo dolerite key points are the dominant features that form wetlands. The wetlands in these areas have several cut-off meanders and backswamps, which indicate processes such as flood peak reduction, run-off interception and sediment trapping, which are important wetland functions in these areas. The wetlands in this area are considered to have an important role in regulating the quality and quantity of water from these high-altitude catchments (Dely *et al.*, 1995). A closer visual assessment of the historical boundary concludes that the area is well predicted by the model and supported by a high average probability value (Table 4.5; Figure 4.11).

5.2 Verification of the modelled probability layer

The assessments in the study focused on determining how well the modelled probability layer predicted wetland areas. The modelled probability layer of wetland occurrence was a mosaic of probability values which ranged from 2–98%. Although there are some wetland-rich areas in KZN, only 6% of the total province area had a wetland probability value exceeding 80%. This translates into a total area of 5 520 km^2 in KwaZulu-Natal, which exceeds the more conservative figure of 4 200 km^2 established by Rivers-Moore and Cowden (2012). The ROC analysis results indicated that the model predicts wetlands well, and therefore this difference in area could be the result of incomplete mapping efforts throughout the KZN province, which stems from the lack of resources for mapping.

The visual assessments revealed that the final probability layer has the potential to identify new wetland areas, and that a mosaic of probability values can be found in a delineated wetland extent. Wetlands typically have different wetness zones within a single wetland system or extent area. Relating this back to the literature, Marnewecke and Kotze (1999) mentioned in DWAF's 'Guidelines for delineation of wetland boundaries and zones', that wetlands have a transitional nature, where the boundary of the wetland is often not clearly apparent. A wetland has zones of wetness, ranging from permanently wet, to seasonal, to temporarily wet. The modelled probability layer does not explicitly indicate the wetland boundary but shows a continuum of increasing probability values from areas of low probability to high probability. The average wetland extent was primarily represented by high probability values in the modelled probability layer, as well as a mosaic of mid and lower probability values (Table 4.5). It could be interpreted that lower to mid probability values express zones of uncertainty, which could potentially be temporary or seasonal wetland areas found within a wetland system.

The 2011 KZN wetland layer has only mapped the extent of wetlands that are visually discernable in satellite imagery, therefore wetland systems that have been modified into dams, cultivated land or urban areas are not mapped (Escott pers. comm.) The result is a KZN wetland layer that is fragmented in nature because of anthropogenic factors that have produced disconnected ecosystems through land cover modifications. In comparison, the modelled probability layer predicts the presence of wetland systems regardless of any visual (spectral) differences in a satellite image, and rather uses information based on topographic, climatic and hydrologic variables within known wetland extents. In the visual assessments, the results not only indicate the probable wetland area, but the map indicates highly probable connectivity between wetland ecosystems. In Appendix C (Figures 6.2 and 6.4) the current wetland layer portrays a rather fragmented wetland picture, while the modelled probability layer illustrates greater historic connectivity between the different wetland systems. The importance of this can be related back to the literature review (section 2.1.5), where Nel *et al.* (2011a) and Dudgeon *et* *al.*(2006) mention that the challenge in freshwater conservation planning is relating wetlands to upstream drainage networks, surrounding riparian zones, and downstream reaches; focusing on representing biodiversity in isolated areas is conceptually flawed. If the probability model shows the strong potential to identify historic connections between wetland systems then it is possible that the model could identify historic wetland clusters, which are important groups allowing for ecological processes such as the migration of frogs and insects between ecosystems (Nel *et al.*, 2011a).

5.3 The applicability of the modelled probability layer

Wetland mapping approaches are dominated by satellite imagery classification approaches, which involve a range of different complex methodologies in which spectral ratios and values are classified to identify wetland areas (Li and Chen, 2005). Many methodologies draw on the support of ancillary data to improve the accuracy of mapping and classifying wetlands (Kulawardhana *et al.,* 2007). Li and Chen (2005) used altitude and slope as two important determinants of where wetlands are likely to occur in their rule-based wetland mapping approach. Bwangoy *et al.* (2010) combined the use of derived topographical indices with optical and radar remotely sensed data to map wetlands in the Congo Basin, emphasising that topographic information significantly contributed to the classification-tree procedure in mapping wetlands. Wright and Gallant (2007) researched the value of using ancillary environmental data in improving the classification of wetland areas using remote sensing, and concluded that error rates dropped incrementally as DEM-derived terrain variables and other ancillary layers were added in the classification of wetlands. Evident from the studies briefly mentioned above is that ancillary data can add value and prove instrumental in the success of wetland mapping approaches. Topological entities are commonly used as ancillary data and have been valuable in the success of numerous wetland mapping techniques (Ibrahim, 2009; Islam *et al.*, 2008; Pantaleoni *et al.*, 2009). Five of the nine variables used in the model were topological variables (*altitude*, *slope*, *terrain units*, *landform*, and *soil depth*) and the verification of the modelled probability layer concluded that it predicts wetland areas well; therefore, there is value in mentioning that the modelled probability layer derived in this study has the potential to serve as an important ancillary data layer to further improve wetland mapping accuracies in KZN by identifying more wetland areas, and hence improving the current wetland layer for KZN.

The KZN wetland layer was updated by comparing the entire wetland dataset to SPOT 5 2008, a relatively moderate-resolution imagery (Escott pers. comm., 2011). The shortcomings of this process are that it is tedious, and that numerous wetlands are missed because they do not reflect obvious differences relative to the surrounding vegetation. The modelled probability layer has the potential to earmark potential wetland sites, which could expedite the process and reduce the possibility of missing a wetland.

In terms of wetland management and conservation, the final modelled output has the potential to predict wetlands in already modified areas. The approach does not use vegetation structure or spectral entities in its prediction; therefore, the final model output predicts the likelihood of wetland occurrence regardless of any land cover modification. Boyd Escott (pers.comm., 2011) believes that the south coast of KZN is weakly mapped in terms of its wetlands and that there is room for improvement. In Figure 6.5 (Appendix C) it is apparent that this is the case. The south coast of KZN has undulating terrain and settings not highly conducive to wetland formation, which may account for the low number of wetlands identified, but the main reason for this is that many wetland areas along the south coast have been modified into sugarcane cultivation. The highly probable wetland areas identified by the model (and not by the 2011 wetland layer) could potentially be the extents of historic wetlands that have been modified into sugarcane, but to confirm this will require further research and field verification. However, if this is the case, then the modelled probability layer could prove an essential tool in wetland management and rehabilitation. Closer visual assessment in section 4.4.2 further confirmed the model's success in potentially identifying the historical wetland extent. The eThekwini Municipality has already incorporated the final probability model to help locate and map wetlands within the municipal boundaries. eThekwini Municipality believes that the wetland probability layer is valuable in earmarking high-density wetland areas that should be investigated further through a field visit. The modelled probability layer holds the potential to locate and identify historical wetlands, especially in the urban modified areas (Botes pers. comm., 2012).

Wetlands occur in positions in the landscape where a positive water balance exists frequently and long enough for it to favour biota adapted to hypoxic soil conditions and standing water (Walters pers. comm., 2012). The modelled probability layer appears to represent these areas quite well. The modelled layer could therefore guide the process of identifying wetlands through the classification of satellite imagery. Acting as a mask, the model could minimise the misclassification of pixels or the high number of errors caused by spectral confusion. Supervised classification of satellite imagery involves using a training set of pixels in order to identify similar pixels; a simple binary classification of the modelled layer into wetland and non-wetland areas (above a certain threshold) could provide the necessary training sites in order to identify new wetland areas more accurately.

The modelled probability layer could also be used as a standalone wetland map. Creating a standalone map would also involve reclassifying the wetland layer into a binary dataset using a threshold that best represents wetland areas. This threshold is dependent on the user. As section 4.4.3 illustrates, there is a trade-off between the accuracy of the model's prediction of wetland extent and the accuracy of its prediction of wetland occurrence, which needs to be taken into account when creating such a map. For example, a user who is interested in using the model to identify new wetland areas, and who is less concerned about the model's ability to predict wetland extent, would opt for a higher cut-off probability value, producing a wetland map with a higher accuracy in predicting wetland occurrence and a lower accuracy in modelling the correct wetland extent.

The relationship between the two accuracies and probability value (Figure 4.12) could be used to assess the maximum accuracy of the model at any probability value. The optimised trade-off point is the probability value where both the accuracy of wetland occurrence and extent are the same. Importantly, this optimised trade-off point differs between wetland types (Figure 4.12). Thus, the model predicted depressions, which have the highest optimised trade-off point of all wetland types, with the highest accuracy, and channelled valley bottoms and flats with the lowest accuracy. Inferences could be made that depression-type wetland systems are usually fed through groundwater discharge, precipitation or surface runoff and have little or no flow out of the system (with outflow usually only taking place through evaporation or groundwater recharge), and are therefore almost always wetlands (Collins, 2005). On the other hand, wetlands characterised by channelled valley bottoms and flats flow through fluvially-deposited sediment. The soil deposited can range from poorly drained to excessively drained, and therefore the soil depositions found in the landscape, together with the permanent, seasonal or ephemeral/episodic stream flow, do not always result in a wetland system (DWAF, 2007). The significance of finding that different HGM types have different accuracies in the model is important in terms of its applicability, but it is recommended that further research be done on why there is this difference in accuracies. A possible explanation for this difference is that some of the input variables used in the model have a stronger correlation to certain HGM wetland types and the biophysical processes that define the different types. Such additional research could provide clarity on why there is a difference in the accuracies of the different HGM types and should be explored further.

5.4 Limitations

The method and approach used to achieve the objective were novel, and the final output of this objective was achieved: indicating wetland areas with high probability values at a resolution of 20 m. The method and approach were successful, but a few limitations were noted in the development of this probability model.

The BN used in calculating the conditional probability table required that the quantitative variables in the dataset be redefined into qualitative states, which resulted in some loss of information. This limitation was also noted by Uusitalo (2007:314), who states that "BNs can only deal with continuous data in a limited manner" and that "discretisation can only capture rough characteristics of the original distribution and we may lose statistical power in the relationship between variables". It is therefore important how the continuous data is discretised because of the risk that the power of the variable data identified using PCA will be lost, with the states now representing the data not being able to accurately define or predict wetland location. "There is no satisfactory automatic discretisation technique or method for Bayesian networks" (Uusitalo, 2007:314). Future research adopting this approach should focus on working out defining value ranges within continuous data that signify important breaks within the data, so that generalisation (caused by discretisation) is avoided to some extent.

Thought should also be given to the potential limitations posed by the use of the 2011 KZN wetland layer in this study. As mentioned in the methods, the coverage draws on multiple sources with varying capture scales and temporal scales. Although a 'cleaning' process was initiated to standardise the coverage, the coverage is bound to have errors and inconsistencies based purely on its large extent. Another concern is the fact that the coverage requires extensive ground-truthing (Escott pers. comm., 2013), as there are no reports on the accuracy of the coverage in terms of the extent and location of its wetlands. In this study an assumption was made that the wetland coverage used was accurate in terms of its wetlands mapped, and the rough accuracy assessment performed on the KZN wetland coverage only provides a confidence estimate. Although the modelled probability layer performed well in this study, ways in which the KZN wetland layer could have influenced the output of this study should still be considered.

5.5 Recommendations

Although the verification assessments in this study provided valuable insight into the model's accuracy in predicting wetlands, the assessments were desktop-driven. It is therefore recommended that the final wetland probability layer be further verified using ground-truthing exercises. Its exact applicability will only be realised when the model is used in contemporary projects in the field of wetland conservation and management, from which conclusive statements can be drawn regarding the strengths and weaknesses of this modelled probability layer.

In this study the probability of wetland occurrence was successfully modelled at a provincial scale. It is recommended that the model be verified at various scales (i.e. provincial and municipal scales), and that the strengths and weaknesses be defined at each scale. Scale issues may play a role in generalisation of the data caused by the discretisation process, so therefore finer-scale models (i.e. municipal rather than provincial scale) would require different resolutions of class intervals in the predictive variables. This would require the user to rerun the model only within the boundaries of the new project area. The result would be a refined model and probability layer for the project area, and predictions are that the result will be more accurate in predicting the probability of wetland occurrence than when modelled at provincial scale. The incorporation of the modelled probability layer to identify wetland areas in eThekwini Municipality provides a good starting point for assessing the model's applicability at municipal scale, and the feedback and critique provided by eThekwini will be valuable in fully assessing the applicability of the modelled probability layer. Knowledge of scale-dependant accuracies could assist with the adaptation of the method and approach to site-specific areas.

The method and approach used in this study could potentially be adopted nationally; the success of such an undertaking is, however, limited to which input variables are available at a national scale. Attempting this approach on a national scale will only result in further generalisation of the data over a larger domain. This is potentially problematic, as it could result in lowering the predictive power of the model, especially for different HGM types, particularly because wetlands across southern Africa are driven and defined by a variety of different hydrologic, physiochemical, climatic and geomorphologic environments. As reviewed in section 2.1.3, wetlands in South Africa are found across a range of different landscapes, from arid areas with low rainfall and high evapotranspiration in the west, to high water-yield areas along the escarpment, to landscape influenced by subcontinental uplift 20 million and five million years ago. Therefore, an alternative suggestion to a "one-size-fits-all" national model is to develop models for each province or region, or per wetland region (as described by Cowan, 1995), and to piece together the separate probability layers to establish a national coverage. The success and accuracy of the provincial coverages is likely to differ between provinces because of regional differences in the availability and quality of input variables. KZN has a vast variable dataset which proved valuable in the success of the final output as well as the accuracy assessments of the final coverage. The available input variable dataset provided a solid and robust learning case file used to derive conditional probabilities in the BN, which added to the success of this study.

Finally, it is recommended that the Bayesian approach used in this study be extended to provide information on the most likely HGM type associated with a wetland presence probability value. A new modelled probability layer that indicates the likelihood of a wetland being a certain HGM wetland type could be valuable additional information for the model in this study. The idea is that geomorphic and hydrologic surrogates be used in the same general approach described in this study, to determine the conditional probability of a specific HGM wetland type occurring in a specific area in the landscape. This output would be a valuable dataset in classifying the modelled probability layer into probable HGM wetland types. The HGM classification will provide a tool for measuring changes in the functions of wetland ecosystems due to impacts by proposed projects, and restoration, creation, and/or enhancement (Brinson *et al.*, 1997). Additionally, there would be value in exploring the idea of determining/using landscape indicators of wetland condition to create attribute information on the condition of a wetland in a highly probable wetland area (Rivers-Moore pers. comm., 2013).

The data emanating from the above study and the recommendations will help to improve and refine the wetland layer for KZN and could therefore potentially contribute to addressing EKZNW spatial conservation priorities and to improving wetland management in general.

6 References

Abdi, H. and Williams, L.J. 2010. Principal Component Analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2: 1–47.

Adam, E., Mutanga, O. and Rugege, D. 2010. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: A review. *Wetland Ecology and Management*, 18: 281–296.

Aguilera, P.A., Fernandez, A., Fernandez, R., Rumi, R. and Salmeron, A. 2011. Bayesian networks in environmental modelling. *Environmental Modelling and Software*, 26: 1376–1388.

Batzer, D.P. and Sharitz, R.R., eds. 2006. *Ecology of freshwater and estuarine wetlands*. Berkeley, CA: University of California Press.

Begg, G.W. 1986. *The wetlands of Natal part 1: An overview of their extent, role and present status*. Pietermaritzburg: Natal Town and Regional Planning Commission, Report 68.

Begg, G.W. 1988. *The wetlands of Natal part 2: The distribution, extent and status of wetlands in the Mfolozi catchment.* Pietermaritzburg: Natal Town and Regional Planning Commission, Report 71.

Begg, G.W. 1989. *The wetlands of Natal part 3: The location, status and function of the priority wetlands of Natal*. Pietermaritzburg: Natal Town and Regional Planning Commission, Report 73.

Beyer, H.L. 2004. *Hawth's analysis tools for ArcGIS* . Available from: www.spatialecology.com/htools [Accessed 5 July 2011].

Boone, K. 2004. *Bayesian statistics for dummies.* Available from: web.vu.union.edu [Accessed 17 July 2012].

Botes, W. 2012 (November). Environmentalist, eThekwini Municipality. Personal Communication. Durban.

Brinson, M.M., Lee, L.C., Rheinhardt, R.D., Hollands, G.G.,Whigham, D.F. and Nuttler W.D. 1997. A summary of common questions, misconceptions, and some answers concerning the hydrogeomorphic approach to functional assessment of wetland ecosystems: scientific and technical issues. *Society of Wetland Scientists*, 17(2): 16–21.

Bruyninckx, H. 2002. *Bayesian probability*. Belgium: K.U. Leuven, Department of Mechanical Engineering.

Bwangoy, J.B., Hansen, M.C., Roy, D.P., Grandi, G. and Justice, C.O. 2010. Wetland mapping in the Congo Basin using optical and radar remotely sensed data and topographical indices. *Remote Sensing of Environment*, 114(1): 73–86.

Cain, J. 2001. *Planning improvements in natural resources management: Guidelines for using Bayesian networks to support the planning and management of development programmes in the water sector and beyond.* Walingford, UK: Centre for Ecology & Hydrology.

Carletta, J. 1996. Assessing agreement on classification tasks: The Kappa statistic. *Comput Linguist*, 22: 249–254.

Chhokar, K.B., Pandya, M. and Raghunathan, M. 2004. *Understanding environment*. New Delhi: Sage.

Collins, N.B. 2005. *Wetlands: The basics and some more*. Bloemfontein: Free State Department of Tourism, Environmental and Economic Affairs.

Colvin, C., Le Maitre, D., Saayman, I. and Hughes, S. 2007. *Introduction to aquifer dependent ecosystems in South Africa*. Water Research Commission of South Africa. Report no. 301/07. Pretoria: CSIR.

Cowan, G.I. 1995. Wetland regions of South Africa. In: D.I. Cowan, ed., *Wetlands of South Africa*. Pretoria: Department of Environmental Affairs and Tourism.

Dahl, T.E. and Allord, G.J. 1997. Technical aspects of wetlands: History of wetlands in the conterminous United States. *US Geological Survey Water-Supply*, 2425: 19–26.

Davis, D. 2001. *Functions and values of wetlands*. Washington, DC (USA): United States Environmental Protection Agency.

DEAT. 1997. *South African National Wetland Inventory*. Wetlands Conservation Programme. Available from: www.ngo.grida.no/soesna/nsoer/resource/wetland/inventory.htm [Accessed 26 November 2012).

DEAT. 2006. *South Africa's National Biodiversity Strategy and Action Plan*. Pretoria: Department of Environmental Affairs and Tourism..

DEAT. 2007. *South Africa Environment Outlook: A report on the state of the environment.* Pretoria: Department of Environmental Affairs and Tourism.

Dely, J., Kotze, D., Quinn, N. and Marder, J. 1995. *A pilot project to compile an inventory and classification of wetlands in the Natal Drakensberg Park.* Pietermaritzburg: Institute of Natural Resources, Report 101.

De Voogt, K., Kite, G., Droogers, P. and Murray-Rust, H. 2000. Modelling water allocation between wetlands and irrigated agriculture: Case study of the Gediz basin, Turkey. Colombo, Sri Lanka: International Water Management Institute, IWMI working paper 001.

Dini, J. 2012 (October). Director, Freshwater Programme (SANBI). Personal Communication. Pretoria.

Dini, J., Cowan, G. and Goodman, P. 1998. *South African National Wetland Inventory. Proposed wetland classification system for South Africa*. Pretoria: Department of Environmental Affairs and Tourism.

DLA-CDSM. 2006. *1:50 000 Inland water bodies and rivers*. Pretoria: Department of Land Affairs – Chief Directorate: Surveys and Mapping.

Dlamini, W.M. 2010. A Bayesian belief network analysis of factors influencing wildfire occurrence in Swaziland. *Environ Model Software,* 25:199–208.

Driver, A., Maze, K., Rouget, M., Lombard, A.T., Nel, J.N., Turpie, J.K., Cowling, R.M., Desmet, P., Goodman, P., Harris, J., Jonas, Z., Reyers, B., Sink, K.J. and Strauss, T. 2005. *National Spatial Biodiversity Assessment 2004: Priorities for biodiversity conservation in SA.* Pretoria: SANBI.

Driver, A., Sink, K.J., Nel, J.N., Holness, S., Van Niekerk, L., Daniels, F., Jonas, Z., Majiedt, P.A., Harris, L. and Maze, K. 2012. *National Biodiversity Assessment 2011: An assessment of South Africa's biodiversity and ecosystems. Synthesis report*. Pretoria: SANBI and the Department of Environmental Affairs.

Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z., Knowler, D.J., Leveque, C., Naiman, R.J., Prieur-Richard, A., Soto, D., Stiassny, M.L.J. and Sullivan, C.A. 2006. Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews*, 81(2): 163–182.

DWAF. 2007. *Manual for the assessment of a Wetland Index of Habitat Integrity for South African floodplain and channelled valley bottom wetland types*. Compiled by M. Rountree (ed), C.P. Todd, C.J. Kleynhans, A.L. Batchelor, M.D. Louw, D. Kotze, D. Walters, S. Schroeder, P. Illgner, M. Uys and G.C. Marneweck. Report no. N/0000/00/WEI/0407. Pretoria: Resource Quality Services, Department of Water Affairs and Forestry (DWAF).

Eeley, H.A.C., Lawes, M.J. and Piper, S.E. 1999. The influence of climate change on the distribution of indigenous forest in KwaZulu-Natal, South Africa. *Journal of Biogeography*, 26: 595–617.

Ellery, W.N., Grenfell, M.C., Grenfell, S.E., Kotze, D., McCarthy, T., Tooth, S., Grundling, P.-L., Beckedahl, H., Le Maitre, D. and Ramsay, L. 2008. *WET-Origins: Controls on the distribution and dynamics of wetlands in South Africa*. WRC Report No. TT334/09. Pretoria: Water Research Commission.

Environmental Systems Research Institute (ESRI). 2007. *ArcGIS Desktop: Release 9.3*. Redlands, CA (USA): ESRI.

Escott, B. 2011. *Landform map for KZN based on the 90m SRTM DEM (v4 edited)*. Pietermaritzburg: Ezemvelo KZN Wildlife.

Escott, B. 2012 (March). GIS Analyst, Ezemvelo KZN Wildlife. Personal Communication. Pietermaritzburg (RSA).

Ewart-Smith, J.L., Ollis, D.J., Day, J.A. and Malan, H.L. 2006. *National wetland inventory: development of a wetland classification system for South Africa.* WRC Report No. K8/65. Pretoria: Water Research Commission.

Fawcett, T. 2006. An introduction to ROC analysis. *Pattern Recognition Letters*, 27: 861–874.

Fry, J.C. 1993. *Biological data analysis*. New York: Oxford University Press.

Gill, J. 2002. *Bayesian methods: A social and behavioural science approach*. Florida: Chapman and Hall/CRC.

GISCOE. 2001. *20m GISCOE DTM Data*. Pretoria: GISCOE (Pty) Ltd..

Grenfell, M.C., Ellery, W.N., Garden, S.E., Dini, J. and van der Valk, A.G. 2007. The language of intervention: A review of concepts and terminology in wetland ecosystem repair. *Water SA*, 33: 43–50.

Grenfell, M.C., Ellery, W.N., and Grenfell, S.E. 2009: Valley morphology and sediment cascades within a wetland system in the KwaZulu-Natal Drakensberg Foothills, Eastern South Africa, *CATENA*, 78 (1): 20–35.

Grêt-Regamey, A. and Straub, D. 2006. Spatially explicit avalanche risk assessment linking Bayesian networks to GIS. *Natural Hazards and Earth System Sciences,* 6: 911‒926.

Guisan, A. and Zimmermann, N.E. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2): 147‒186.

He, H.S., D.C. Dey, X. Fan, M.B. Hooten, J.M. Kabrick, C.K. Wikle, and Z. Fan (2007). Mapping pre-European settlement vegetation using a hierarchical Bayesian model and GIS. Plant Ecology, 191: 85-94.

Heimlich, R.E., Wiebe, K.D., Claasen, R., Gadsby, D. and House, R.M. 1998. *Wetlands and agriculture: private interests and public benefits*. Washington, DC: US Department of Agriculture.

Hooten, M.B., Larsen, D.R. and Wikle, C.K. 2003. Predicting the spatial distribution of ground flora on large domains using a hierarchical Bayesian model. *Landscape Ecology,* 18: 487‒502.

Hughes, R.H. and Hughes, J.S. 1992. *A directory of African wetlands.* Gland, Switzerland; Nairobi, Kenya; and Cambridge, UK: IUCN (the World Conservation Union), UNEP (the United Nations Environment Programme) and WCMC (the World Conservation Monitoring Centre).

Ibrahim, K. 2009. Assessment of wetlands in Kuala Terengganu District using Landsat TM. *Journal of Geography and Geology,* 1(2): 33‒40.

Islam, M.A., Thenkabail, P.S., Kulawardana, R.W., Alankara, R., Gunasinghe, S., Edussriya, C. and Gunwardana, A. 2008. Semi-automated methods for mapping wetlands using Landsat ETM+ and SRTM data. *International Journal of Remote Sensing*, 29(24): 7077‒7106.

Jenks, G.F. and Caspall, F.C. 1971. Error on chloroplethic maps: Definition, measurement, reduction. *Annals of American Geographers,* 61: 217‒244.

Jewitt, D. 2011. Landcover accuracy assessment photography – 2011. Unpublished Imagery Dataset, Biodiversity Research & Assessment Division. Pietermaritzburg: Ezemvelo KZN Wildlife.

Jewitt, G.P.W and Kotze, D.C. 2000: Wetland Conservation and Rehabilitation as Components of Integrated Catchment Management in the Mgeni Catchment, KwaZulu-Natal, South Africa. In: Bergkamp, G., Pirot, J.Y. and Hostettler, S (eds.) *Intergrated Wetlands and Water Resources Management, Proc. Workshop held at the 2nd Int. Conf. on Wetlands and Development.* November 1998, Dakar, Senegal.

Johnson, V.M. 2003. *Geographic information: How to find it, how to use it*. Westport, CT: Greenwood Press..

Joubert, R. 2009. *The origin and dynamics of Wakkerstroom Vlei, Mpumalanga*. Unpublished thesis: University of KwaZulu-Natal.

Junhua, L. and Wenjun, C. 2005. A rule-based method for mapping Canada's wetlands using optical, radar and DEM data. *International Journal of Remote Sensing,* 26 (22): 5051‒5069.

Jury, M.R. 1998. Statistical analysis and prediction of KwaZulu-Natal climate. *Theoretical and Applied Climatology*, 60: 1‒10.

King, L. 1978. A geomorphology of central and southern Africa. *Monographiae Biologicae,* 31: 1‒17.

Kidd, M. 2011. *Environmental law*. Cape Town: Juta..

Knight, A.W., Tindall, D.R. and Wilson, B.A. 2009. A multitemporal multiple density slice method for wetland mapping across the state of Queensland, Australia. *International Journal of Remote Sensing*, 30(13): 3365–3392.

Kocabas, V. and Dragicevic, S. 2007. Enhancing a GIS cellular automata model of land use change: Bayesian networks, influence diagrams and causality. *Transactions in GIS*, 11(5):681–702.

Kotze, D.C. and Breen, C.M. 1994. Wetlands and people. What values do wetlands have for us and how are these values affected by our land-use activities*? WETLAND-USE booklet 1*. Share-Net. Howick: WESSA.

Kotze, D.C. and Breen, C.M. 1996. *Wetlands and people: what values do wetlands have for us and how are these values affected by our land-use activities?* Pietermaritzburg: University of KwaZulu-Natal.

Kragt, M.E. 2009. A beginners guide to Bayesian network modelling for integrated catchment management (Technical report no. 9). Hobart: Landscape Logic.

Kulawardhana, R.W., Thenkabail, P.S., Vithanage, J., Biradar, C., Islam, M.A., Gunasinghe, S. and Alankara, R. 2007. Evaluation of the wetland mapping methods using Landsat ETM+ and SRTM data. *Journal of Spatial Hydrology*, 7(2): 62–96.

KwaZulu-Natal Provincial Planning Commission (KZNPPC). 2011. *Provincial Growth and Development Strategy*. South Africa: KZNPPC.

Landmann, T., Schramm, M., Colditz, R.R., Dietz, A. and Dech, S. 2010. Wide area wetland mapping in semi-arid Africa using 250-meter MODIS metrics and topographic variables. *Remote Sensing*, 2: 1751–1766.

Lee, S., Choi, J., and Min, K. 2002. Landslide susceptibility analysis and verification using the Bayesian probability model. *Environmental Geology*, 43: 120–131.

LeRoux, S.D. not dated. *Bioclimatic Groups of Natal*. Available from: www.kzndae.gov.za [Accessed 27 August 2013].

Li, J. and Chen, W. 2005. A rule-based method for mapping Canada’s wetlands using optical, radar and DEM data. *International Journal of Remote Sensing*, 26(22): 5051–5069.

Liu, Y., Guo, H., Mao, G. and Yang, P. 2008. A Bayesian hierarchical model for urban air quality prediction under uncertainty. *Atmospheric Environment*, 42: 8464–8469.

Lunetta, R.S., Balogh, M.E. and Merchant, J.W. 1999. Application of multi-temporal Landsat 5 TM imagery for wetland identification. *Photogrammetric engineering and remote sensing*, 65: 1303–1310.

Lyon, J.G. 1993. Practical handbook for wetland identification and delineation. Florida (USA): Lewis Publishers.

Macfarlane, D. M., Walters, D. and Cowden, C. 2011. *A wetland health assessment of KZN’s priority wetlands: An unpublished report prepared for EKZNW.* Pietermaritzburg: Biodiversity Conservation Planning Division, Ezemvelo KZN Wildlife.Margules, C.R. and Pressey, R.L. 2000. Systematic conservation planning. *Nature*, 405: 243–253.

Marnewecke, G. and Kotze, D. 1999. *Resource-directed measures for protection of water resources: Wetland ecosystems – Appendix W6: Guidelines for delineation of wetland boundary and wetland zones.* Pretoria: Department of Water Affairs and Forestry.

May,D., Wang, J., Kovacs, J. and Mutter, M. 2002. Mapping wetland extent using IKONOS satellite imagery of the O'Donell Point region, Georgian Bay, Ontario. Proceedings of the 25th Canadian Symposium on Remote Sensing, Canadian Aeronautics and Space Institute.

Mbona, M. 2011. *National Wetland Inventory life and history up to 2011*. National Wetland Indaba 2011 Poster Presentation. Pretoria (RSA).

McCarthy, T.S., Hancox, P.J. 2000. Wetlands. In: T.C. Partridge, R.R. Maud,eds. *The Cenozoic of Southern Africa*. Oxford: Oxford University Press.

Metz, M. 2008. *Soil moisture estimation for KwaZulu-Natal.* Pretoria: ARC-Institute for Soil, Climate and Water, GW/A/2009/33.

Michaud, J.P. 2001. *At home with wetlands: A landowner's guide (2nd Edition)*. Olympia, WA: Washington State Department of Ecology.

Millenium Ecosystem Assessment (MEA). 2005. *Ecosystem and human well-being: Wetlands and water synthesis*. Washington, DC: World Resources Institute.

Mitsch, W.J. and Gosselink, J.G. 1993. *Wetlands (2nd Edition)*. New York: Van Nostrand Reinhold Publishers.

Mitsch, W.J. and Gosselink, J.G. 2000. *Wetlands (3rd Edition)*. New York: John Wiley & Sons.

Mitsch, W.J. and Gosselink, J.G. 2007. *Wetlands (4th Edition).* New York: John Wiley & Sons, Inc.

Moore, P.D. 2006. *Biomes of the Earth: Wetlands*. New York: Chelsea House/Infobase Publishing.

National Research Council (NRC). 1995. *Wetlands: Characteristics and Boundaries*. Washington, D.C.: National Academy Press.

Nel, J.L., Driver, A., Strydom, W.F., Maherry, A., Petersen, C., Hill, L., Roux, D.J., Nienaber, S., van Deventer, H., Swartz, E. and Smith-Adao, L.B. 2011a. *Atlas of Freshwater Ecosystem Priority Areas in South Africa: Maps to support sustainable development of water resources*. Gezina: Water Research Commission.

Nel, J.L., Murray, K.M., Maherry, A.M., Petersen, C.P., Roux, D.J., Driver, A., Hill, L., van Deventer, H., Funke, N., Swartz, E.R., Smith-Adao, L.B., Mbona, N., Downsborough, L. and Nienaber, S. 2011b. *Technical report for the National Freshwater Ecosystem Priority Areas project*. Gezina: Water Research Commission.

Nel, M. not dated. *Spreading the wetlands work*. Available from: www.wetland.org.za [Accessed 24 August 2012].

Nel, W. 2008. Observations on daily rainfall events in KwaZulu-Natal Drakensberg. *Water South Africa*, 34 (2): 271–274.Nel, W. and Sumner, P.D. 2007. Intensity, energy and erosivity attributes of rainstorms in the KwaZulu-Natal Drakensberg, South Africa. *South African Journal of Science*, 103: 398–402.

O'Hara, C. G. 2002: Remote sensing and geospatial application for wetland mapping, assessment, and mitigation. In Morain, S. and Budge, A. (eds) *Integrated Remote Sensing at the Global, Regional and Local Scale*. Denver, CO: International Society for Photogrammetry and Remote Sensing.

Olhan, E., Gün, S., Ataseven, Y. and Arisoy, H. 2010. Effects of agricultural activities in Seyfe Wetland. *Scientific Research and Essay,* 5(1): 9–14.Oxford English Dictionary. 2002. *The Oxford English Reference Dictionary*. Oxford: Oxford University Press.

Ozesmi, S.L. and Bauer, M.E. 2002. Satellite remote sensing of wetlands. *Wetland Ecology and Management,* 10: 381–402.

Pantaleoni, E., Wynne, R.H., Galbraith, J.M. and Campbell, J.B. 2009. A logit model for predicting wetland location using ASTER and GIS. *International Journal of Remote Sensing*, 30(9): 2215–2236.

Patrick, M.J. and Ellery, W.N. 2006. Plant community and landscape patterns of a floodplain wetland in Maputaland, Northern KwaZulu-Natal, South Africa. *African Journal of Ecology*, 45: 175–183.

Pillay, D.L. 2001. An investigation into mapping wetlands using satellite imagery: the case of Midmar sub-catchment. Unpublished MSc dissertation. University of KwaZulu-Natal, Pietermaritzburg.

Ramsar Convention Secretariat. 1993. *The Ramsar Convention on Wetlands: Its history and development.* Switzerland: Ramsar Convention Secretariat.

Ramsar Convention Secretariat. 2006. *The Ramsar Convention Manual: a guide to the Convention on Wetlands (Ramsar, Iran, 1971)* (4th Edition). Switzerland: Ramsar Convention Secretariat.

Ramsar Convention Secretariat. 2007. *Wetland inventory: A Ramsar framework for wetland inventory. Ramsar handbooks for the wise use of wetlands, Vol. 12 (3rd Edition).* Switzerland: Ramsar Convention Secretariat.

Ramsar Convention Secretariat. 2010. *Ramsar's fact sheets on wetland ecosystem services.* Switzerland: Ramsar Convention Secretariat.

Ramsar Convention Secretariat. 2011. *The Ramsar Convention Manual: a guide to the Convention on Wetlands (Ramsar, Iran, 1971), (5th Edition).* Switzerland: Ramsar Convention Secretariat.

Rebelo, L.M., Finlayson, C.M. and Nagabhatla, N. 2009. Remote sensing and GIS for wetland inventory, mapping and change analysis. *Journal of Environmental Management*, 90(7): 2144–2153.

Ricchetti, E. 2000. Multispectral satellite image and ancillary data integration for geological classification. *Photogrammetric Engineering and Remote Sensing*, 66(4):429–435.

Rivers-Moore, N.A. 2013 (January). Wetland and Freshwater Specialist, Institute of Natural Resources. Personal Communication. Pietermaritzburg (RSA).

Rivers-Moore, N.A. and Cowden, C. 2012. Regional prediction of wetland degradation in South Africa. *Wetlands Ecology and Management*, 20(5): 1–14.

Rivers-Moore, N.A. and Goodman, P.S. 2010. River and wetland classification for freshwater conservation planning in KwaZulu-Natal. *African Journal of Aquatic Science*, 35(1): 61–72.

Rivers-Moore, N.A., Goodman, P.S. and Nel, J.L. 2010. Scale-based freshwater conservation planning: towards protecting freshwater biodiversity in KwaZulu-Natal, South Africa. *Freshwater Biology*, 56: 125–141.

Rogers, K.H. 1997. Freshwater wetlands. In: R.M. Cowling, D.M. Richardson and S.M. Pierce, eds. *Vegetation of southern Africa*. Cambridge: Cambridge University Press.

Rountree, M., Thompson, M., Kotze, D., Batchelor, A. and Marneweck, G. 2009. *WET-Prioritise: Guidelines for prioritising wetlands at national, regional and local scales*. Gezina: Water Research Commission.

Republic of South Africa (RSA). 1983. *Conservation of Agricultural Resources Act (Act 43 of 1983)*. Pretoria: Government Printer.

Republic of South Africa (RSA). 1996. *Constitution of SA (Act 108 of 1996)*. Pretoria: Government Printer.

Republic of South Africa (RSA). 1998a. *National Water Act (Act 36 of 1998).* Government Printer.

Republic of South Africa (RSA). 1998b. *National Environmental Management Act (Act 107 of 1998)*. Government Printer.

Republic of South Africa (RSA). 2004. *National Environmental Management: Biodiversity Act*. *(Act 10 of 2004).* Government Printer.

Ryo, M., Peng, G. and Bing, X. 2012. Spectral mixture analysis for bi-sensor wetland mapping using Landsat TM and Terra MODIS data. *International Journal of Remote Sensing*, 30(11): 3373–3401.

SANBI. 2009. Further development of a proposed National Wetland Classification System for South Africa. Primary project report prepared by the Freshwater Consulting Group (FCG) for the South African National Biodiversity Institute (SANBI).

Schulze, R.E. 1982. *Agrohydrology and climatology of Natal*. Pretoria: Water Research Commission..

Schulze, R.E. 1997. *South African atlas of agrohydrology and climatology (2003 rainfall grids).* Pretoria: Water Research Commission.

Scott-Shaw, C.R. and Escott, B.J. 2011. *KwaZulu-Natal Provincial Pre-Modification Vegetation Type Map – 2011.* Unpublished GIS coverage [kznveg05v2_1_11_wll.zip]. Pietermaritzburg: Biodiversity Conservation Planning Division, Ezemvelo KZN Wildlife.

Secretariat of the Convention on Biological Diversity (SCBD). 2006. *Global Biodiversity Outlook 2*. Montreal (Canada): Convention of Biological Diversity.

Taylor, A.R.D., Howard, G.W. and Begg, G.W. 1995. Developing wetland inventories in southern Africa: A review. *Plant Ecology*, 118(1,2): 57–79.

Thompson, M.W. 1994. *The use of current remote sensing technology to support wetland inventory and monitoring: Feasibility and application in South Africa – Part 1. Guidelines for the use of remote sensing in the mapping and monitoring of wetlands in South Africa.* Pretoria: Council for Scientific and Industrial Research (CSIR).

Thompson, M.W., Marneweck, G., Bell, S., Kotze, D., Muller, J., Cox, D. and Clark, R. 2002. *A methodology proposed for a South African National Wetland Inventory. Report prepared for South African Wetlands Conservation Programme.* Pretoria: Department of Environmental Affairs and Tourism.

United Nations, 2010: *The Millennium Development Goals: Report 2010*. United Nations. New York.

United States Department of Agriculture (USDA). 2008. *Hydrogemorphic wetland classification system: An overview and modification to better meet the needs of the natural resources conservation service. Technical Note 190-9-76.* Washington, DC: Natural Resources Conservation Service.

University of KwaZulu-Natal (UKZN). Not dated. The geology of KwaZulu-Natal. Geology Education Museum. Available from: stec.ukzn.ac.za/GeologyEducationMuseum [Accessed 27 August 2013].

University of KwaZulu-Natal (UKZN). 2011. *Cartography department GIS dataset*. School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Pietermaritzburg.

Uusitalo, L. 2007. Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling*, 203: 312–318.

Van der Berg, H.M., Weepener, H.L. and Metz, M. 2009. *Spatial modelling for semi-detailed soil mapping in KwaZulu-Natal*. Pretoria: Department of Agriculture, 21.1.1/07LUSM-03/ISCW.

Van der Valk, A. 2006. *The biology of freshwater wetlands*. Oxford: Oxford University Press.

Votteler, T.H. and Muir, T.A. 1996. Wetland protection legislation. In: J.D. Fretwell, J.S. Williams and P.J. Redman, eds. *National Water Summary on Wetland Resources. USGS Water Supply Paper 2425*. Washington, DC: US Geological Survey.

Walters, D. 2012 (August). Senior Wetland Ecologist/Specialist, Mondi Wetland Programme. Personal Communication. Howick (RSA).

Wilen, B.O., Carter, V. and Jones, J.R. 2002. Wetland management and research: Wetland mapping and inventory. *National Water Summary on Wetland Resources*, 2425: 1–11.

Wilen, B.O. and Tiner, R.W. 1993. Wetlands of the United States. In D.F. Whignam, D. Dgkyjova and S. Hejny, eds. *Wetlands of the world I – Inventory, ecology and management.* Dordrecht (Netherlands): Khuver Academic Publishers.

Winstanley, T.J. 2001. *Wetland rehabilitation manual: Legal component*. Cape Town: EnAct International.

Woodhouse, S., Lovett, A., Dolman, P. and Fuller, R. 2000. Using GIS to select priority areas for conservation. *Computers, Environment and Urban Systems*, 24: 79–93.

Wright, C. and Gallant, A. 2007. Improved wetland remote sensing in Yellowstone National Park using classification trees to combine TM imagery and ancillary environmental data. *Remote Sensing of Environment*, 107(4): 582–605.

Xie, Y., Sha, Z. and Yu, M. 2008. Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology*, 1(1): 9–23.

Yan, W. and Tinker, N.A. 2004. An integrated biplot analysis system for displaying, interpreting, and exploring genotype x environment interaction. *Crop Science*, 45(3): 1004–1016.

Zhou X.H., Obuchowski N.A., McClish D.K. 2002. *Statistical methods in diagnostic medicine*. New York: Wiley.

6.1 Appendix A – Background of input variables

6.1.1 Legend of qualitative input variables

Table 6.1: Descriptors of the classes of all the qualitative input variables

SOIL ASSOCIATION

1	Bare rock and rock outcrops dominant
2	Litho soils dominant
3	Litho soils and red apedal soils co-dominant
4	Litho soils and yellow-brown apedal soils co-dominant
5	Litho soils and structured soils co-dominant
6	Red apedal soils dominant (including soils with plinthic and neocutanic horizons)
7	Yellow-brown apedal soils dominant (including soils with plinthic and neocutanic horizons)
8	Red apedal and structured soils complex
9	Yellow-brown apedal and structured soils complex
10	Duplex soils (non plinthic) and or pedocutanic soils dominant
11	Structured soils dominant (vertic and melanic soils dominant or co-dominant with other structured soils)
12	Significant occurrence of hydromorphic soils and prismacutanic soils
13	Significant occurrence of hydromorphic soils
14	Water

SOIL DEPTH

1	Shallow soils dominant
2	Medium-deep soils dominant
3	Deep soils dominant
4	Water

SOIL MOISTURE

1	Very dry
2	Dry
3	Medium
4	Moist
5	Very moist

TERRAIN UNITS

1	Crest
2	Midslope convex
3	Midslope concave
4	Foot slope
5	Valley bottom

SLOPE (4 CLASSES)

1	Low
2	Low to medium
3	Medium to high
4	High

CLAY SOIL

1	Very low clay content dominant
2	Low clay content dominant
3	Medium clay content dominant
4	High clay content dominant
5	Very high clay content dominant
6	Water

HYDROMORPHIC SOILS

0	No
1	Yes

LANDFORM CLASSES

1	Canyons
2	Shallow valleys
3	Upland drainage
4	U-shaped valleys
5	Plains
6	Open slopes
7	Upper slopes
8	Local ridges
9	Midslope ridges
10	Mountain tops

GEOLOGY

1	AMPHIBOLITE
2	ARENITE
3	BASALT
4	CONGLOMERATE
5	DOLERITE
6	GABBRO
7	GNEISS
8	GRANITE
9	GREENSTONE
10	MARBLE
11	MUDSTONE
12	OLIVINE GABBRO
13	PERIDOTITE
14	QUARTZITE
15	RHYOLITE
16	SAND
17	SCHIST
18	SERPENTINITE
19	SHALE
20	SILTSTONE
21	TILLITE
22	TONALITE
23	WATERBODY

6.1.2 Descriptions of wetland HGM types assessed in the study

The following descriptions are taken from the *Primary Project Report* of the *Proposed National Wetland Classification System* (SANBI, 2009: 33-35):

1. **Channelled valley-bottom wetland:** a mostly flat valley-bottom wetland dissected by and typically elevated above a channel (see channel). Dominant water inputs to these areas are typically from the channel, either as surface flow resulting from overtopping of the channel bank/s or as interflow, or from adjacent valley-side slopes (as overland flow or interflow). Water generally moves through the wetland as diffuse surface flow, although occasional, short-lived concentrated flows are possible during flooding events. Small depressional areas within a channelled valley-bottom wetland can result in the temporary containment and storage of water within the wetland. Water generally exits in the form of diffuse surface flow and interflow, with the infiltration and evaporation of water from these wetlands also being potentially significant (particularly from depressional areas). The hydrodynamic nature of channelled valley-bottom wetlands is characterised by bidirectional horizontal flow, with limited vertical fluctuations in depressional areas.
2. **Unchannelled valley-bottom wetland:** a mostly flat valley-bottom wetland area without a major channel running through it, characterised by an absence of distinct channel banks and the prevalence of diffuse flows, even during and after high rainfall events. Water inputs are typically from an upstream channel, as the flow becomes dispersed, and from adjacent slopes (if present) or groundwater. Water generally moves through the wetland in the form of diffuse surface flow and/or interflow (with some temporary containment of water in depressional areas), but the outflow can be in the form of diffuse or concentrated surface flow. Infiltration and evaporation from unchannelled valley-bottom wetlands can be significant, particularly if there are a number of small depressions within the wetland area. Horizontal, unidirectional diffuse surface-flow tends to dominate in terms of the hydrodynamics.
3. **Floodplain wetland:** the mostly flat or gently sloping wetland area adjacent to and formed by a Lowland or Upland Floodplain river, and subject to periodic inundation by overtopping of the channel bank. For purposes of the classification system, the location adjacent to a river in the Lowland or Upland Floodplain Zone is the key criterion for distinguishing a floodplain wetland from a channelled valley-bottom wetland. Water and sediment input to floodplain wetland areas is mainly via overtopping of a major channel, although there could be some overland or subsurface flow from adjacent valley side-slopes (if present). Water movement through the wetland is dominantly horizontal and bidirectional, in the form of diffuse surface flow and interflow, although there can be significant temporary containment of water in depressional areas (within which water movement is dominantly vertical and bidirectional). Water generally exits as diffuse surface flow and/or interflow, but infiltration and evaporation of water from a floodplain wetland can also be significant, particularly if there are a number of depressional areas within the wetland.

4. **Depression:** a landform with closed elevation contours that increases in depth from the perimeter to a central area of greatest depth, and within which water typically accumulates. Dominant water sources are precipitation, ground water discharge, interflow and (diffuse or concentrated) overland flow. For 'depressions with channeled inflow', concentrated overland flow is typically a major source of water for the wetland, whereas this is not the case for 'depressions without channeled inflow'. Dominant hydrodynamics are (primarily seasonal) vertical fluctuations. Depressions may be flat bottomed (in which case they are often referred to as 'pans') or round-bottomed (in which case they are often referred to as 'basins'), and may have any combination of inlets and outlets or lack them completely. For 'exorheic depressions', water exits as concentrated surface flow while, for 'endorheic depressions', water exits by means of evaporation and infiltration.
5. **Flat:** a near-level wetland area (i.e. with little or no relief) with little or no gradient, situated on a plain or a bench in terms of landscape setting. The primary source of water is precipitation, with the exception of flats along the coast (usually in a plain setting) where the water table (i.e. groundwater) may rise to the surface or near to the surface in areas of little or no relief because of the location near to the base level of the land surface represented by the presence of the ocean.
6. **Seep:** a combination of hillslope seeps and valleyhead seeps. Hillslope seep: a wetland area located on (gently to steeply) sloping land, which is dominated by the colluvial (i.e. gravity-driven), unidirectional movement of material down-slope. Water inputs are primarily from groundwater or precipitation that enters the wetland from an up-slope direction in the form of subsurface flow. Water movement through the wetland is mainly in the form of interflow, with diffuse overland flow ('sheetwash') often being significant during and after rainfall events. Water leaves a 'hillslope seep with channelled outflow' mostly by means of concentrated surface flow, whereas water leaves a 'hillslope seep without channelled outflow' by means of a combination of diffuse surface flow, interflow, evaporation and infiltration. Valleyhead seep: a gently-sloping, typically concave wetland area located on a valley floor at the head of a drainage line, with water inputs mainly from subsurface flow (although there is usually also a convergence of diffuse overland water flow in these areas during and after rainfall events). Horizontal, unidirectional (down-slope) movement of water in the form of interflow and diffuse surface flow dominates within a valleyhead seep, while water exits at the downstream end as concentrated surface flow where the valleyhead seep becomes a channel.

6.2 Appendix B – Pre-model results

Table 6.2.1: The descriptive statistics of the dataset extracted in known wetland areas

	AVERAGE	MIN	MAX	STD DEVIATION	MEDIAN	CLASS
Soil association	8	1	14	3.54	8	Red apedal and structured soils complex
Soil depth	2	1	5	0.74	2	Medium-deep soils dominant
Soil moisture	3	1	5	0.57	3	Medium
Terrain units	4	1	5	1.26	3	Midslope concave
Slope (4 classes)	1	1	4	0.63	1	Low
Clay soils	4	1	15	1.06	4	High clay content dominant
Hydromorphic soils	0	0	1	0.37	0	Hydromorphic soil
Solar radiation	22.16	18.44	25.83	1.34	22.27	n/a
Mean annual temperature	17.43	9.00	22.00	2.81	17.00	n/a
Mean annual precipitation	878.82	505.00	1759.00	162.99	864.00	n/a
Heat units (Winter)	1006.52	-4.00	1831.00	537.58	843.00	n/a
Heat units (Summer)	1903.11	298.00	2803.00	482.30	1803.00	n/a
Evapotranspiration	105.56	78.00	131.00	10.54	103.00	n/a
Evaporation	1748.08	1280.64	2078.10	143.29	1717.80	n/a
DEM	857.39	1.28	2476.65	608.23	1015.60	n/a
Aspect (degrees)	176.71	0.00	360.00	102.89	170.10	n/a
Flow accumulation	320.01	0.00	82528.00	2190.71	5.00	n/a
Flow direction	29.17	1.00	192.00	39.58	8.00	n/a
Slope (degrees)	3.69	0.00	63.43	4.39	2.26	n/a
Sinks	133.69	0.00	102897.00	3265.85	0.00	n/a
Landform classes	5	1	10	1.52	5	Plains
Groundwater	20.80	0.00	47.90	7.25	21.06	n/a
Geology	11	1	24	6.94	11	Mudstone

Table 6.2.2: The dataset defined into states of high, medium and low (only 20 of 44 771 dataset records shown below for illustration purposes)

Soil_depth	Terrain_units	Solar_radiation	Mean_precip	Evapotranspiration	Altitude	Slope_degree	Landform	Hydromorphic_soil	Prob_of_wetland_occurence
Medium	Medium	High	Medium	Low	High	Medium	Medium	NO	HIGH
Low	Medium	Medium	Medium	Low	Medium	Medium	Medium	NO	HIGH
Low	Medium	Medium	Medium	Low	Medium	Medium	Medium	NO	HIGH
Low	Medium	Medium	Medium	Low	Medium	Medium	Medium	NO	HIGH
Medium	Low	Medium	Medium	Low	Medium	Low	Medium	NO	HIGH
Medium	Medium	Medium	Medium	Low	Medium	Low	Medium	NO	HIGH
Medium	Medium	Medium	Medium	Low	Medium	Low	Low	NO	HIGH
Medium	Low	Low	High	High	Low	Low	Low	NO	HIGH
Medium	Low	Low	High	High	Low	Low	Low	NO	HIGH
Medium	Low	Low	Medium	High	Low	Low	Low	NO	HIGH
Medium	Medium	High	Medium	Medium	Medium	Low	Low	NO	LOW
Low	High	High	High	Medium	Medium	Medium	Medium	NO	LOW
Low	Medium	Low	Medium	Low	Medium	Low	Medium	NO	LOW
Low	Medium	High	High	Low	High	High	Medium	YES	LOW
Low	Medium	Low	Medium	Medium	Low	Low	Medium	NO	LOW
Low	Medium	Low	Medium	Low	Low	Medium	Medium	NO	LOW
Low	Medium	High	Medium	Medium	Medium	Medium	Medium	NO	LOW
Low	Medium	Medium	Medium	High	Low	Medium	Medium	NO	LOW
Low	Medium	High	Low	Low	Medium	High	Medium	NO	LOW
Medium	Medium	Low	Medium	Low	Low	Medium	Medium	NO	LOW

Table 6.2.3: The Conditional Probability Table (CPT) derived from NETICA using the Bayesian network (only 20 of 13 122 CPT records shown below for illustration purposes)

									Wetland Probability	
Altitude	**Slope_degree**	**Evapotranspiration**	**Soil_depth**	**Solar_radiation**	**Mean_precip**	**Landform**	**Hydromorphic_soil**	**Terrain_units**	**HIGH**	**LOW**
High	Low	Low	Medium	High	High	Low	NO	Medium	84.7458	15.2542
High	Low	Low	Medium	High	High	Low	NO	Low	88.4615	11.5385
High	Low	Low	Medium	High	Medium	High	YES	High	50	50
High	Low	Low	Medium	High	Medium	High	YES	Medium	50	50
High	Low	Low	Medium	High	Medium	High	YES	Low	90	10
High	Low	Low	Medium	High	Medium	High	NO	High	29.4118	70.5882
High	Low	Low	Medium	High	Medium	High	NO	Medium	79.2453	20.7547
High	Low	Low	Medium	High	Medium	High	NO	Low	50	50
High	Low	Low	Medium	High	Medium	Medium	YES	High	50	50
High	Low	Low	Medium	High	Medium	Medium	YES	Medium	84.6154	15.3846
High	Low	Low	Medium	High	Medium	Medium	YES	Low	94.4444	5.55556
High	Low	Low	Medium	High	Medium	Medium	NO	High	57.1429	42.8571
High	Low	Low	Medium	High	Medium	Medium	NO	Medium	78.7879	21.2121
High	Low	Low	Medium	High	Medium	Medium	NO	Low	92.1429	7.85714
High	Low	Low	Medium	High	Medium	Low	YES	High	66.6667	33.3333
High	Low	Low	Medium	High	Medium	Low	YES	Medium	78.9474	21.0526
High	Low	Low	Medium	High	Medium	Low	YES	Low	96.7033	3.2967
High	Low	Low	Medium	High	Medium	Low	NO	High	56.8965	43.1034
High	Low	Low	Medium	High	Medium	Low	NO	Medium	89.5669	10.4331
High	Low	Low	Medium	High	Medium	Low	NO	Low	98.2332	1.76678

Table 6.2.4: The CPT modified into logarithmic codes to derive unique values that represent conditional probabilities (only 20 of 13 122 logarithmic modified records shown below for illustration purposes)

										Wetland Probability	
Altitude	Slope_degree	Evapotranspiration	Soil_depth	Solar_radiation	Mean_precip	Landform	Hydromorphic_soil	Terrain_units	LOG SUM (UNIQUE VALUE)	HIGH	LOW
3	10	100	2,000	30,000	300,000	1,000,000	10,000,000	200,000,000	211332113	84.7458	15.2542
3	10	100	2,000	30,000	300,000	1,000,000	10,000,000	100,000,000	111332113	88.4615	11.5385
3	10	100	2,000	30,000	200,000	3,000,000	30,000,000	300,000,000	333232113	50	50
3	10	100	2,000	30,000	200,000	3,000,000	30,000,000	200,000,000	233232113	50	50
3	10	100	2,000	30,000	200,000	3,000,000	30,000,000	100,000,000	133232113	90	10
3	10	100	2,000	30,000	200,000	3,000,000	10,000,000	300,000,000	313232113	29.4118	70.5882
3	10	100	2,000	30,000	200,000	3,000,000	10,000,000	200,000,000	213232113	79.2453	20.7547
3	10	100	2,000	30,000	200,000	3,000,000	10,000,000	100,000,000	113232113	50	50
3	10	100	2,000	30,000	200,000	2,000,000	30,000,000	300,000,000	332232113	50	50
3	10	100	2,000	30,000	200,000	2,000,000	30,000,000	200,000,000	232232113	84.6154	15.3846
3	10	100	2,000	30,000	200,000	2,000,000	30,000,000	100,000,000	132232113	94.4444	5.55556
3	10	100	2,000	30,000	200,000	2,000,000	10,000,000	300,000,000	312232113	57.1429	42.8571
3	10	100	2,000	30,000	200,000	2,000,000	10,000,000	200,000,000	212232113	78.7879	21.2121
3	10	100	2,000	30,000	200,000	2,000,000	10,000,000	100,000,000	112232113	92.1429	7.85714
3	10	100	2,000	30,000	200,000	1,000,000	30,000,000	300,000,000	331232113	66.6667	33.3333
3	10	100	2,000	30,000	200,000	1,000,000	30,000,000	200,000,000	231232113	78.9474	21.0526
3	10	100	2,000	30,000	200,000	1,000,000	30,000,000	100,000,000	131232113	96.7033	3.2967
3	10	100	2,000	30,000	200,000	1,000,000	10,000,000	300,000,000	311232113	56.8965	43.1034
3	10	100	2,000	30,000	200,000	1,000,000	10,000,000	200,000,000	211232113	89.5669	10.4331
3	10	100	2,000	30,000	200,000	1,000,000	10,000,000	100,000,000	111232113	98.2332	1.76678

Table 6.2.5: Reclass table used to reclassify spatial layer into conditional probabilities (only 20 of 13 122 unique value records shown below for illustration purposes)

UNIQUE VALUE	Wetland Probability HIGH
211332113	84.7458
111332113	88.4615
333232113	50
233232113	50
133232113	90
313232113	29.4118
213232113	79.2453
113232113	50
332232113	50
232232113	84.6154
132232113	94.4444
312232113	57.1429
212232113	78.7879
112232113	92.1429
331232113	66.6667
231232113	78.9474
131232113	96.7033
311232113	56.8965
211232113	89.5669
111232113	98.2332

6.3 Appendix C – Visual assessment results

6.3.1 Visual assessment case study sites

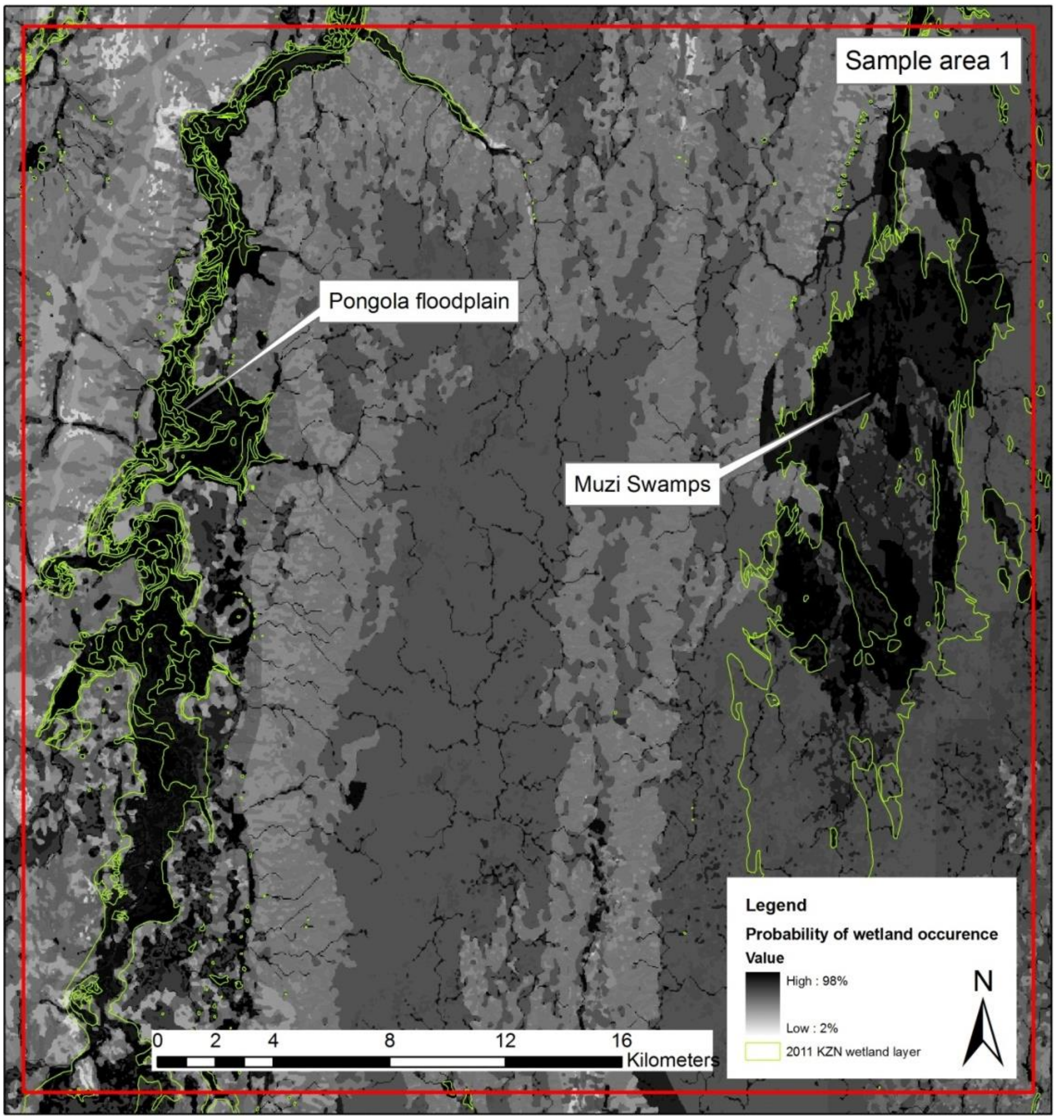


Figure 6.1: Case study Site 1 located in the Makatini flats and the Zululand coastal plains.

Significance in Assessment: This area contains two of the biggest wetland areas in KZN, namely the Muzi swamp and the Pongola floodplain.

In Site 1, the modelled probability layer had clearly identified the Pongola floodplain and Muzi Swamp. Comparing the probability layer to the current KZN wetland layer, the probability layer predicted more probable wetland areas along the Pongola floodplain, especially the southeastern parts of the system (Figure 6.1). The probability layer had under-represented the Muzi Swamps south of the system, and it is difficult to pinpoint why, but it could be the low relief of the system as a whole.

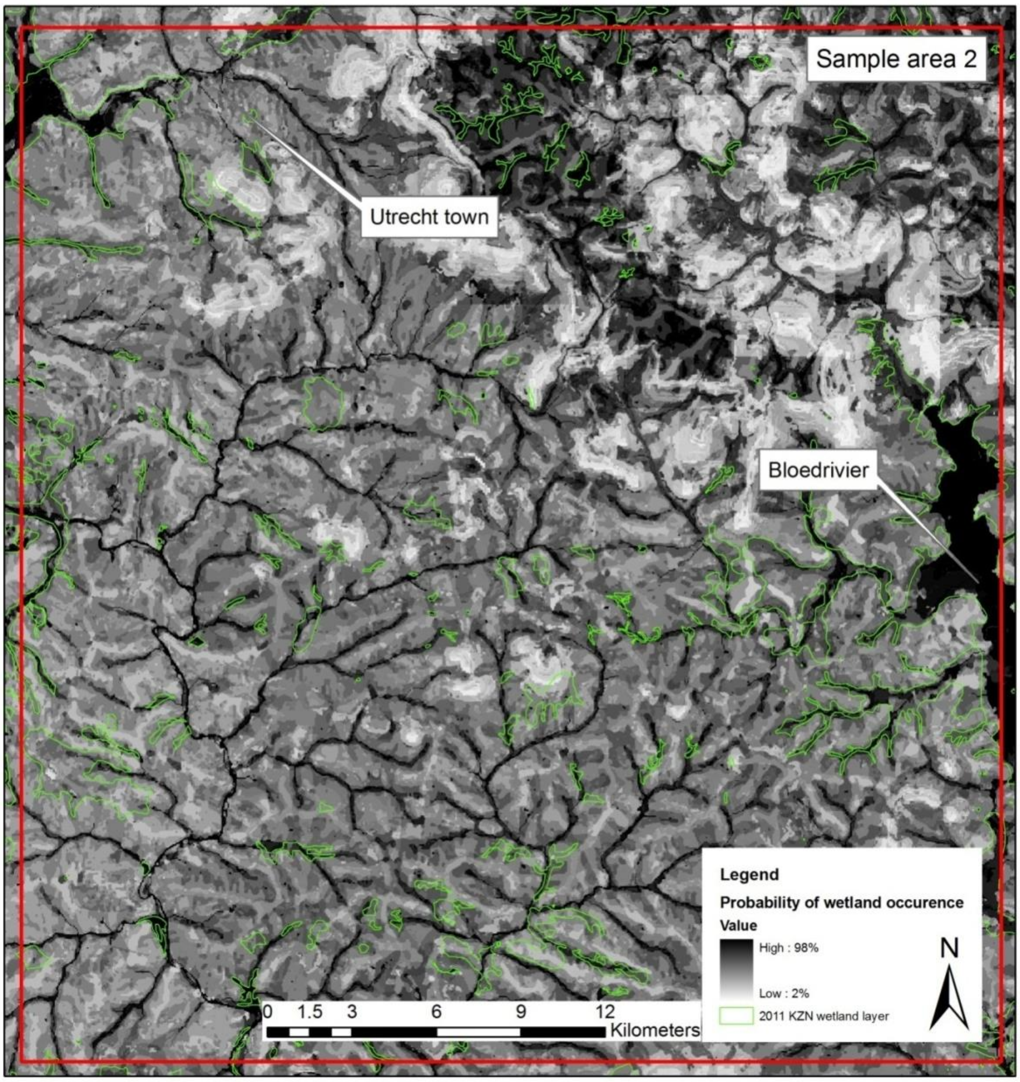


Figure 6.2: Case study Site 2 located in the upper reaches of the Buffalo River in the Tugela basin.

Significance in Assessment: Illustration of possible overestimation of wetland area by predicting riparian areas as wetlands.

In Site 2, the probability model had correctly predicted the Blood River wetland system but visual assessment revealed that the model had also identified drainage lines (i.e. riparian areas) as highly probable wetland areas (Figure 6.2). This may present a certain limitation regarding the modelled layer in which it identifies steeper flowing areas as probable wetlands, resulting in a potential overestimation of predicted wetland area.

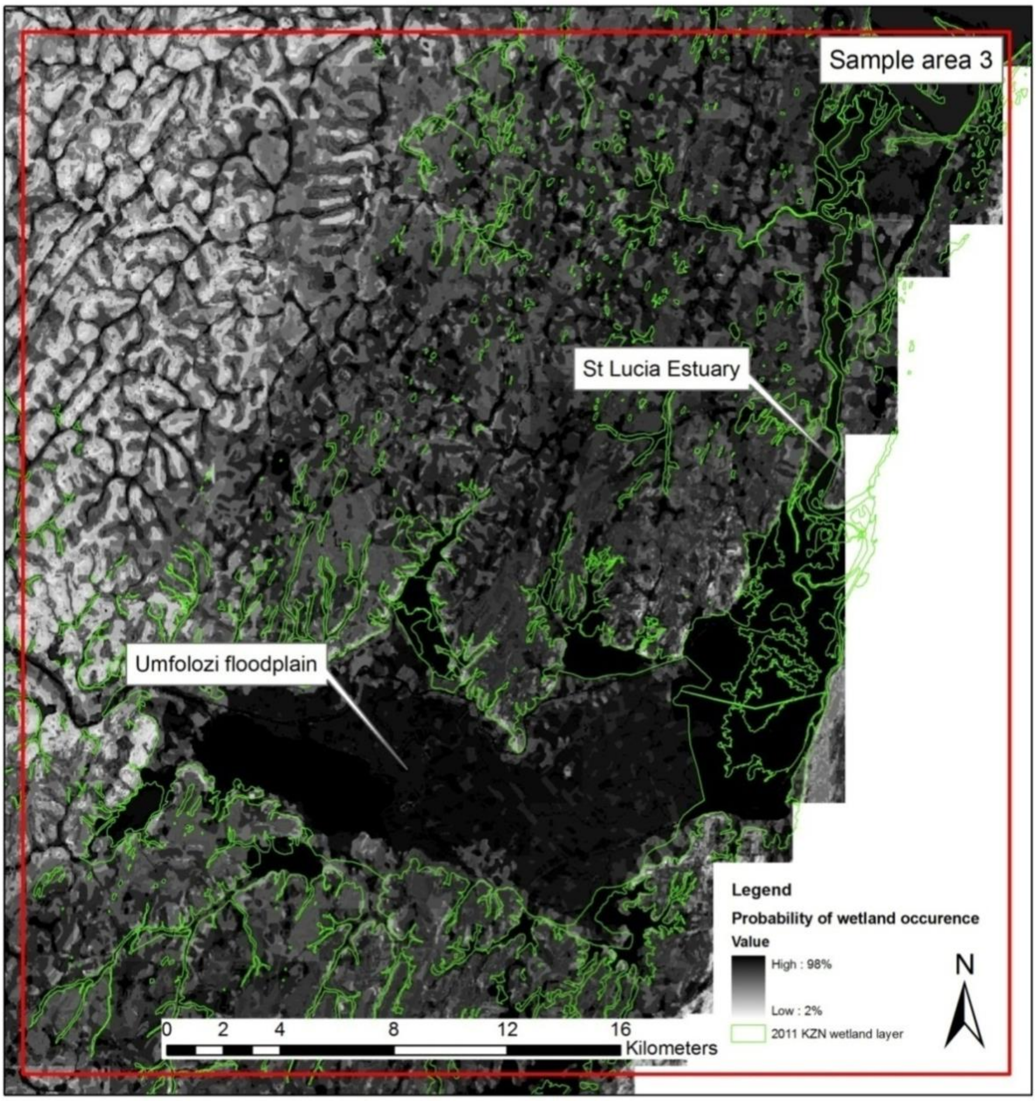


Figure 6.3: Case study Site 3 located in the Zululand coastal plains with the Mfolozi swamps.

Significance in Assessment: Assessment of the wetland prediction in a highly modified agricultural environment and complex coastal systems.

In Site 3, the probability layer had identified a mosaic of different wetland areas shown by the patchy shades of grey along the coast. Coastal plain wetland areas present different wetland types ranging from seasonal to permanent wetlands and open-water bodies, and the modelled layer had shown this matrix and complexity in wetland distribution in this study site. It was difficult to quantify a measurement of accuracy of the wetlands predicted in this study site, but in comparison to the already identified wetland areas (current KZN wetland layer), visual assessment corroborates that the model had in fact identified the majority of the wetland areas (Figure 6.3). In Figure 6.3 the patchy mosaic of high and low probability values indicates evidence of the dune cordons indicative of the formation of the wetland areas along the coast, creating patches of permanent to temporary wetness zones. Overestimation of highly probable wetland areas was highly likely. An evident limitation of the model in this study site was the undefined pixelated coastline visible in this case study site, which resulted from using coarse-resolution data variables in the model. This limitation could not be avoided.

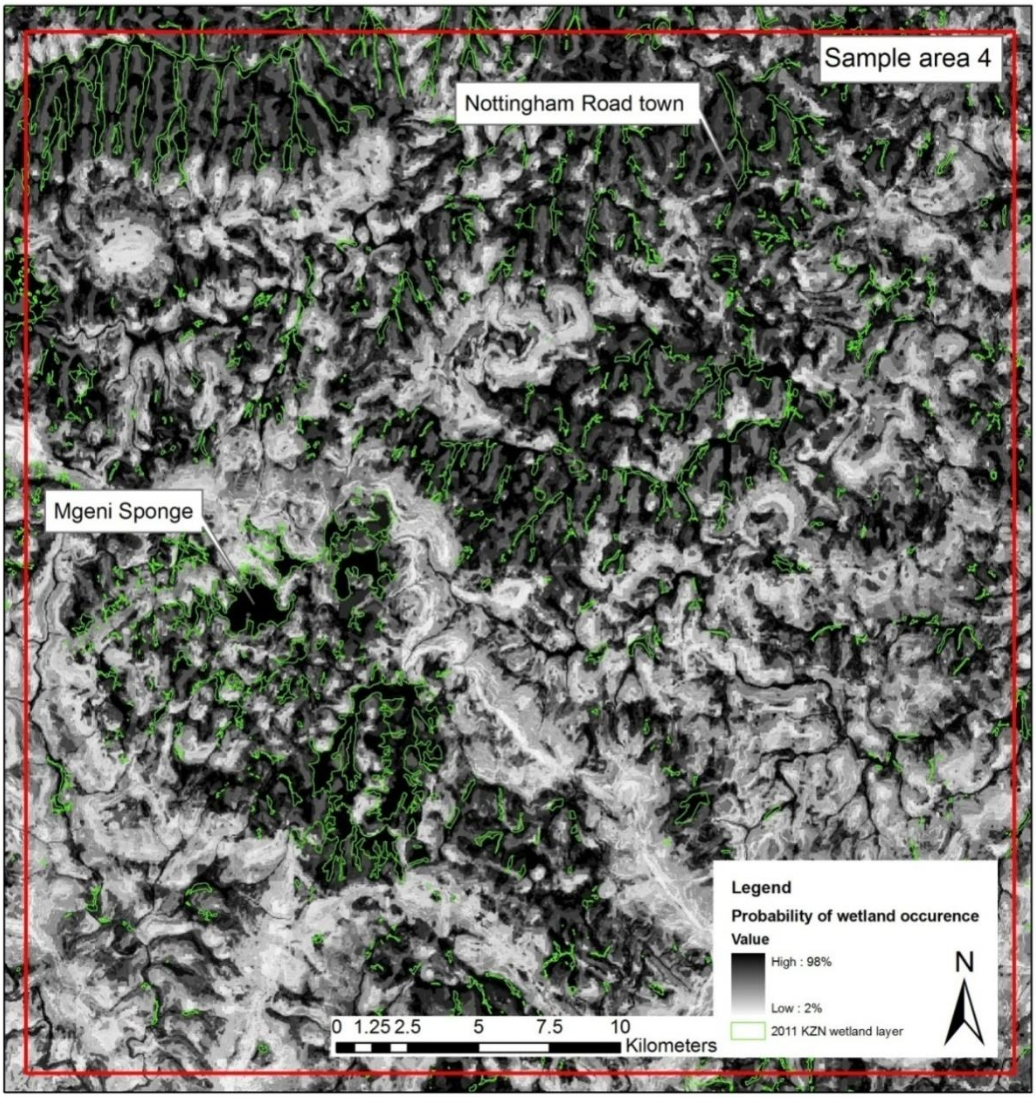


Figure 6.4: Case study Site 4 positioned in the Mgeni catchment with the Mgeni Sponge

Significance in Assessment: This area covers the Mgeni Sponge and the Mgeni River, recognised as the most important river system in KZN. This is the most extensively mapped area with regard to wetland extent.

Site 4 was regarded as the most extensively mapped wetland area in KZN in terms of the 2011 KZN wetland coverage (Escott, 2011 pers. comm.). The modelled probability layer mirrored the mapped wetland areas (current KZN wetland layer) well and indicated even more probable wetland areas (Figure 6.4). The modelled layer had been relatively successful in identifying the Mgeni Sponge formed on the Impendle Plateau of Karoo dolerite.

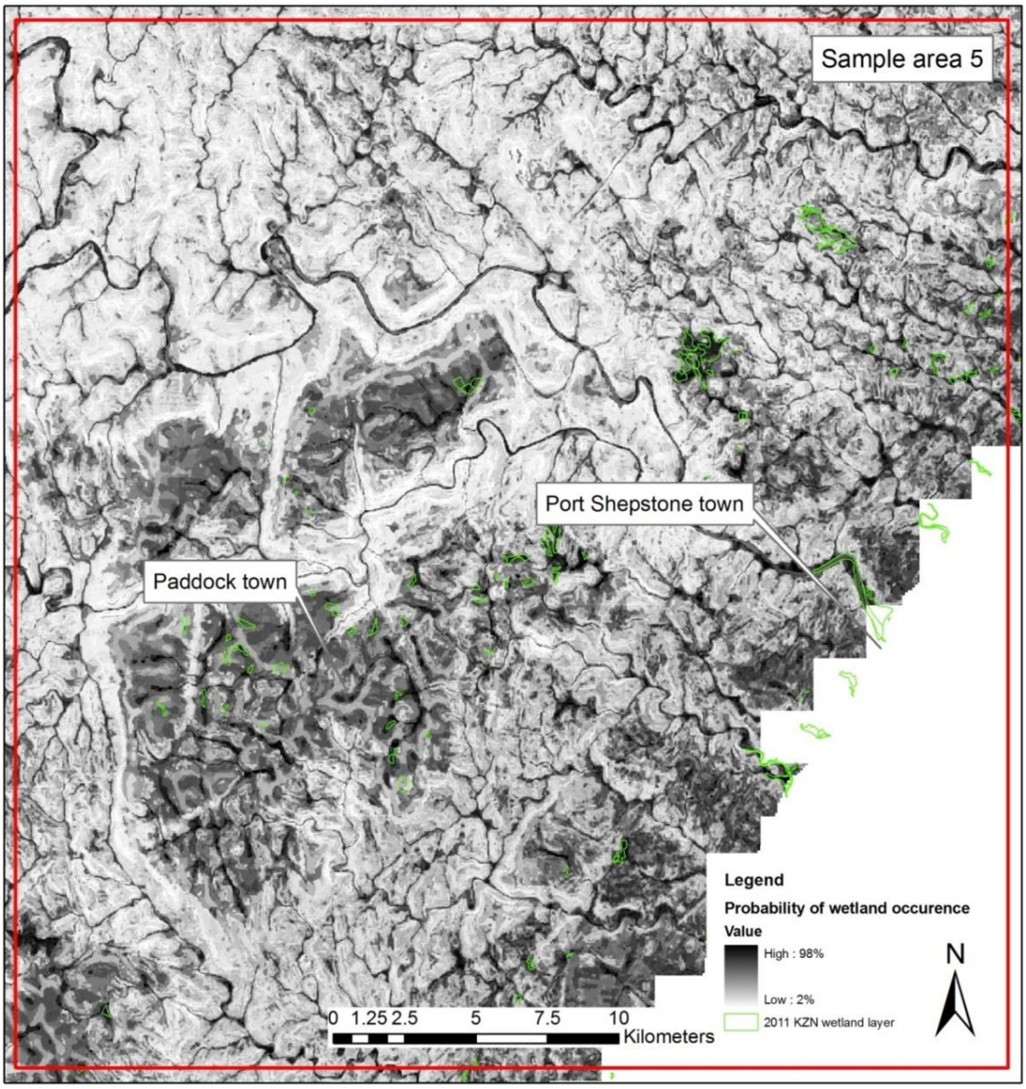


Figure 6.5: Case study Site 5 positioned in the Mzimkulu catchment, where the Mzimkulu River exits into the Indian Ocean.

Significance in Assessment: EKZNW suggests that this area requires improvement in the number of wetlands mapped (Escott, 2011 pers comm.). This area has largely been modified into agriculture.

Site 5 was regarded as a poorly mapped area in terms of the 2011 KZN wetland coverage (Escott, 2011 pers. comm.). The modification of wetlands into agricultural cropland has made it difficult to identify wetlands in this area. The general probability of the site area was quite low (between 0% and 40% probability), which could be as a result of the undulating relief found in that area. The model identified highly probable wetlands areas around the town of Paddock, which had not been identified as wetlands by the current KZN wetland layer (Figure 6.5). Again, the limitation of the modelled layer in this study site was the undefined pixelated coastline.