Student thesis series INES nr 640

The hunt for the elusive Cantharellales:

How to stake the odds in the mycophiles favour, using Species Distribution Modelling and GIS.

Photo: Maiken Frank-Hansen

Maiken Frank-Hansen

2024
Department of
Physical Geography and Ecosystem Science
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Maiken Frank-Hansen (2024).
The hunt for the elusive Cantharellales: How to stake the odds in the mycophiles favour, using Species Distribution Modelling and GIS.
Bachelor degree thesis, 15 credits in Physical Geography and Ecosystem Analysis
Department of Physical Geography and Ecosystem Science, Lund University

Level: Bachelor of Science (BSc)

Course duration: *March* 2024 until *June* 2024

Disclaimer

This document describes work undertaken as part of a program of study at the University of Lund. All views and opinions expressed herein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

The hunt for the elusive Cantharellales:
How to stake the odds in the mycophiles favour, using Species Distribution Modelling and GIS.

Maiken Frank-Hansen

Bachelor thesis, 15 credits, in Physical Geography and Ecosystem Analysis

Supervisor:
Anna Terskaia
Department of Physical Geography and Ecosystem Science,
Lund University

Exam committee:
Cecilia Akselsson
David Tenenbaum
Department of Physical Geography and Ecosystem Science,
Lund University

Acknowledgements

I would like to thank my supervisor Anna Terskaia for her help, guidance, and feedback during this thesis. Our weekly meetings and subsequent “fika” with interesting conversation always provided a relaxing atmosphere to an at times stressful period. Furthermore, I would like to thank Dr. Sergei Dudov for his valuable input and suggestions on how to improve the model output in Maxent and providing me with the necessary reference material. Thanks, should also be extended to both my examiners, Prof. Cecilia Akselsson, and Prof. David Tenenbaum for their valuable input. I would also like to recognise Martin Weibull and Karin Larsson for helping me with technical support regarding Maxent and ArcGIS Pro respectively. Finally, I would like to thank my family who have cheered me on and supported me in all the different ways only a family could.

Abstract

For centuries, people have appreciated the taste of wild mushrooms which can be found in nature throughout the European continent and the Fennoscandian peninsula. This is especially so for those of the order of *Cantharellales*, where you find the Golden Chanterelle among others. Today the commercial aspect of mushroom harvesting has resulted in regulations and guidelines set out by the Nordic Council of Ministers.

With the help of Geographic Information Systems (GIS) and Species Distribution Models (SDM) it is now possible to ascertain a probability of occurrence that helps guide the mushroom pickers to the best spots in a designated area. This thesis aims to map and evaluate how accurately the SDM model, Maxent, will predict the locations of five specified *Cantharellales* species. By combining six different environmental factors with collected sample points from the 2023 mushroom season for Svedala Municipality and external sample points, collected over a 20-year time period from GBIF and Inaturalist, a model was developed. Both are online international networks that provide open access to species data for scientist and researcher as well as land managers and the public. The model produced a good fit for Svedala Municipality and a fair fit for Scania County. The evaluation was based on Areas Under the Curve values (AUC) derived from the Receiver Operating Characteristic curve (ROC). The curves values range from 0 to 1 and the results from the respective models were a mean AUC of 0.762 for Scania County and a 0.875 for Svedala Municipality.

Of the six environmental factors, the land cover layer was expected to have the highest influence establishing the model and within the model. This was also confirmed, however, the National Land Cover Data of 2018 clearly outperformed that of the more generalised land cover of the CORINE land cover classification. Additionally, specific local characteristics of an area were deemed as rather significant, in terms of its influence and contribution of data to the model.

Overall, the model indicated that it was possible to achieve a valid prediction of *Cantharellales* mushroom occurrence using open-source datasets and SDM modelling. With the development of future web applications that help landowners track the effects their management techniques have on the fungi habitats; the growth of this specific ecosystem service will benefit both the landowners and the mycophiles.

Key Words: *Geographic Information Systems, GIS, Modelling, Species Distribution Model, SDM, Maxent, Mushrooms, Fungi, Cantharellales, The Golden Chanterelle, Cantharellus cibarius, Winter Chanterelle, Craterellus tubaeformis, Yellow Foot, Craterellus lutescens. The Black Trumpet, Craterellus cornucopioides*

Table of Contents

1 Introduction

Wild mushrooms and fungi have been appreciated and consumed by humans for hundreds of years (Hall et al., 2003; Heilmann-Clausen et al., 2015; Kozarski et al., 2015). Among the top five, two in particular are sought after by both mushroom enthusiasts and gourmet chefs; *Cantharellus cibarius*, commonly known as the Golden Chanterelle and *Craterellus cornucopioides*, known by several different names, such as the Black Trumpet, the Trumpet of Death and Horn of Plenty (Holmberg & Holmberg, 2021). For a long time, mushrooms of the order *Cantharellales*, have been intensively harvested in the forests of Europe for commercial purposes (Peintner et al., 2013; Ribeiro et al., 2009). The Nordic Council of Ministers of the Scandinavian countries have deemed it necessary to set mushroom guidelines and issue advice for four *Cantharellales* species; The Golden Chanterelle (*Cantharellus cibarius*), Winter Chanterelle (*Craterellus tubaeformis*), Yellow Foot (*Craterellus lutescens*), The Black Trumpet (*Craterellus cornucopioides*) (Gry & Petersen, 2012), with the exception of the Terracotta *Hedgehog* which is not included in the regulations. All five will be described in more detail in this thesis.

Despite attempts to cultivate one of these wild mushrooms and arguably one of the most sought mushrooms, the Golden Chanterelle (*Cantharellus cibarius*), the efforts to grow them commercially have proven fruitless due to the complex interactions between several factors such as the presence of certain bacteria, foreign microbes and the interrelationship linking Chanterelle mushrooms and its host trees (Kozarski et al., 2015). Therefore, the efforts to increase the chances for the mycophile, which is a wild mushroom or fungi enthusiast, whose hobby it is to hunt and forage for them (Peintner et al., 2013), could be considered by many of a somewhat high importance.

Though it is possible to identify studies about commercial modelling of mushrooms, where the use of fuzzy logic is applied in the cultivation of oyster mushrooms (Cikarge & Arifin, 2018; Faizollahzadeh Ardabili et al., 2016), applying parameters such as temperature and timed irrigation (specifically in greenhouses), there seems to be a gap when it comes to studies and reports about modelling predictions of forest fungi fruit bodies. For this purpose, Geographic Information Systems (GIS) in combination with Species Distribution Modelling (SDM), using presence only data and environmental factor layers, will be used for identifying the ecological niche and predicting locations for the 5 *Cantharellales* mushrooms in the forests of the Svedala Municipality and more widely in Scania County. One thesis with a similar topic of habitat modelling by Aina (2022) exclusively based the modelling on environmental parameters and did not include any presence data, focusing predominantly on results that could be correlated with ecosystem services, by applying a combination of the Analytical Hierarchy Process (AHP), Weighted Overlay Analysis (WOA) and the Pairwise Comparison Matrix method (PCM). Other studies such as the ones by Dramani et al. (2022) or Khaund and Joshi (2016) focussed more on habitat conservation or the development of mapping methods to optimise the harvesting of *Cantharellales* for the benefit of the local or indigenous people of their respective region. With the study of Khaund and Joshi (2016) taking place in the state of Meghalaya, India and the study by Dramani et al. (2022), carried out in the country of Benin in Africa, only Aina (2022) has environmental factors with ranges that are similar to those of Svedala Municipality and Scania County.

This thesis hopes to contribute to, and in some way, fill the mentioned gap to encourage others, with a basic or more in depth understanding of GIS and modelling, to explore the possibilities these programs provide for predicting mushroom locations.

1.1 Aim

The aim of this bachelor thesis is to map the locations of possible sites where *Cantharellales* mushrooms can be found with the use of SDM and GIS. Two objectives have been established to accomplish this aim:

The first objective is to evaluate how accurately the SDM model will predict the locations of the specified *Cantharellales*.

The second objective is to investigate if different land cover- and soil moisture layers will influence the model predictions. Land cover is considered the most influential influence on mushroom location according to literature, while soil moisture is relevant due to its availability as two differently classified layers of the same dataset.

The objectives will be achieved by utilising specific environmental factor variables, such as said land cover- and soil moisture layers, topographic layers, soil type layer, and collected presence data samples points, to the species distribution model Maxent. The produced outputs and results will be analysed as well as visualised with the help of tables, maps, and graphs.

1.2 Motivation

The definition of *Cantharellales* mushrooms for this bachelor thesis are defined by the author as wild mushrooms that grow in groups within a smaller area, that can be eaten raw (Elliott et al., 2022) or eaten without any other preparation than frying in a pan. Furthermore, wild mushrooms are officially classified as some of the best edible fungi one can find in the country of Sweden.

An additional reason for writing this thesis is to make a model and a methodology that is more accessible for mycophile to utilise and thus with the help of GIS and SDM identify possible locations where the good mushrooms can be found.

1.3 Outline of Report

The thesis is structured in the following way, with sections on Background and Theory, Data and Methodology, Methods, Results, Discussion, and Conclusions. The Background and Theory in short describes fungal structure, the five different *Cantharellales* species, the species distribution model, and the designated study area. The Data and Methodology section outlines the different environmental factors as well as providing an overview of the Maxent model including which method to use for evaluating its results. The Methods section is a description of the different steps that were involved in producing a valid model. The Results displays and analyses the output from the best-fit model, while the Discussion includes and evaluates the results, while also covering certain pros and cons, limitations as well as other aspects related to modelling fungi. Finally, the Conclusion lays out the principal findings of the report.

2 Background and Theory

To allow a general understanding of the interconnection between mycorrhizae, and how that relates to the *Cantharellales* fungi, the first part of this section provides some of the available theory on fungi mycorrhizae. This is done to highlight that the fruit bodies of a mushroom is only a small part of a larger organism that is depend on a specific type of land cover. The five selected *Cantharellales* species and the Species Distribution Model are also described in more detail. Lastly, an introduction to the study areas, Bokskogen and Holmeja forest are also included.

2.1 Mycorrhizae

Although we might not see it, wild mushrooms and their mycelium and subsequently the mycorrhizae network are a huge part of the ecosystem (Mossberg et al., 1987). Mycelium intertwines the soil substrate (called the hartig net (Söderström, 1992)), dead plants and decomposing animals, breaking it down into smaller pieces. They can also be considered living and functioning as a parasite or coexisting in a symbiotic relationship. In all, fungi play a vital role in the nutritional supply for many organisms. One type of fungi that produces fruit bodies above ground can be labelled as mycorrhizae mushrooms (from the Greek word *mykes* meaning: mushrooms and *rhiza* meaning: root). A lot of plants, but especially forest tree species, coexist or live in a symbiotic relationship with fungi. The mycelium surrounds the roots of the tree and the tree as a result benefits, as this facilitates the nutrient uptake from the ground (called the mantle (Söderström, 1992)). Subsequently, this in turn benefits the agroforestry business (Mossberg et al., 1987). According to Smith and Read (2008) the majority of boreal and northern temperate forest tree species are considered to be living in a symbiotic relationship with ectomycorrhizal fungi (EMF). Aside from the fruit bodies of the mushroom (sporocarps), the EMF is comprised of the ectomycorrhizal extrametrical mycelium (EMM), the mantles and the hartig nets (Söderström, 1992). Additionally, Söderström (1992) notes that hundreds of metres of EMM have been identified within a single gram of humus soil.

2.2 *Cantharellale* mushrooms

Of the 29000 species of basidiomycotes (a division of fungi) that are known, the fungi that belong to the *Cantharellales* order belong to the biggest subclass, the agaricomycetes (Margulis & Chapman, 2009). The basidium is the club-shaped microscopic reproductive structure, which is also where the division gets its name, and is what differentiates it from other species of fungi and from different phyla (taxonomic rank at the level below Kingdom and above Class). All basidiomycotes are considered heterotrophs, that is, they get their energy which they need to survive from other organisms, such as agricultural crops and forest trees. A majority form a symbiotic relationship, termed ectomycorrhizae, as previously described. The main benefit for both parties is that the fungal hyphae, which are the branching filaments that the mycelium consists of, can reach further into the soil substrate in the pursuit of nutrients, such as phosphorus and nitrogen. Through the hartig net, the nutrients and other inorganic nutrients are migrated from the fungi to the plant host. In return, the fungi receive vital carbohydrates, derived from photosynthesis of the plant (Margulis & Chapman, 2009).

The environmental criteria of wild mushrooms are very different. Some species will only grow in very specific locations, while others are able to grow in both coniferous- and deciduous forests, yet under very different environmental conditions (Mossberg et al., 1987).

2.2.1 Golden Chanterelle (*Cantharellus cibarius*)

Figure 1: A Golden Chanterelle (Cantharellus cibarius) captured on the 18th of Sep. in Bokskogen a deciduous forest dominated by beech and oak trees.

Of all the wild mushrooms, the Golden Chanterelle (Figure 1) is by many considered the most known and desirable by mycophile (Holmberg & Näslund, 1988). It is possible to find in most environments, such as beech-, oak-, birch-, pine- and coniferous forests, even along trodden paths and roads. It does however require the nearness of a tree as it is a mycorrhizae mushroom (Mossberg et al., 1987). Like the rest of the wild mushrooms described below, the Golden Chanterelle grows in a group setting (Holmberg & Näslund, 1988). The colour of the Golden Chanterelle is egg yolk yellow and the transition between the stem (stipe or stalk) and the cap (pileus) is hard to define (Mossberg et al., 1987). The underside of the cap is filled with rills and shallow ridges and the cap itself is at first rounded when the fruit body is young and later takes the shape of a funnel as the fruit body matures. The cap can grow as wide as ten centimetres and the stem between five to ten centimetres (Mossberg et al., 1987). The season for picking Golden Chanterelle starts at the beginning of July and ends at the end of November (Holmberg & Näslund, 1988; Mossberg et al., 1987).

2.2.2 Winter Chanterelle (*Craterellus tubaeformis*)

Figure 2: Winter Chanterelles (Craterellus tubaeformis) captured on the same date and in the vicinity of the Golden Chantrelle (Cantharellus cibarius) in Figure 1.

Despite its name, the Winter Chanterelle (Figure 2) is a prominent wild autumn fungusi. Like the Terracotta Hedgehog, it grows in coniferous forests with mossy ground cover (Mossberg et al., 1987). The cap (pileus) displays a grey-brown colouring, thus disguising itself amongst any brown litter that may cover the surface, making it hard to spot. However, once a few have been identified in an area, one will most likely find several others as one of its characteristics is that it grows in larger numbers. The branches of the mycelia of the Winter Chanterelle can spread for several hundred kilometres Mossberg et al. (1987). The Winter Chanterelle can also grow within areas that have a high concentration of limestone while at the same time thriving in rather acidic environments (Mossberg et al., 1987). It is found throughout all of Scandinavia and Finland, although it is less common in the northern parts. It has been a highly sought wild mushroom in the Nordic countries since the start of the millennia, despite being viewed as less than desirable on the continent (central Europe)(Holmberg & Holmberg, 2021; Mossberg et al., 1987). It is easily mixed up with the Yellow Foot, but this is a minor matter as both are considered excellent for eating. The cap (pileus) is thin and not as 'fleshy' as, for instance, that of the Golden Chanterelle and, when it is new, will present with a slight inward bend with a small indentation of the middle, with a brownish colour (Mossberg et al., 1987). As it matures, the indentation will be more pronounced, and the border of the cap will exhibit a wave-like shape.

It can grow to a width of six centimetres and the stem (stipe or stalk) can grow to between two to seven centimetres tall, which is hollow throughout and flattens as the mushroom matures, with the colour ranging from yellowish to yellow-green (Mossberg et al., 1987). The growing season ranges from the beginning of September to the end of November (Holmberg & Näslund, 1988; Mossberg et al., 1987).

2.2.3 Yellow Foot (*Craterellus lutescens*)

Figure 3: Yellow Foots (Craterellus lutescens) documented on the 8th of Sep. in a section of beech forest in the Holmeja forest.

This wild mushroom is, as mentioned above, easily confused with the Winter Chanterelle, whilst also being a highly sought-after wild mushroom throughout Scandinavia (Mossberg et al., 1987). The cap of a Yellow Foot (Figure 3) resembles the form of the Black Trumpet but is more flared out with small wart-like protrusions on top. The underside of the cap will predominantly showcase an orange-yellow colour. Like the Winter Chanterelle, once one has been located, several can be found in the near vicinity of the first one discovered, being very good at hiding amongst the ground litter. It grows in damp conditions, often near or in the vicinity of marshes, and for the most part in coniferous forest (Mossberg et al., 1987). The cap (pileus) will grow between three to six centimetres in width and the stem (stipe or stalk) can reach the height of five to ten centimetres (Holmberg & Näslund, 1988). They are usually found at the beginning of August to the middle of October (Holmberg & Näslund, 1988; Mossberg et al., 1987).

2.2.4 Black Trumpet (*Craterellus cornucopioides*)

Figure 4: Black Trumpets (Craterellus cornucopioides) located on the 1st of Sep. in the area the Yellow Foots in Figure 3.

The Black Trumpet (Figure 4), with its very characteristic cap (pileus), could be likened to the shape of the horn on an old gramophone, hence it is also known as The Horn of plenty. It is also the shape of the mushroom that is the reason for its given Latin name, *cornucopioides*, which in Latin means horn (Holmberg & Näslund, 1988). It can be found in several types of deciduous forests, such as oak, hazel and beech, but it is also possible to find it in mixed coniferous forests near birch and aspen trees. Seen from above, the cap (pileus) gives a sort of brown-blackish complexion with small scaly structures and the width of the cap ranges between three to eight centimetres (Holmberg & Näslund, 1988). The more water the mushroom body contains the darker the colour will be and vice versa (Holmberg & Holmberg, 2021). The stem (stipe or stalk) features a grey-black colour and with older specimens it can even seem whitish due to spore coverage (Holmberg & Näslund, 1988). It is possible to find the black trumpet from the beginning of August to the end of October (Holmberg & Näslund, 1988; Mossberg et al., 1987). The Black Trumpet is considered a favourite amongst gourmet chefs (Holmberg & Holmberg, 2021).

2.2.5 Terracotta Hedgehog (*Hydnum rufescens*)

Figure 5: Two Terracotta Hedgehogs (Hydnum rufescens) found on the 9th of Aug. in the moss-covered undergrowth of a segment of spruce trees in the Holmeja forest.

The Terracotta Hedgehog (Figure 5) is a wild mushroom that can grow in a variety of environmental conditions. It will grow in the southern parts of the beech forests of Scandinavia as well as in coniferous forest with moss-covered undergrowth (Mossberg et al., 1987). It is even possible to find it in alpine birch forest. Notably, it is commonly found in areas with high levels of ground moisture. Despite its very characteristic teeth on the underside of the cap (pileus), the Terracotta Hedgehog is of the same family as the Golden Chanterelle, which as mentioned above, has gills instead of teeth (Holmberg & Holmberg, 2021). They do however fill the same function for spore dispersal. The cap will display a colour that is a mixture of orange and yellow, with the same colouring of its teeth, and with a size of up to seven centimetres in width. The stem (stipe or stalk) can grow to a height of two to seven centimetres, with a width of one and a half centimetres (Mossberg et al., 1987). The Terracotta Hedgehog usually grows from August to the end of October (Holmberg & Näslund, 1988; Mossberg et al., 1987).

It is worth noting that all of the above-mentioned wild mushrooms are very rarely attacked by insect larvae (Holmberg & Näslund, 1988; Holmberg & Holmberg, 2021).

2.3 Species Distribution Modelling

Species Distribution Modelling (SDM) was originally developed as a conservation mapping tool in addition to inventorying resources (Miller, 2010). Over the years however, it has evolved into a multitude of different connotations along with the expanding availability and variations of environmental data, statistical methods, and the employment of digitisation in the study of life, as they are structured in GIS.

As a concept, SDM works by quantifying the correlation between a set of given environmental factors (layers) and the linked distribution of flora and fauna (Miller, 2010). The output or "environmental profile" extracted from the SDM is a valuable tool that can be used to describe the significance of specific environmental factors, influencing the species distribution and, predict species' distribution in an area where no samples have been collected (Franklin, 2010). SDM is also used to examine environmental change and the resulting consequences of those changes to the specific ecosystem in question (Franklin, 2010).

When using any type of SDM application, it is very important to have a clear idea about what one is specifically trying to model or map and the associated delineations in addition to any set premises (Miller, 2010).

Despite the fact that the ecological niche (for more information please read Franklin (2010)) is a vital concept within the framework of SDM, it is instead "habitats" that more correctly relate to what is frequently modelled in SDM applications.

A habitat map that results from the correlation between the given set of environmental factors and the input of known sample points will additionally produce a curve or gradient. The curve or gradient is essentially comprised of several gradients that each represents the species' response to the conditions of the environment along the gradient, and therefore characterise the dispersal in a particular habitat and special environment. The curve is typically characterised by a minimum and maximum threshold, but also includes a critical value where the curve "flattens out", indicating a limit to the expansion of the modelled species (Franklin, 2010).

There are three types of environmental gradients, direct (e.g. temperature and pH), resource (resources that are used or consumed such as nutrients) and indirect, which would be a gradient that correlates due to it being location-specific (e.g. elevation and latitude)(Austin, 1980). Miller (2010) defines how the ecological niche concepts are translated into the SDM in the following way: "A simple way of relating the concept of an ecological niche to SDM is that the niche describes a species' fitness in environmental space, a statistical method quantitatively describes that environmental profile, and the resulting predicted map translates the environmental profile into some measure of suitability in geographic space".

Thus, the advantage of SDM and the combination of GIS, is how the model renders the relationship between a given species and its environment to its placement in the geographic space (Miller, 2010). The most frequently used data inputs in SDM are biological- and environmental data. The biological data will describe the distribution of one or more species and are usually measured at three levels, nominal, such as presence/absence or type, ordinal, like ranked abundance (which is a component of biodiversity) and ratio, as in abundance or richness (Miller, 2010). With several input layers a ratio of 75:25 of the samples from the collected biological data is recommended in order to test the validity of the model, that is, 75% of the sample points for modelling and 25% of the sample points for testing the validity of the model (Franklin, 2010). According to Franklin (2010) the equation for calculating the percentages of testing and training respectively is:

$$1/(1+\sqrt{(p-1)}) \tag{1}$$

where p represents the number of predictors (factors/environmental layers). Therefore, if five predictors are used the ratio would be 67% and 33%.

Of the environmental data, the subsequent are frequently used: climate data, topography, geology, soil type and land cover (Miller, 2010). Although GIS is crucial for the processing, handling and management of input and output data, which is used and produced by the SDM, the statistical analysis is predominantly handled in a separate software or program. Lastly, three main categories of modelling algorithms exist. First, there is traditional regression, for instance, generalised linear models (GLM), generalised additive models (GAM) and multivariate adaptive regression splines (MARS). Secondly, machine-learning models and thirdly, maximum entropy which deals with presence-only datasets (Miller, 2010).

2.4 Study area

Both Bokskogen and Holmeja forest (Figure 6), which are the forested areas where data points for this thesis have been collected, are located in the Municipality of Svedala. The Municipality of Svedala is one out of 33 in the county of Scania, and lies centrally in the southwest of Scania, located east of Malmö Municipality and south of Lund Municipality (Länsstyrelsen Skåne).

2.4.1 Bokskogen (Torup)

The forest in Bokskogen is dominated by beeches, some of them more than 150 years old, but also includes oak, ash elm and other deciduous tree species (Malmö Stad, 2023). The undergrowth is covered during spring by white and yellow wood anemone (*Anemonoides nemorosa* and *Anemone ranunculoides*) and later during the summer by Common Dog-violet (*Viola riviniana*) and Yellow archangel (*Lamium galeobdolon*) (Malmö Stad, 2023). It is also possible to see ground cover consisting of Dog’s Mercury (*Mercurialis perennis*) a plant which is commonly found in woodland with a tree composition that is distinctive of Bokskogen, as it prefers to live under oak, beech ash and elm (Malmö Stad, 2023). There are also several other rare and endangered plant species one might encounter, such as the Lesser Hairy Brome (*Bromus benekenii*), the protected Wild Bird's-nest Orchid (*Neottia nidus-avis*) in addition to several types of wild mushrooms and lichens (Malmö Stad, 2023).

Part of the forest is passively managed and consequently is left to develop freely without human interference. Thus, old trees that have fallen over or a snag (a dead tree left standing upright), are left to decompose naturally, as they contain a myriad of life such as insects, larvae, fungi, and bacteria (Malmö Stad, 2023). Some of the insects such as the Large Black Longhorn Beetle (*Stictoleptura scutellatae*) and the Red-horned Cardinal Click Beetle (*Ampedus rufipennis*) are red-listed (Brunet, 2005). Since May of 2019 parts of the forest have been converted into a nature reserve (Malmö Stad, 2023). Roughly a third of the forest is protected in accordance with Natura 2000, a network of protected areas covering Europe’s most valuable and threatened species and habitats according to Malmö stad. More specifically it is the beech forest and its rich field flora that have been inducted (European Environment Agency, 2023). In the old days, the forest was used for grazing and fodder for the animals of the estate of Torup, as well as for collecting firewood and timber (Malmö Stad, 2023). The forest is owned by the city of Malmö and covers an area of approximately 350 ha (Brunet, 2005).

2.4.2 Holmeja skogen

The Holmeja forest has belonged to the Skabersjö estate since the mid-1660s (Brunet, 2005) and is situated in the eastern part of the Svedala Municipality (Jansson, 2006). The forest is estimated to be about 1400 ha and is dominated in large part by the lake, Fjällfotasjön (Jansson, 2006). The lake itself is surrounded by peatland in the northern parts and elsewhere by unrestrained deciduous swamps, with areas of higher elevation covered by beech- and spruce forest (Jansson, 2006). The owners of the estate conduct agroforestry management practices with periodic logging and clearcutting throughout the entire forest (Jansson, 2006). Additionally, part of Skåneleden runs through the forest’s northern section and connects with Bokskogen in the west, as part of the North to South trail (Region Skåne, 2024). The forest in general is a mixture of planted deciduous- and coniferous forest, with species such as beech, spruce, pine, larch and birch (Jansson, 2006). Historically, the landscape has been used for hunting game and fishing, which in some respects are still currently in place, however, agroforestry is presently the main influence on the surrounding ecosystem (Jansson, 2006).

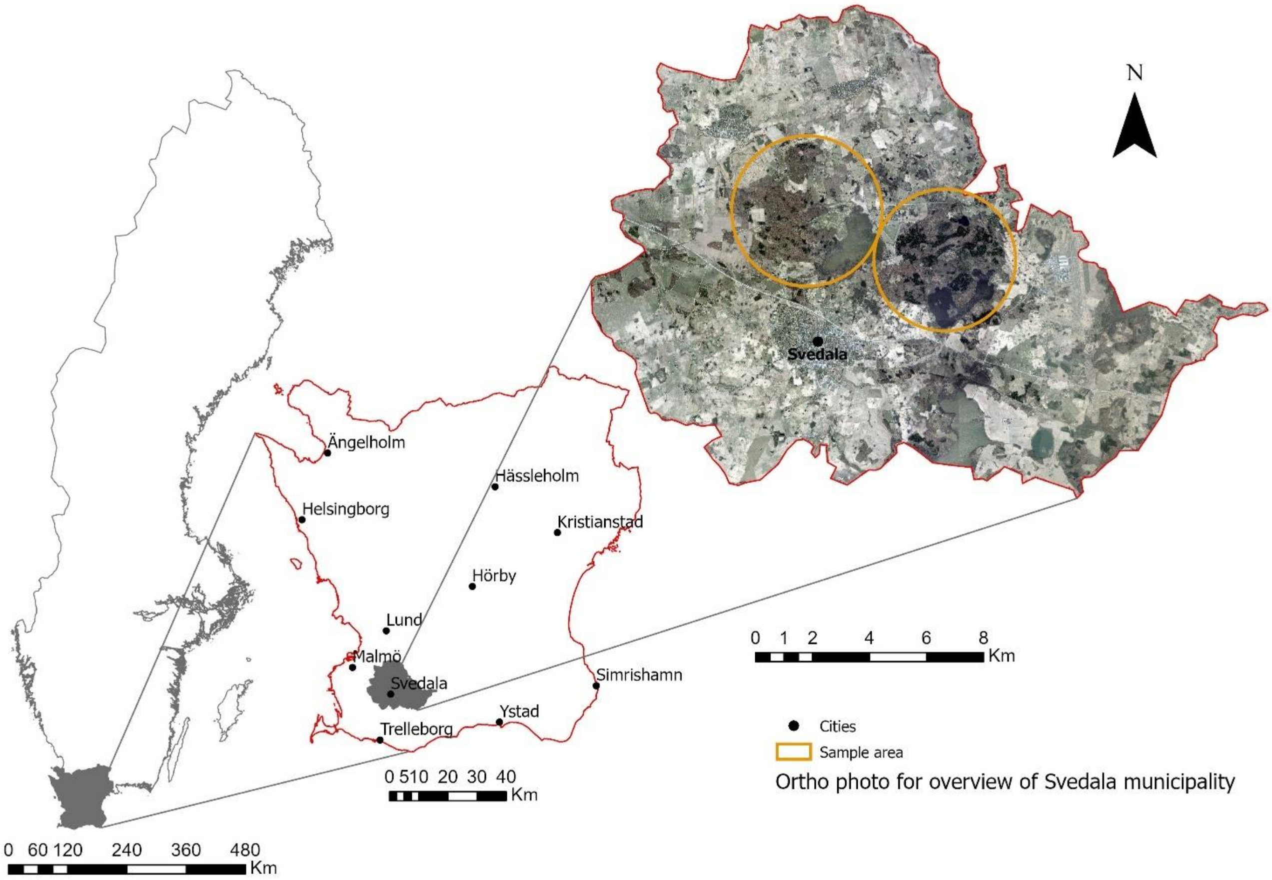

Figure 6: The two forest areas, Bokskogen and Holmeja forest, circled in yellow in the part of Svedala Municipality, where the sample points for the thesis were collected.

3 Data and Methodology

The modelling will be carried out using the Maxent Java application, which is self-contained, meaning the jar file can be run as an exe-file (Phillips et al., 2017). The primary ecological factors needed for this type of modelling was identified by both Dreisbach et al. (2002) and Phillips and Dudík (2008) as the following: vegetation cover, topography, soils and remotely sensed variables such as temperature and precipitation. An additional factor that should be considered, is the mycorrhizae and more specifically the ectomycorrhizal fungi (EMF) of the above-mentioned fungi. Hall et al. (1997) defined a habitat as "the resources and conditions present in an area that produces occupancy - including survival and reproduction - by a given organism". Because certain EMF are tightly linked with the presence and absence of certain tree species, these tree species can therefore be used as an indicator for the potential of several species of wild mushrooms as stated by Dreisbach et al. (2002). However, due to the lack of available data with such specificity, the model will not be able to produce any output that could be analysed for this specific purpose in mind.

3.1 Maxent

The choice of the program Maxent for modelling the potential presence of *Cantharellales* mushrooms, was in part the result of two characteristics, its primary use of presence-only data and its free availability. This would make it possible for a larger group of people, who might want to use the methodology described in this thesis, to do so with little or no hindrance.

As a self-contained Java application for SDM, Maxent has based its modelling predominantly on occurrences records combined with environmental variables in other words land cover and topography (Phillips et al., 2017). As such, many have used it and articles such as the journal article by Phillips et al. (2006) have been cited more than 6000 times according to Phillips et al. (2017).

Maxent uses machine learning response and as such is meant to model predictions from what in the world of machine learning would be considered fragmented data (Baldwin, 2009). As such, Maxent estimates what it considers the most uniform grouping, i.e. maximum entropy of a given set of sampling points (Phillips et al., 2006). Since the maximum entropy algorithm is viewed as deterministic, it will display a distribution which intersects at the maximum entropy probability distribution according to Phillips et al. (2006). In other words, the output that Maxent produces represents how significantly better a proposed model fits the sample data in comparison to a more uniform distribution. In addition to this, a further advantage of Maxent is the possibility of using both continuous and categorical factors (Phillips, 2005). However, as with most SDM programs, Maxent tends to "overfit", which means it predicts distributions which are centred around the sample points (Baldwin, 2009). This can be corrected in the settings, but simulations done by Phillips and Dudík (2008) have shown that default configurations execute to the same level as adjusted settings. There are four different output formats: raw, cumulative, logistic and cloglog. Prior to the release of the 3.4.0 version of Maxent, the logistic format was the recommended output format as it provided an estimate of the probability of occurrences based on the incorporated environmental factors (Phillips & Dudík, 2008). However, since the release of the 3.4.0 and later versions of Maxent, the introduction of the cloglog output, has replaced the previous output formats as the most advocated option (Phillips et al., 2017). With a range of 0 to 1 (score) as explained further by Phillips (2005), it is recommended that the results are projected into a GIS program as it helps with the interpretation. Accordingly, a score of 1 would indicate a high probability of occurrence while 0 would indicate zero probability of occurrence. As with any SDM program, Maxent being one, the importance of knowing how each factor influences the presence of a given modelled species cannot be understated (Baldwin, 2009). It is equally important to know, which factor might have the greatest influence on the model, and in what way (Baldwin, 2009). There are two ways this can be accomplished. First, Maxent provides the contribution of each factor in percentage (Phillips, 2005). But it is of note, that this method tends to indicate a higher importance for highly correlated layers (Baldwin, 2009). The second method is the jackknife technique (Young et al., 2011). This alternative is performed when enabled in the program. It excludes one factor at a time when running the program model (Baldwin, 2009). Hence, the information it provides is on how big a part of unique information each variable provides the model. Baldwin (2009) observes that it also indicates the performance of each factor in terms of its importance in explaining the species distribution. The program primarily requires the following input data; geographical sample data of species occurrences, environmental factors and thirdly, which is optional, a sampling bias grid (Syfert et al., 2013). A sampling bias grid, corrects the model by applying 'background data' that consists random locations which have been weighted by the sampling bias grid (Phillips et al., 2009).

In ArcGIS Pro version 3.1 and newer, a tool called Presence-only Prediction (Maxent) can be accessed as part of the Spatial statistics toolbox (for more information please see (Esri, n.d.)).

3.2 ROC (AUC)

The AUC, depicted as the area under the under the receiver operating characteristic (ROC) curve represents the measurement of the discriminatory capacity of a classification model (Jiménez-Valverde, 2011). A ROC plot can be illustrated in a diagram by entering available sensitivity values, which would be true positive fractions, on a y-axis against their equivalent false positive fractions (1- specificity (*Sp*)) on the x-axis (Fielding & Bell, 1997). This would be done for all available thresholds.

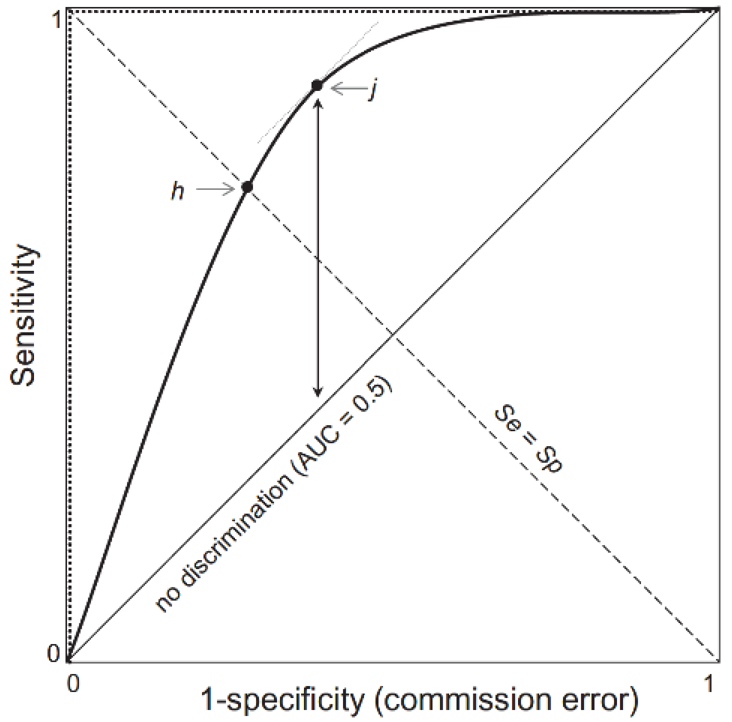

Figure 7: The idealized receiver operating characteristic (ROC) curves used for evaluating a Species Distribution Models (SDM) adapted from Jiménez-Valverde (2011).

Figure 7 depicts the idealised receiver operating characteristic (ROC) curves. Indicated by the dotted line is the curve of perfect discrimination. Reversely, the thick black line represents the curve of a model showing imperfect discrimination. Where the diagonal line illustrates no discrimination, which would equal area under the ROC curve (AUC) = 0.5 and where perpendicular specificity (*Sp*), shown as a dashed line, equals sensitivity (*Se*). *h* would be the point where the ROC curve crosses *Se* = *Sp* and where the tangent of the curve, *j*, equals 1 (thin grey line)(Jiménez-Valverde, 2011).

The curve and its significance would be specified by the area under the curve (AUC), which is unitless (Baldwin, 2009). The generated values vary from 0.5 to 1.0, with values near 0.5 indicating a fit that would be no better than what could be expected if it was random. Values close to 1.0 or equal to 1.0 would indicate an almost perfect or perfect fit. Although rare, it would be possible to achieve a result lower than 0.5, which in turn would mean that the model fits worse than random according to Engler et al. (2004). In general, the value of the AUC would not reach the value of 1.0 (Fielding & Bell, 1997).

3.3 Environmental factors

The following subsections describe each individual environmental factor that was entered into the Maxent model. Further information and sources for each layer, as well as additional layers and data used for enhancing the interpretation of the maps, can be found in Table A1 of the appendices. For clarification, regarding maps and illustrations, those were all produced by the author using the available data or the subsequent data output produced by the author via the subsequently mentioned programs.

3.3.1 The Digital Elevation Model (The GSD-Elevation data, Grid 2+)

The Digital Elevation Model (DEM)(Figure 8A) utilised in this thesis consists of a terrain model in a grid with the resolution of 2 metres (Lantmäteriet, 2019a). The elevation data was collected with the use of an aerial laser scanner and subsequently processed into a terrain model in the form of a raster layer. The calculation was done through a linear interpolation in a Triangulated Irregular Network (TIN). The terrain model was based on laser points which were classified as ground and water points. The model has been continuously updated with the help of aerial image matching in addition to laser data. The accuracy of the DEM, based on the points measurements, in the plane and height was 0.3 metres and 0.1 metres respectively as described in the accompanying metadata (specifically for data captured with Aerial Laser scanning (LLS). For Aerial image matching (FBM) the numbers differ with a dependency on image resolution. Thus, the accuracy in the plane would be 1 time the image resolution and the accuracy in height 1,5 times the image resolution. The height projection of the data was provided in RH2000. In accordance with the HMK's (Swedish handbook in surveying and mapping [Handbok i mät- och kartfrågor]) classification system, the terrain model is counted as standard level 1 (Lantmäteriet, 2019a). The subsequent slope- (Figure 8B) and aspect (Figure 9) layers were based on the DEM layer as described in the methodology.

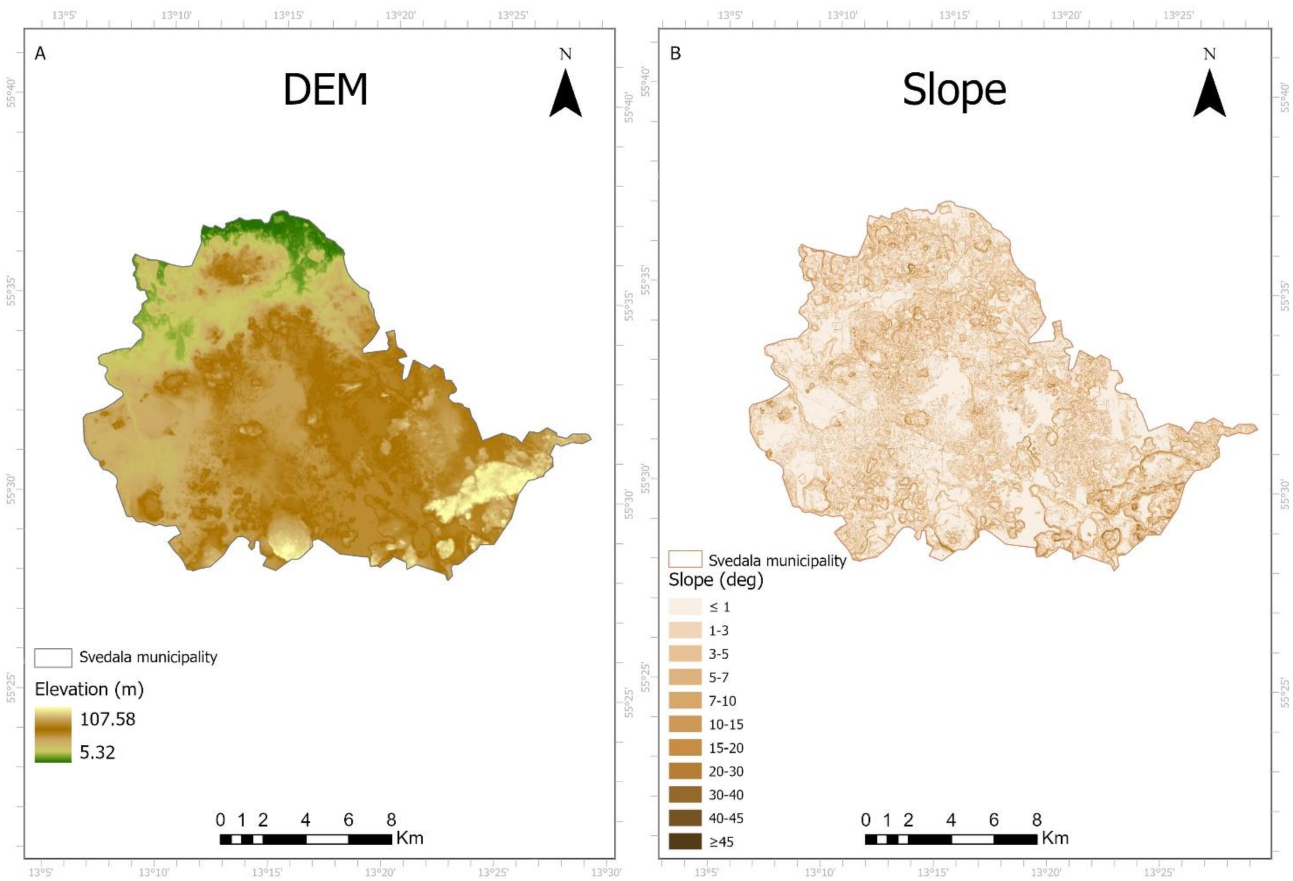

*Figure 8: **A** shows the DEM for the Svedala Municipality at a 10m spatial resolution, indicating a maximum elevation of around 108m and a minimum around 5m. **B** depicts the different slopes in degrees and their variations within the Municipality. From nearly level (≤1°) to very steep (≥45°).*

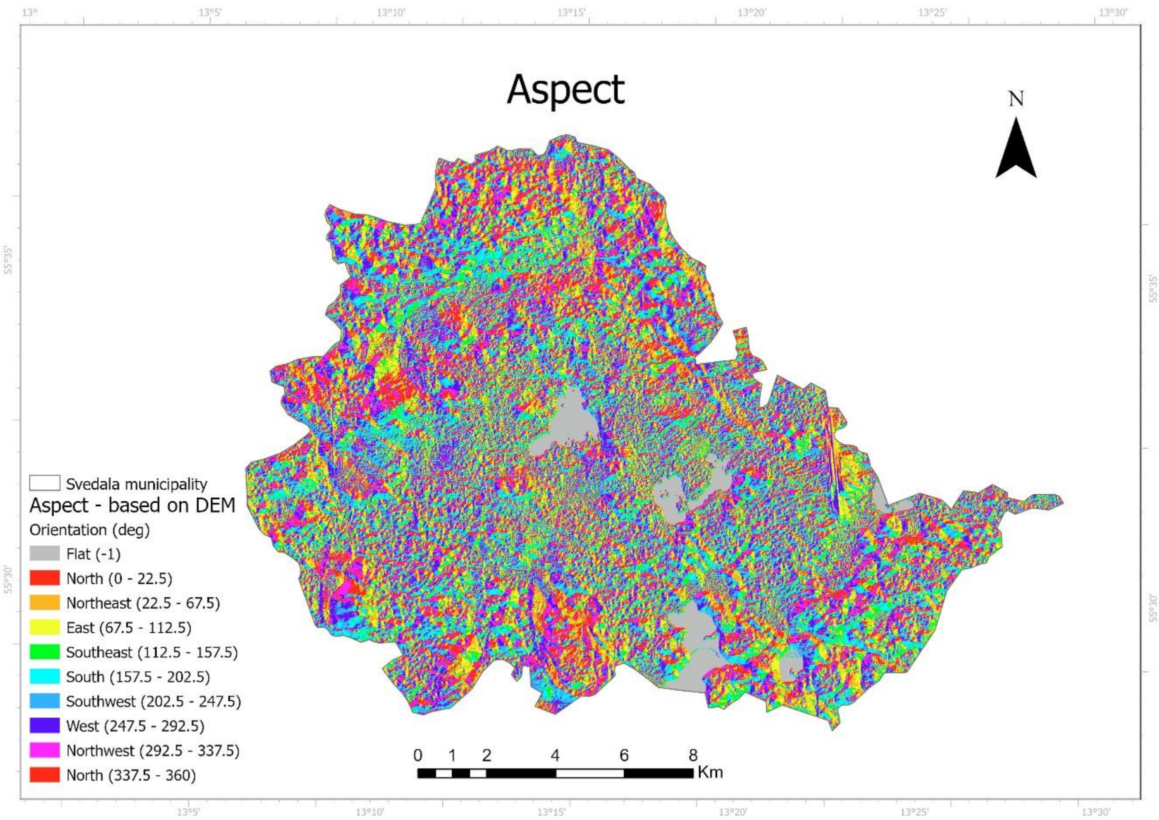

Figure 10: The aspect based on the DEM layer identifies the orientation (compass direction), in degrees, that the corresponding downhill slope faces for each location.

3.3.2 **Soil map (**Soil types **[**JORDARTER] 1:25 000-1:100 000**)**

The main soil map layer (Figure 10), designated JG2, provided by the Department of Geological Survey of Sweden, was produced to convey a comprehensive picture of the earth deposits in and around the ground surface (Sveriges geologiska undersökning, 2018). The layer depicts the type of soil which would normally be found at the mapped soil depth of 0.5 metres, although the actual depth of the soil layer likely surpasses this for most areas. The layer was provided in vector format and contained an attribute table with among other information, a code for the soil type plus a text description. The borders between different areas of earth deposits should be viewed as transitional or less defined in the terrain since the transitional zone can cover a distance or width of 50 metres or more (Sveriges geologiska undersökning, 2018).

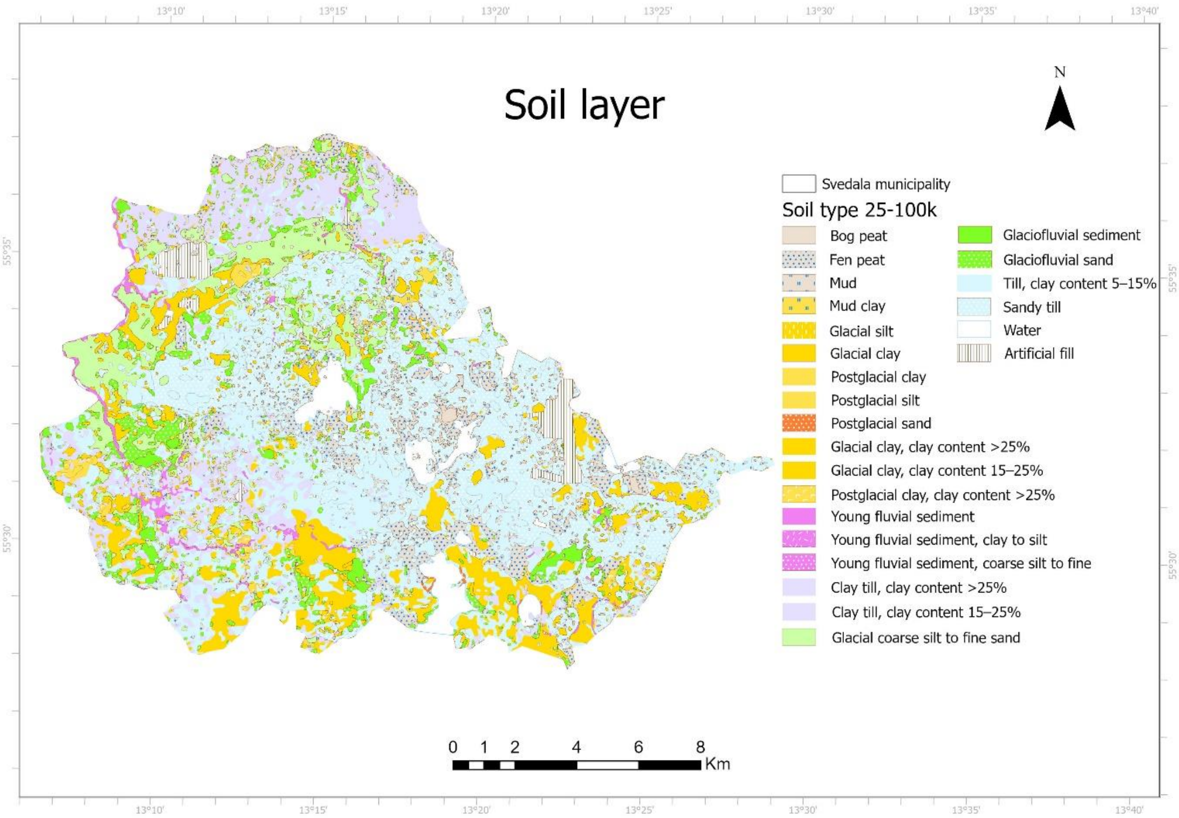

Figure 11: The soil map shows the extend of the soil types in and around the surface (0.5m depth) including any instance of larger rocks at ground level. The layer is shown in its original vector format.

3.3.3 **Soil moisture layers (**SLU soil moisture layer [SLU markfuktighetskarta] and the SLU classified soil moisture layer [SLUMFKKlassad]**)**

The data folder provided for soil moisture contained two soil moisture layers that described the registered moisture for the ground surface (Institutionen för skoglig ekologi och skötsel (SLU), 2020). The layers were provided in raster format and produced based on a combination of data from the Land Surveys national laser scanning, including sample squares from The Swedish National Forest Inventory (Swedish NFI). The spatial resolution of the two raster layers were 2 metres and the values for each raster cell described the mean soil moisture during the year the data was captured. Both raster layers were based on hydrological modelling which in turn were based on the GSD-Elevation data, Grid 2+. From the national GSD-Elevation data, the water flow could be modelled with the assumption that the water would follow the topography and flow downwards. Through the modelling process of flow accumulation, an estimation of the potential amount of water that might gather in one square (pixel) was established. The value for each cell provided in the raster layer indicated how many ha of land would flow into that specific cell. The SLU soil moisture layer (Figure 11A) [SLU markfuktighetskarta] indicated the probability that a pixel in the layer would be classified as wet, where 0 indicates completely dry soil and 100 indicates full saturation. Furthermore, the SLU soil moisture layer [SLU Markfuktighetskarta] depict different soil moisture with the set colouring, as dry land are assigned the colour red, moist land the colour yellow or turquoise, healthy-moist land the colour green and wet land the colour blue. To present a softer transitioning, the colouring has been “blended” together.

The classified SLU soil moisture layer (Figure 11B) [SLU markfuktighetskarta (MFKKlassad)] was instead divided into three classes, dry-health (1), health-moist (2) and moist-wet (3). Open water and water bodies were given the value of 4 (Institutionen för skoglig ekologi och skötsel (SLU), 2020).

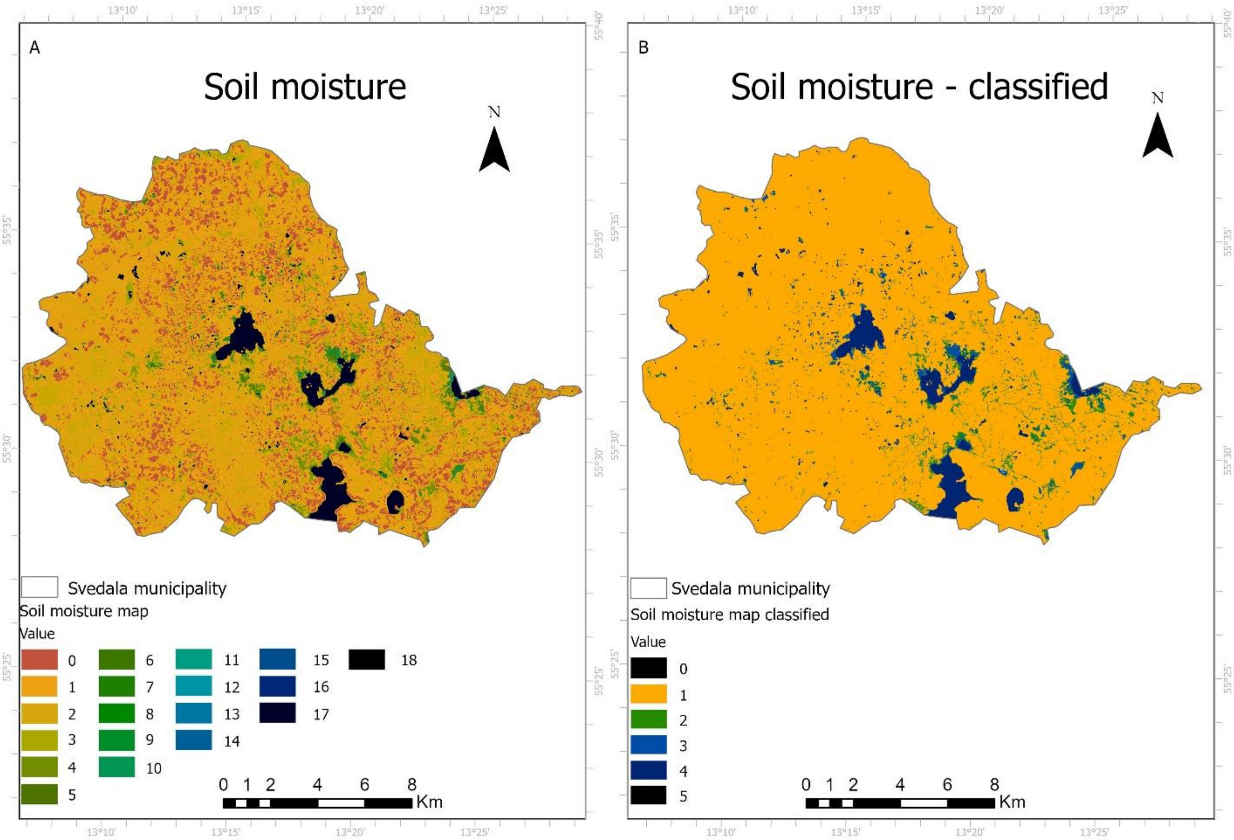

Figure 12: The two SLU soil moisture layers at a 10m spatial resolution, where ***A*** *has been developed to display continuous values (to create smother transitions between the classes) and* ***B*** *categorial values. For map A, 0 indicates dry which gradually increase in saturation with 18 indicating the highest levels of saturation in the area. For map B, 0 and 1 are dry-health which shifts towards moist-wet (3) and ends with 4 and 5 as open water or water bodies.*

3.3.4 The CORINE Land cover layer (CLC 2018)

The producers of the Corine Land Cover 2018 (CLC 2018) layer employed the technique of satellite imagery data interpretation, with the vegetation and land types presented in 44 different classes (Lantmäteriet, 2019b). Of the 44 vegetation classes, 35 were deemed relevant for Sweden and organised in three levels with five main classes listed as Artificial Areas, Agricultural Areas, Forest and Semi-natural Areas, Wetlands and lastly Water Bodies. The mapping was delivered as a geographic cut-out of 50 by 50 km map sheets for the county (Scania). A standardised form of delivery, regardless of the county selected by the user. The height projection of the data was provided in RH2000. The images used for the CLC 2028 vector layer were taken from Sentinel 2, registered during the year 2017, and compared with the images used to produce the previous version, CLC 2012. The satellite images, that the land cover is based on, were of a spatial resolution of 10 metres, except for some images with a 20 or 25 metre resolution. The CLC 2018 vector layer (Figure 12A) presents the land cover for the given reference years, including any changes between the year 2012 and 2018. As such, no intended updates are planned for the CLC 2018 vector layer (Lantmäteriet, 2019b).

3.3.5 National Land Cover Data 2018 (NMD 2018)

The mapping for the National Land Cover Data 2018 (NMD 2018) was carried out during the years 2017-2019 with the plan for an update of the land cover every fifth year (Naturvårdsverket, 2020). The NMD 2018 raster layer (Figure 12B) based its mapping on 25 thematic classes subdivided into three hierarchical levels. The raster was provided in a 10-metre spatial resolution. The mapping was based on the combination of satellite imagery from the Sentinel 2 and laser scanning. On occasion, this would be complemented with information from existing map material (Naturvårdsverket, 2020). The layer used in the thesis was the base layer. The classification of each land cover class category was individually assessed and given one of the following quality grades, very good, good, acceptable, and low (Naturvårdsverket, 2020). For further information of both land class cover classifications please read the respective product description.

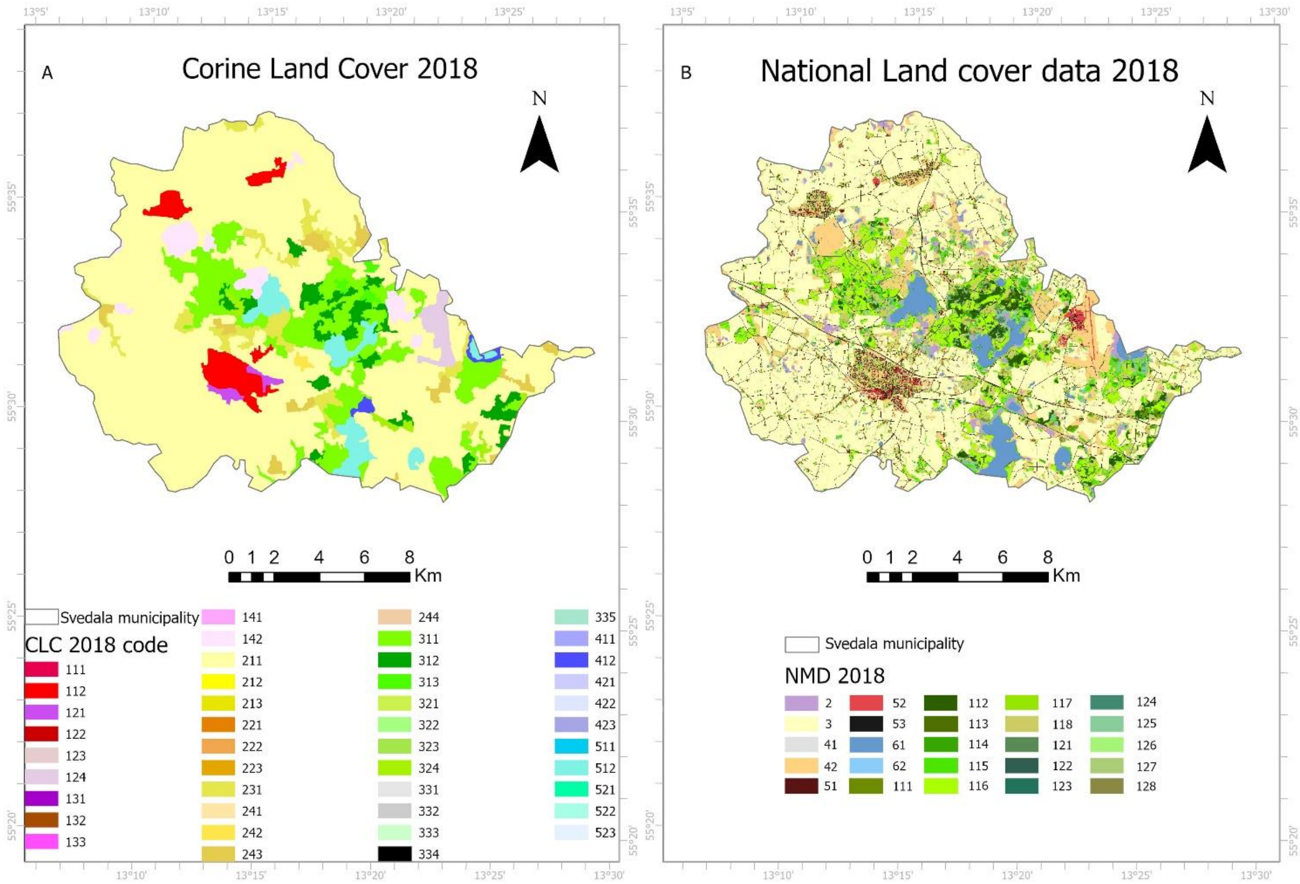

*Figure 13: The current CORINE land cover layer from 2018 (the layer is shown in its original vector format) (**A**) with its classifications ranging from red and purple colours for built-up areas, yellow and brown colours for agricultural lands, green colours for forests and blue colours for water bodies. The National land cover data from 2018 at a 10m spatial resolution (**B**), which is also the most current national land cover, displays dark red to black colours as built-up areas, yellow to orange as open ground and agricultural land, purple and blue colours as wetland and water bodies, and green colours as forest.*

All layers were provided in the national projection of SWEREF 99 TM. Additionally, the projection was kept and applied to all layers used in the thesis as this would facilitate the possibility of modelling the entirety of Scania County. The maps for Scania County environmental factors can be viewed in Figure A1 to A5 of the appendices. It would also make the methodology of the thesis applicable to any given area of Sweden.

3.3.6 Precipitation and Temperature

Mean monthly and daily air temperature was acquired from the Swedish Meteorological and Hydrological Institute (SMHI) weather station at Sturup airport (Station number 53300), located at an elevation of 72 metres above sea level at latitude 55.5231 and longitude 13.3787. For air temperature measurements this was the closest weather station to the area of Bokskogen and Holmeja skogen. Measurements were recorded 2 metres above the ground. The daily mean temperature was registered at midnight the following day. The monthly mean temperature was recorded on the first day of the following month. The values of the measurements have been roughly verified by the institute and the unit was Celsius (°C). The station has been active since the 1st of January 2008 (Swedish Meteorological and Hydrological Institute, 2023a).

Precipitation data was not available at Sturup Airport and thus was obtained from the Swedish Meteorological and Hydrological Institute (SMHI) weather station located in the city of Malmö (Station number 52350). The measurements were taken at 06:00 in the morning the following day. Measurements were recorded 2 metres above the ground. The values of the measurements have been verified and confirmed by the institute and the unit was millimetres (mm). The station is located at an elevation of 20 metres above sea level at latitude 55.5715 and longitude 13.0708. The station has been active since the 1st of November 1995 (Swedish Meteorological and Hydrological Institute, 2023b).

3.3.7 Sample data

Within the Municipality of Svedala, as described in the subsection Study area, lies two forested areas, Bokskogen (in association with the Torup castle) and Holmeja forest. The two forested areas are the locations where the points for the model were collected during the 2023 wild mushroom season (from August to October).

The locations were recorded with a mobile phone, model ASUS Zenphone 10, and the Google Maps app (version 11.123.0103. The sampling method was road sampling along paths and gravel roads within the forested areas. The use of a mobile phone and the Google maps app was justified by that fact that the needed accuracy for the recorded sample points location would equal that of a commercial GPS, given that the location the points were recorded in was a forest with tree coverage (Methakullachat & Witchayangkoon, 2019; Wing et al., 2005; Zandbergen & Barbeau, 2011) A mobile phone with a map app or any other app with coordinate recording capabilities would also be the most accessible equipment most people and mycophile would readily have access to. The coordinates were in the projection of WGS84.

While the Winter Chanterelle and Yellow Foot mushrooms typically grow in coniferous forests, the sample points collected for this thesis revealed several instances where these species were found within beech forest vegetation. Notably, some of these occurrences were in close proximity to water sources.

3.3.8 External sample data from the Global Biodiversity Information Facility (GBIF) and Inaturalist

Sample data with point locations for *Cantharellales* was downloaded from GBIF and Inaturalist. GBIF is described as an international network and data infrastructure that aims to provide anyone, anywhere, with open access to data about life on earth (GBIF, n.d.). Equally, Inaturalist is a platform that provides valuable open data to the likes of researchers, land managers and the public (INaturalist, 2022). These had been collected over a time period of 20 years and shared on the respective portal. The points, part of a shared database, were all recorded during their respective mushroom season. The points were all recorded in the WGS84 projection (GBIF, 2000-2023; INaturalist, 2000-2023).

4 Methods

When aiming to accomplish any kind of modelling, in the context of SDM, the first step should be to establish which data input would be deemed as necessary to obtain reliable output data. Dreisbach et al. (2002) noted that most studies that focusses on fungi habitats are conducted at a resolution of less than 1 m but provided no reference to this statement. They also noted that vegetation, topography, and soils and in some terms, climate factors, are the categories that are most often associated with fungi in terms of ecological factors. Vegetation or land cover would be considered as the primary factor. Topographic factors would include elements such as elevation, slope, and aspect. For earth deposits, data akin to soil properties and soil moisture would be useful. The latter two were considered secondary factors. Climate factors are complex. Here temperature and moisture (being the result of precipitation) would according to Dreisbach et al. (2002) be considered the most relevant. All in all, these factors may be important to *Cantharellales* occurrence.

Due to the complexity of climate factors and their effects on the timing of mushroom formation, including the subsequent effects on the output as indicated by Dreisbach et al. (2002), a decision was made to exclude said data from the modelling. Instead, they were included as standalone graphs to discuss its possible influence on the sample data for the 2023 *Cantharellales* season. Thus, the layers described in the subsection 3.3.5 of the datasets section 3.5, were downloaded (daily precipitation data from SMHI from station Malmö A and mean monthly temperature data from SMHI from station Malmö-Sturup airport).

. ArcGIS Pro version 2.7.3 was used for all pre and post GIS processing. The processing of CSV files and production of graphs was conducted in Microsoft Excel version 2403 (Build 17425.20176) and WordPad version 23H2 (Build 22631.3447). The Maximum Entropy Modelling of Species Geographic Distributions version 3.4.3 November 2020 (Maxent) was used to produce all modelling outputs.

4.1 The process of developing a model

The DEM raster layer was created by combining the sections of GSD elevation data 2+ segments into one raster layer and harmonised to the extent of the boundaries of the county Scania. An additional layer for the Municipality of Svedala was also created. Both layers were resampled to the spatial resolution of 10 metres to harmonise with the land cover layers. Afterwards, aspect and slope layers where generated based on the DEM layer.

Regarding the soil layer, the Jordarter_25_100k_jg2_south layer was dissolved into three categories: JG2, JG2_TX and symbol to create a layer where each soil type was represented as one row in the attribute table. This layer was subsequently clipped with the borders for Scania and Svedala respectively and converted into a raster layer with the spatial resolution of 10 metres.

Both soil moisture layers, SLU soil moisture [Markfuktighetskarta] and SLU classified soil moisture [Markfuktighetskarta (MFKKlassad)], came in tiled sections. For each layer, the tiles were merged into one mosaiced raster, and assigned a spatial resolution of 10 metres to match the resolution of the land cover layers.

For the CORINE land cover layer, a selection was made so as the previously geographical delineations were covered and afterwards harmonised. It also had to be converted from vector format into a raster format and reclassified for the correct values to be included in the final conversion into the ascii file format.

The land cover layer from The Swedish Environmental Protection Agency came in the format of a raster and pre-delimited to the county of Scania. Afterwards it was harmonised to the limits of the Svedala Municipality.

The collected sample point coordinates were imported from an excel file and reprojected from WGS84 to SWEREF99 TM. As the coordinates were recorded in decimal degrees (dd) they had to be converted into metres, as SWEREF99 TM is a projected coordinate system. This was accomplished by calculating the geometry in the attribute table. This was necessary for the points to match the extent of the raster layers. The revised attribute table was exported into two separate csv-files where the five individual *Cantharellales* species names were replaced with only that of *Cantharellales*, as this was how the data was intended to be modelled. This was because collecting a larger number of each of the five *Cantharellales* species points could not be achieved given the time constraints and limited forested area easily accessed in proximity to the city of Malmö (home location). Additionally, whichever of the *Cantharellales* mushroom species one would find at the modelled predicted location, would be considered an equally good find. One file contained points, to be used for training the model and the other file containing validation points, was used to assess the accuracy of the model output. To facilitate the option of applying the model to an area outside the limits of the Municipality of Svedala, sample points from GBIF and Inaturalist were downloaded and harmonised in accordance with the description above.

To evaluate the model for any predisposition to sample bias, a buffer layer was created around the sample points. The buffer was in the form of a polygon layer which had to be converted into a raster and harmonised in accordance with the rest of the input factors. To evaluate which distance should be used, three different options were tested for the extent of the Svedala Municipality, a 50m-, a 250m- and a 500m buffer. After a visual evaluation in conjunction with the values of the AUC, the third option of a 500m buffer was selected.

Prior to all layers being converted into ASCII files, the decision was made to set the extent of the DEM layer. The reason for this was that Maxent requires the extent of every single layer to be identical. Otherwise, the program would not be able to run. The same thing was required by the sample data files. The ASCII file format is the only raster format accepted by the Maxent program (.asc file ending format in ESRI).

Each layer was combined into the following setup, as shown in Table 1, to evaluate how different land cover- and soil moisture layers would impact the generated model in Maxent.

Table 1: The setup for each environmental factor layer combination of the respective modelled areas. While the Land cover and Soil moisture layer would vary, the Soil-, Elevation-, Slope and Aspect layer would remain the same.

Modelled area		Environmental factors					
		Land cover layer	Soil moisture layer	Soil layer	Elevation layer	Slope layer	Aspect layer
	Scania	CLC2018	SLU Soil moisture	soil type	DEM	Slope	Aspect
	Svedala	CLC2019	SLU Soil moisture	soil type	DEM	Slope	Aspect
	Scania	CLC2020	SLU classified Soil moisture	soil type	DEM	Slope	Aspect
	Svedala	CLC2021	SLU classified Soil moisture	soil type	DEM	Slope	Aspect
	Scania	NMD2018	SLU Soil moisture	soil type	DEM	Slope	Aspect
	Svedala	NMD2019	SLU Soil moisture	soil type	DEM	Slope	Aspect
	Scania	NMD2020	SLU classified Soil moisture	soil type	DEM	Slope	Aspect
	Svedala	NMD2021	SLU classified Soil moisture	soil type	DEM	Slope	Aspect

To organise the output data generated by Maxent, separate folders prior to running each individual setup was created. The reason for this being that Maxent required a preselected folder and that no option for naming output layers individually existed. Thus, the Maxent input and output folders for the runs were as shown in Table 2 and 3.

Table 2: The setup of the first round of runs, with the individual random test percentage and statistical output setting.

Round 1:	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6
Random test percentage	25	33	31	25	33	31
Output format	logistic	logistic	logistic	cloglog	cloglog	cloglog

Table 3: The setup of the second round of runs (where the bias layer and jackknife settings were included), with the individual random test percentage and statistical output setting.

Round 2*:	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6
Random test percentage	25	33	31	25	33	31
Output format	logistic	logistic	logistic	cloglog	cloglog	cloglog

For the first round, one file containing all the sample points were entered and Maxent was allowed to pick the points to be used for training and testing respectively. This option was entered in the settings menu in the random test percentage parameter.

For the second round, a bias layer was added in the advanced settings option and the sample points were divided into two separate files. One file containing points to be used for training the model and another file with sample points aimed for testing the model. The layers aspect, slope, DEM, and SLU soil moisture was set as continuous and land cover, soil type and SLU soil moisture - classified was set as categorical. The overall methodology as described above was compiled into an illustrative description shown in Figure 13.

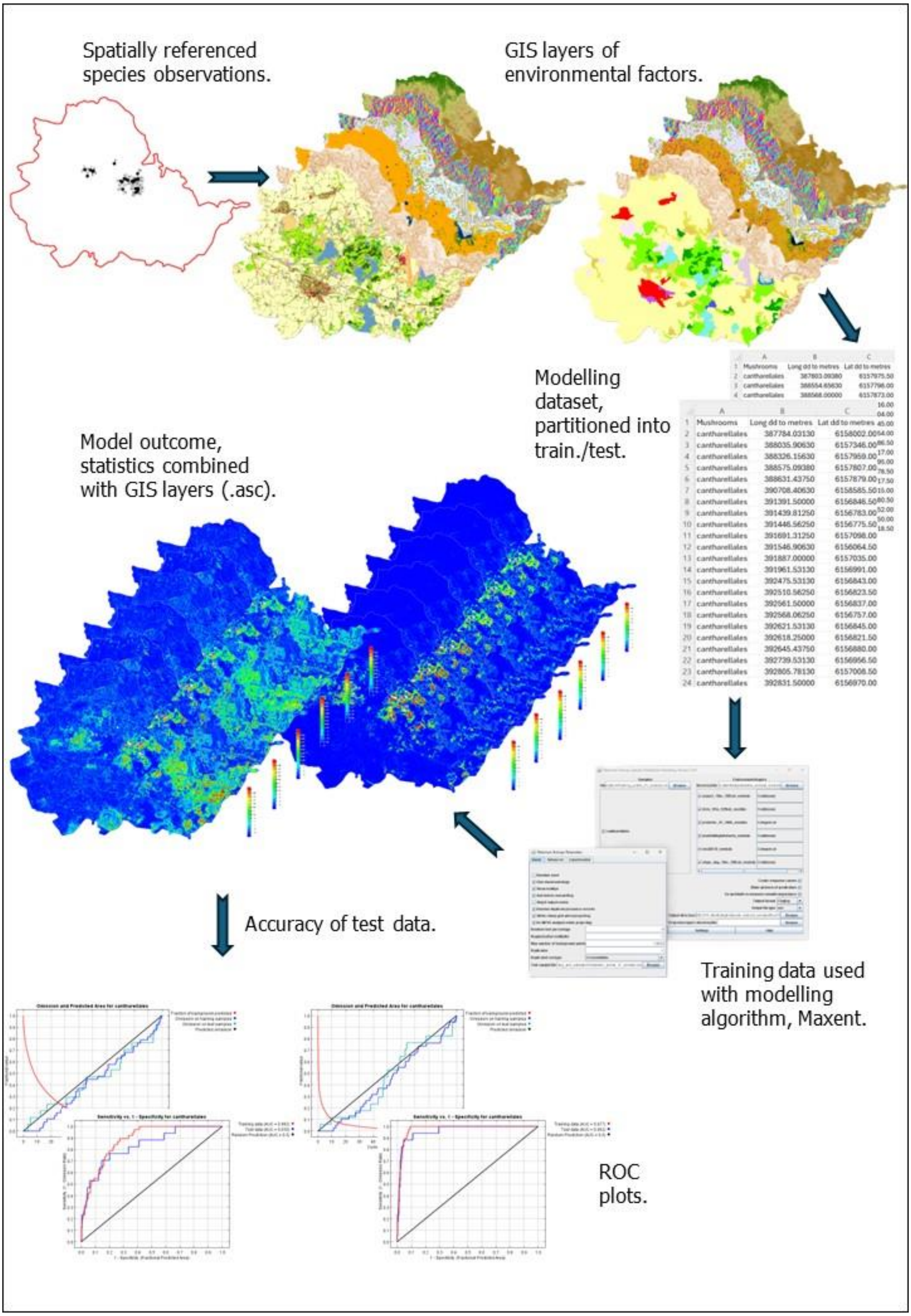

Figure 14: A graphic illustration of the main steps that were undertaken during the process of developing a methodology, to produce a model for Cantharellales *probability occurrences. From sample points (divided into training and test files) and environmental factor layers entered in the Maxent model, to output GIS layers and ROC plots.*

5 Results

In the following section and subsections, the results from the modelling algorithm, Maxent, run on the environmental factors covering the Municipality are presented. They will be presented through tables containing AUC values, derived from ROC plots, that indicate the potential fit of the model. To further the understanding of which environmental factor contributed the most to the training of the model, tables have been produced, including information on how the individual environmental factors permutation importance in the final Maxent model was ranked. Finally, illustrative maps, indicating through a colour scale going from dark blue to deep red, with accompanying values from 0 to 1, depicts the species prediction probability.

5.1 General results

This section helps to provide a summation of the overall results for each of the two areas that were modelled. All the results for each individual runs and corresponding environmental factor layer setup, of the Svedala Municipality and Scania County, can be viewed in the Figures A6 to A18, and Table A2 to A21 (Svedala Municipality) and Figures A19 to A37, and Table A22 to A41 of the appendices.

5.1.1 Svedala Municipality

All AUC values were evaluated and found to be above 0.5 indicating that the model prediction of species probability was better than chance. The first round produced results for the Svedala Municipality model that only varied slightly, with values from 0.974 to 0.978 for the training AUC, resulting in a mean AUC of 0.976. The test AUC values did not reach the same levels but were also above 0.900, with values ranging from 0.924 to 0.965 and, producing a mean AUC of 0.945.

For the second round, the inclusion of the bias layer and the preselection of training- and testing sample points caused the derived training- and test AUC to be lowered, and instead range from 0.827 to 0.904 and 0.683 to 0.879, respectively. Consequently, the mean training- and test AUC was calculated as 0.866 and 0.781.

Analysing the contribution of each environmental factor for building the model, the NMD 2018 land cover always had the highest ranking among the 6 contributing factors, during all runs of round 2. The CLC 2018 land cover was only ranked first for its individual environmental factor combination runs of round 1. For the second round the CLC 2018 was instead ranked as the second or third most important environmental factor. Alternately, the DEM layer was ranked highest. When interpreting the ranked permutation importance, which indicates the importance of the environmental factor within the established model, a somewhat similar picture emerges. Here however both the NMD 2018 and CLC 2018 land cover both ranked highest for all the respective environmental factor combination runs of round 1. Whereas for the second run, both land covers varied in rank from first to fourth. This type of variability between the environmental factors stayed consistent for all the factors regarding the permutation importance.

5.1.2 Scania County

Once more, all AUC values were evaluated and found to be above 0.5 but were overall lower than the AUC values for the Svedala Municipality model. In round 1 the training- and test AUC covered the values from 0.842 to 0.853 and 0.794 to 0.811, resulting in mean AUC values of 0.848 and 0.803, respectively. During the second round of runs the values similarly as the Svedala AUC values were lowered compared to the first round of runs. Here the values ranged from 0.746 to 0.809 for training AUC and 0.623 to 0.753 for test AUC.

A review of each of the environmental factors was also completed for the Scania County model. Compared to the Svedala Municipality model, the CLC 2018 land cover contribution for building the model only ranked second (for its individual environmental factor combinations), whereas the NMD 2018 land cover contribution once again ranked number one in round 1, for its environmental factor combinations. For the combinations where the CLC 2018 land cover was an environmental factor, the soil type layer was the environmental factor that ranked highest. The third highest environmental factor varied between the DEM-, the soil moisture- and the slope layer. Likewise, where the NMD 2018 layer was included in the environmental factor combinations, it was the highest ranked environmental factor. Then the second highest layer would be the soil type layer and third ranking factor would have the same variations as the CLC 2018 land cover combination. The second round produced a more heterogeneous pattern, and only where the NMD 2018 layer was part of the six factors used in the model, was it ranked first. The soil type would rank second and the third highest ranked factor would vary as described for the CLC 2018 combinations. No distinct environmental factor could be singled out specifically, when comparing the factors regarding the permutation importance. Only for the runs 1 to 6 in round 1 with NMD 2018 land cover, SLU Soil moisture layer, soil type, DEM, Slope and Aspect, did the land cover rank highest.

5.2 Selection of best fit model for the Svedala Municipality

When evaluating which factor combination produced the best model with the recommended cloglog setting by Phillips et al. (2017) two runs, 11 and 12, with the environmental factor combinations that included the NMD 2018 land cover had the best results of 0.882, 0.870, 0.903, and 0.904 (Table 4). As observed, run 12 with NMD 2018 and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) had a higher test- than training AUC. It also had the highest test AUC of 0.879 of all four of the best fit models. Consequently, the mean AUC was calculated and the standard deviation for the AUC values were retrieved from the statistical output produced by Maxent (Table 4). Therefore, with the addition of the standard deviation combined with the mean AUC, the Maxent model with the best fit was concluded to be round 2, run 12 of the environmental factor combinations of the NMD 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer). Run 12 also had the advocated division of 69 to 31 percentages between training- and test samples, according to the advised equation (1).

Table 4: The training and test data AUC values for the two runs, 11 and 12 of round 2, for the respective land cover and SLU soil moisture layers (Including DEM, Slope, Aspect, and soil layer). In addition, their respective mean AUC values, and corresponding values for standard deviation. The colouring of the table is intended to match that of the maps in Figure 14, with the colour red indicating a good or excellent result.*

	Svedala			
	NMD 2018 and SLU soil moisture layer*		NMD 2018 and SLU classified soil moisture layer*	
Round 2	11	12	11	12
Training data (AUC)	0.903	0.882	0.904	0.870
Test data (AUC)	0.786	0.835	0.783	0.879
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5

	Svedala			
	NMD 2018 and SLU soil moisture layer*		NMD 2018 and SLU classified soil moisture layer*	
Round 2	11	12	11	12
Mean AUC	0.845	0.859	0.844	0.875
AUC Std. Dev.	0.0631	0.0496	0.0636	0.0336

5.2.1 Analysis of best fit Maxent model for Svedala Municipality

The percentage contribution and percentage permutation importance for the best fit Maxent model, were of the same range, except for the two factor layers with the fourth- and the sixth lowest percentages. The Svedala Municipality model (Table 5) similarly had the same order for the three environmental factors with the highest percentages that contributed to the model as well as within the established model. The NMD 2018 land cover contributed almost half of the data, with 49.9 and 42.5 percentages for each element pertaining to the model. Interestingly, as stated above, the two factor layers with the fourth- and the sixth lowest contribution, Slope and SLU soil moisture layer, switched places. Slope changed from 9 % contributing to the build of the model to 1.7 % within the model. The SLU soil moisture factor similarly switched from contributing 1 % to the build of the model to 6.6 % within the model.

Table 5: The percentage of each environmental factor that contributes to building the model as well as the percentage of permutation importance of each environmental factor in the established model. Specifically for the selected best fit model with the NMD 2018 land cover and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer). The colour red indicating a high % contribution or % permutation importance.*

	Svedala	
	NMD 2018 and SLU classified soil moisture layer*	
	% Contribution	% Permutation importance
NMD 2018	49.9	42.5
DEM	23.9	27.9
Slope	9	1.7
Aspect	2.3	5.2
Soil type	14	16.2
SLU soil moisture layer	1	6.6

Additionally, the environmental factors were ranked in accordance with the percentages each factor contributed in either building the model or how each were ranked while being used inside the confines of the model (Table 6). Subsequently, the NMD 2018 factor layer ranked highest, the DEM factor layer second and the soil type factor layer third. The results produced by the jackknife setting also confirmed that the NMD 2018 land cover-, DEM- and the soil type layer in said order contributed the most information by themselves that was not present in the other factor variables.

Table 6: How each environmental factor layer was ranked based on their respective percentage values for the NMD 2018 land cover and SLU classified soil moisture layer model (Including DEM, Slope, Aspect, and soil layer). The colour red assigned to the highest ranked environmental factor.*

	Svedala	
	NMD 2018 and SLU classified soil moisture layer*	
	Ranked Contribution	Ranked Permutation importance
NMD 2018	1	1
DEM	2	2
Slope	4	6
Aspect	5	5
Soil type	3	3
SLU soil moisture layer	6	4

As shown in the map (Figure 14) the areas with the highest probability of species occurrences of the *Cantharellales*, primarily correlated with areas of land cover that would be associated with the specific order of fungi modelled in this thesis. Aside from areas in near proximity of the recorded sample points, new and other potential sites were shown in both Holmeja forest and Bokskogen. As the land cover of forested area in and around Holmeja forest could objectively be considered larger, the number of possible species occurrences would consequently be larger. In general however, a large part of the Svedala Municipality had probability (species occurrence) values below or around 0.5, which would indicate the possibility of finding *Cantharellales* would equal that of chance. Two land cover classifications were predominant in locations with a high species occurrence probability, spruce forest (NMD 2018 class nr 112) and deciduous forest (NMD 2018 class nr 116). Elevation varied between 60 to 75 m, again at locations with a high species occurrence probability of the second highest environmental factor layer, being the DEM. The third highest ranked environmental factor layer, soil type, showed the highest probability of species occurrences in sandy moraine (Soil type class nr 95).

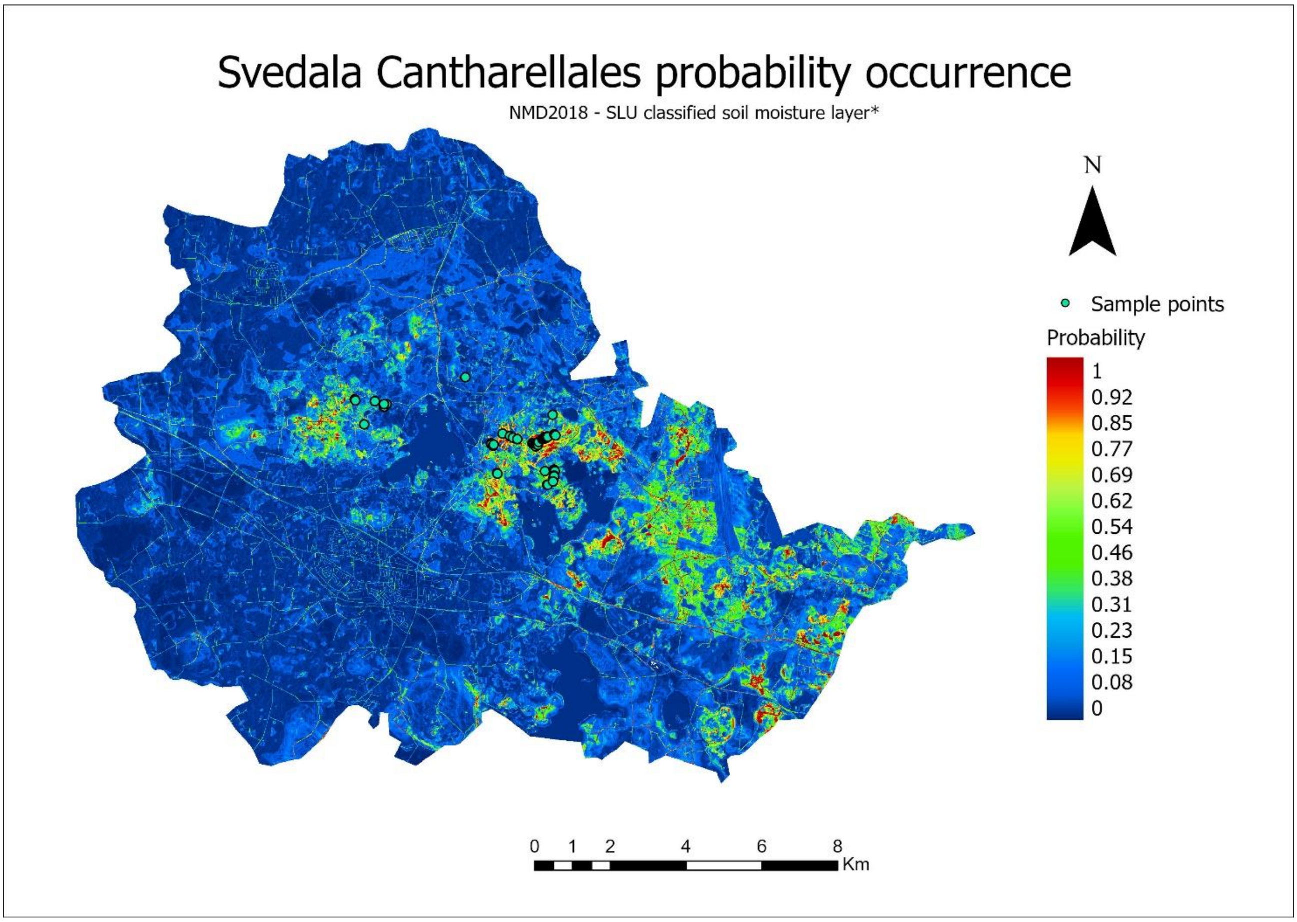

Figure 15: The map shows, with a colour scale ranging from blue to dark red, the species probability occurrence of Cantharellales in any selected area for the Svedala Municipality. This is graphic representation for the environmental factor layer combination of the land cover layer NMD 2018 and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) with the spatial resolution of 10m.*

5.3 Selection of best fit model for Scania County

Similarly, with Scania County, four runs in the second round with the recommended cloglog setting showed the highest AUC values 0.808, 0.771, 0.809, and 0.764 (Table 7). However, compared to the values for the Svedala Municipality, no test AUC values were higher than the training AUC values. Nevertheless, what was comparable between the Svedala Municipality and the Scania County was that the NMD 2018 land cover layer yet again produced the overall highest AUC values for all of the runs it was included in. Still, for the larger area that the Scania County represented, the NMD 2018 and SLU soil moisture layer combination had the altogehter best mean AUC values with 0.808 and 0.771. To separate the two runs, the mean AUC values were once more calculated and combined with the standard deviation (Table 7). Consequently, the Maxent model with the best fit was found to be 12th run in round 2 of environmental factor combinations of the NMD 2018 land cover-and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer). Again, as with the Svedala Municipality model, run 12 applied the advocated division of 69 to 31 percentages between training- and test samples, according to the advised equation (1).

Table 7: The training and test data AUC values for the two runs, 11 and 12 of round 2, for the respective land cover and SLU soil moisture layers (Including DEM, Slope, Aspect, and soil layer). In addition, their respective mean AUC values, and corresponding values for standard deviation. The colouring of the table is intended to match that of the maps in Figure 15, with the colour red indicating a good or excellent result.*

	Scania			
	NMD 2018 and SLU soil moisture layer*		NMD 2018 and SLU classified soil moisture layer*	
Round 2	11	12	11	12
Training data (AUC)	0.808	0.771	0.809	0.764
Test data (AUC)	0.676	0.753	0.662	0.737
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5

	Scania			
	NMD 2018 and SLU soil moisture layer*		NMD 2018 and SLU classified soil moisture layer*	
Round 2	11	12	11	12
Mean AUC	0.742	0.762	0.736	0.751
AUC Std. Dev.	0.0524	0.0441	0.0517	0.0466

5.3.1 Analysis of best fit Maxent model for Scania County

Once more the NMD 2018 land cover layer had the highest percentage contribution and percentage permutation importance for the best fit Maxent model. From there on out though the models for the Svedala Municipality and Scania County differed completely. Although, they do share the same deviation of having four out of six factors inhabit the same position of percentage contribution and percentage permutation importance (Table 8). The NMD 2018 land cover contributed roughly a third of the data to the model, with 36 and 27.7 percentages. The second highest factor in both percentage contribution and percentage permutation importance was the soil type factor with 32 and 25 % respectively. Noteworthy for the Scania Maxent model, was that it was the DEM factor- and the SLU soil moisture layer that switched places in the amount of percentages that they each contributed. Also, compared to the Svedala Municipality area, the slope- and aspect factor layers clearly had the lowest percentage contributions for the Scania County. For the Svedala county this differed instead between three different layers (but still included the same two factors), slope, aspect and SLU soil moisture layer.

Table 8: The percentage of each environmental factor that contributes to building the model as well as the percentage of permutation importance of each environmental factor in the established model. Specifically for the selected best fit model with the NMD 2018 land cover and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer). The colour red indicating a high % contribution or % permutation importance.*

	Scania	
	NMD 2018 and SLU soil moisture layer*	
	% Contribution	% Permutation importance
NMD 2018	36	27.7
DEM	12.5	23.5
Slope	0.6	3.9
Aspect	0.2	0.9
Soil type	32	25
SLU soil moisture layer	18.8	19

Thus, the environmental factors for the Scania County model likewise as the Svedala Municipality model were ranked in accordance with the percentages each factor contributed in either building the model or how each were ranked while being used inside the confines of the model (Table 9). Therefore, the NMD 2018- and soil type factor layer inhabited the first and second highest ranking for both contribution and permutation importance, individually. The jackknife setting again confirmed that the NMD 2018 land cover was the layer that contributed the most information by itself that could not be found in the other factor variables.

Table 9: How each environmental factor layer was ranked based on their respective percentage values for the NMD 2018 land cover and SLU soil moisture layer model (Including DEM, Slope, Aspect, and soil layer). The colour red assigned to the highest ranked environmental factor.*

	Scania	
	NMD 2018 and SLU soil moisture layer*	
	Ranked Contribution	Ranked Permutation importance
NMD 2018	1	1
DEM	4	3
Slope	5	5
Aspect	6	6
Soil type	2	2
SLU soil moisture layer	3	4

As seen in the map (Figure 15) that illustrated the Scania Maxent output, the highest values of species occurrence probability could still be found in the northwestern parts of the Scania County as well as in the Svedala Municipality area. What clearly sets the two models apart, however, were the consequence of having the soil type factor layer ranked second highest. Subsequently, the soil type classification that was predominant in locations with high species occurrence probability was glacial clay, with a clay content higher than 25% (Soil type class 43). Similarly, as with the Svedala Municipality model, the majority of the Scania County had probability (species occurrence) values around 0.5. Notably, because of the AUC values both for training, testing, and the mean averaged around 0.750 in the Scania County model, the individual class for each factor in large does not correspond to the class of the Svedala Municipality model (averaged AUC around 0.850).

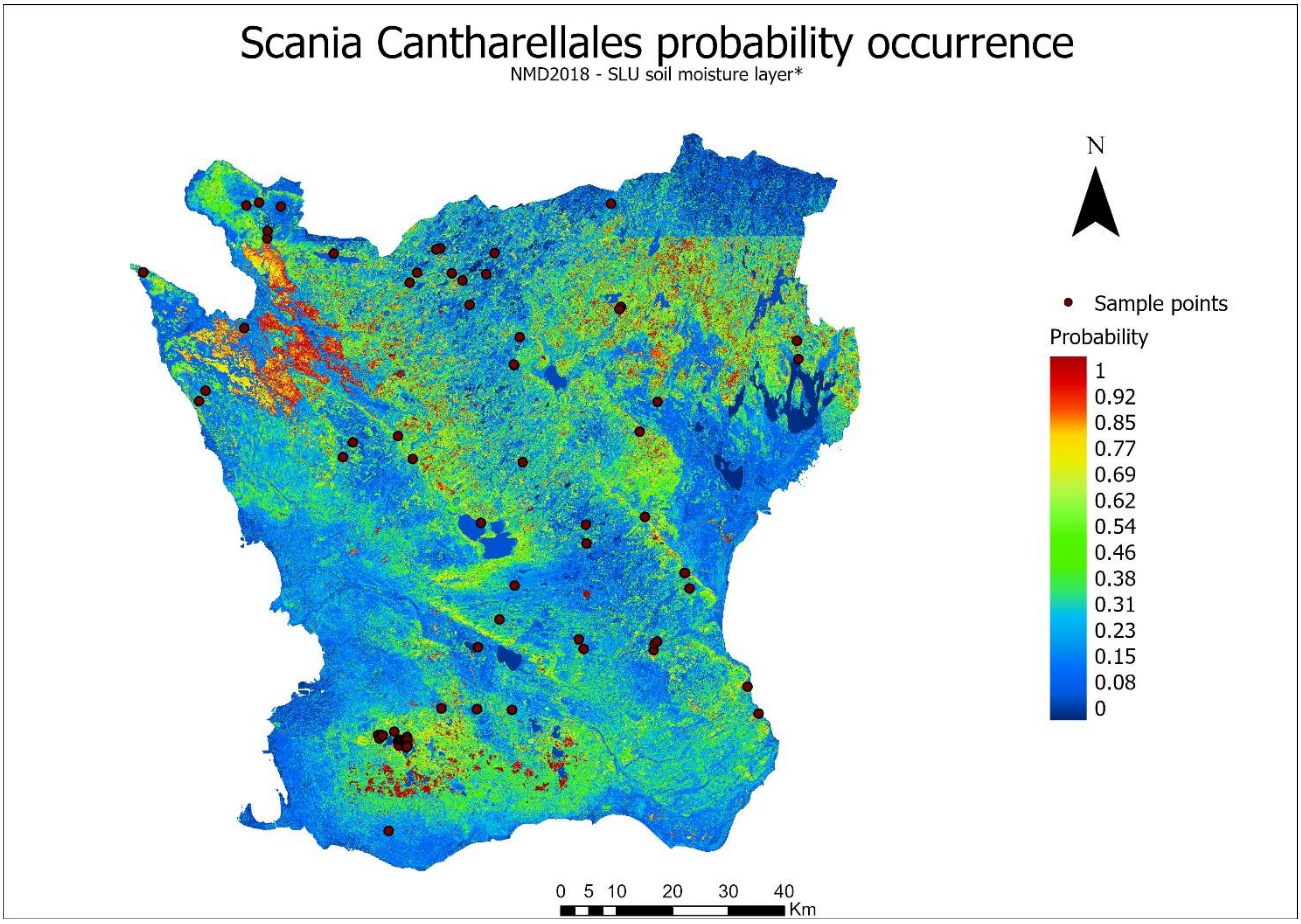

Figure 16: The map shows, with a colour scale ranging from blue to dark red, the species probability occurrence of Cantharellales in any selected area for the Scania County. This is graphic representation for the environmental factor layer combination of the land cover layer NMD 2018 and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) with spatial resolution of 10m.*

5.4 Climate data

Since all climate factors were deliberately not included within any of the Maxent models per the recommendation of the literature, values for precipitation and temperature were illustrated in the following graphs. The intention was to see if they might help indicate how the 2023 *Cantharellales* season had developed.

Precipitation and temperature for the area around Svedala Municipality, recorded a mean value of 3.1 mm (precipitation) as seen in Figure 16 and 16.4°C (temperature) in Figure 17, between the first of July and the first of October. Approximately one week after the highest recording of precipitation (55.5 mm) for the entire summer season, there was also a peak in temperature (20.2°C).

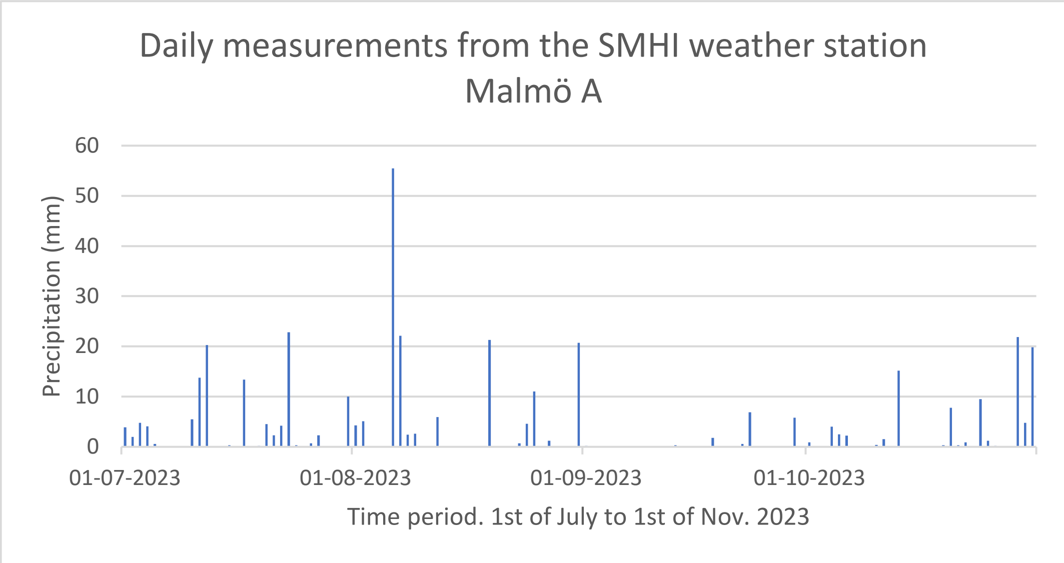

Figure 16: The daily precipitation measurements in mm from the SMHI weather station Malmö A. From the 1st of July to the 1st of Nov 2023.

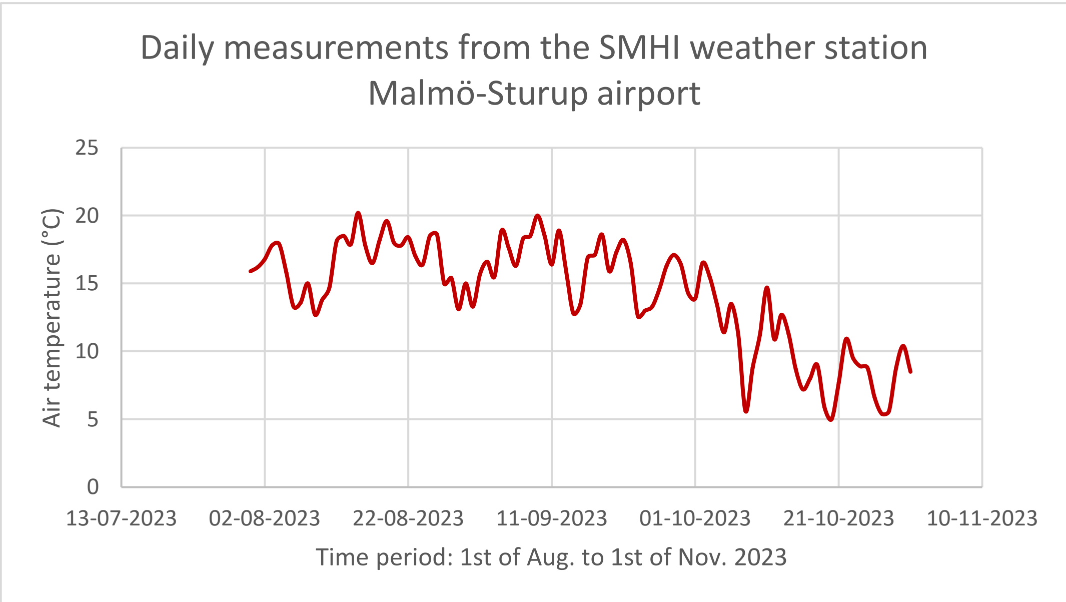

Figure 17: The daily air temperature measurements in mm from the SMHI weather station Malmö-Sturup airport. From the 1st of Aug. to the 1st of Nov. 2023.

6 Discussion

With the aim of this thesis being to map the locations of possible sites where *Cantharellales* mushrooms can be found using SDM and GIS, the following section will expand on and discuss the outcome as previously presented in the Results section. Moreover, some of the benefits and drawbacks encountered during the process of developing the model and improving it will also be described. Furthermore, some of the general limitations regarding all aspects of the thesis, in addition to other relevant features pertaining to the topic, will be discussed.

6.1 Results of the models

Several sources were consulted when it came time to evaluate how good of a model had been achieved with the Maxent program. Specifically, the AUC values. which would be considered the easiest results to assess. Besides the available literature from the main creators of the program, Steven J. Phillips, and Robert E. Schapire, and the html-file generated by the program itself in the output folder, two additional articles were employed. Araújo et al. (2005) and Wang et al. (2022) states that models' predictive accuracies based on independent validations are excellent if AUC is greater than 0.900, good if somewhere around 0.800 to 0.900, but only counts as fair if between 0.700 and 0.800. They also detail that AUC values ranging lower than 0.500, would equal models with no predictive abilities, and models that had AUC values of 1.0 would indicate a model providing perfect prediction probabilities. This scale has also been used in more recent papers by Alemayehu et al. (2024) who added two more grades to the scale established by Araújo et al. (2005). Thus, any AUC values between 0.600 and 0.700 are to be considered poor and any AUC values between 0.500 and 0.600 were invalid. Considering these classifications, the Maxent model determined to have the best fit for Svedala Municipality achieves the assessment of a good model. Both in terms of the AUC value for training (0.870) and testing (0.879), but also when it pertains to the mean AUC value (0.875). The best fit model for the Scania County does not reach the same level. With its mean AUC value of 0.762 (training and test values of 0.771 and 0.753) it can only be considered as a fair model. This also led to areas with what could be viewed as "false positives", that is, areas are indicated to contain a high probability for *Cantharellales* which should not do so.

It is worth noting that there was little to no difference between the results for the AUC values when using either the logistic or cloglog setting within the different environmental factors runs. This can be observed in both the Svedala- and Scania model. The main reason for conducting several different runs, with different combinations of environmental factors was to investigate and develop the model. This, as previously mentioned in section 5.1, can be viewed in more detail in Figures A6 to A18, and Tables A2 to A21 (Svedala Municipality) and Figures A19 to A37, and Tables A22 to A41 of the Appendices. The reason for the numbers being the same, might be due to the factors being used in the models. It might also be attributed to the number of samples for presence observations used in the model. Further runs with either different environmental factors or a number of sample points (presence observations) could share some insight into their respective influences of the respective models. However, in general, both models have proven that they can both produce viable output that works with a limited amount of sample points. Specifically, when taking into consideration the proportion between the number of sample points versus the total area for which they are used.

As one of the objectives of this thesis, the outcome of the two different land covers, the CORINE land cover, and the National land cover from The Swedish Environmental Protection Agency [Naturvårdsverket], confirmed that a national land cover most likely will outperform a more generalised land cover, in this case, a Pan-European land cover.

This should also be taken into consideration when collecting input data for a potential SDM. Therefore, as far as possible the aim should be to make use of environmental factors that are best suited to the area where the model is to be applied. If a model, where the expected environmental factor with the highest impact, only exists in a format primarily intended to be used at a larger scale, this could affect the result and either validate or disqualify an otherwise well-constructed model. It of course all comes down to scale and how current the GIS layers are.

For the Svedala Municipality model, the NMD 2018 land cover layer was clearly the most influential land cover of the two applied in the Svedala Municipality model, which was also confirmed by the jackknife setting.

The main purpose of the jackknife setting, as described in the method section, is to help interpret the output results. Baldwin (2009); Elith et al. (2011); Phillips (2005) all recommend this setting when identifying which variable has important individual effects on the model. It also helped to confirm that the DEM layer in fact had the second highest effect overall. This can be explained by the fact that the sample points are all located in areas of higher elevation compared to their surroundings. It could also explain why the soil type is the third highest and not the second highest factor, a layer, that potentially should have held the second highest position, taking into account that fungi are also highly interconnected to the type of soil it grows in, like that of their symbiosis with certain tree species as explained by Dreisbach et al. (2002), Dyshko et al. (2024) and Pent et al. (2017). Likewise, for Scania County, the NMD 2018 land cover layer was the most influential land cover of the two tested in the model (the NMD 2018 layer and the CLC 2018 layer). Nevertheless, with a much larger area in combination with sample points that did not correlate as excessively with elevation, as most of the sample points for Svedala Municipality did, the soil layer was able to have a larger impact on this model (Scania County).

Another aspect worth additional explanation was the choice of splitting up the sample points into training points and test points (used for validation of the model) instead of relying on the function available in the Maxent model. A training- and testing setup function well for modelling according to Yadav and Shukla (2016). Preferably, a model ought to be validated with an available independent dataset as specified by Sillero et al. (2021). They further explain that the dataset specifically intended for this use, should have a different origin, assembled with a different sampling method, by other people or persons, or from a different period than the 2023 *Cantharellales* season. With the main reason being that an independent set of sample points has a different origin of biases, which in effect makes it able to effectively validate the model (Sillero et al., 2021). The fact is however, that this kind of data is rarely available. Consequently, this is the specific reason for splitting the sample points into the two separate entities. If enough of the external sample points, used for the Scania County model, were to have been located within the Svedala Municipality, they could have filled this function. Sadly however, this was not the case. What could also be observed in the results was that the use of the exact division of sample points per the advised equation (1) also produced the best results for both models, both Svedala Municipality- and Scania County model.

When it comes to the reason for not including climatological data in the model there were a few specific reasons. Very little information is available that can be directly correlated to where and when the fruit bodies of fungi will form (Dreisbach et al., 2002). It is well documented that mushrooms are dependent on temperature as well as the right amount of soil moisture (Mossberg et al., 1987). Mossberg et al. (1987) also explains that the timing of when the fruit bodies will appear can vary a lot from year to year. Different mushroom species also have different requirements in terms of temperature and access to water or moisture according to Mossberg et al. (1987).

Periods of alternately overcast weather that includes rain showers or clear skies combined with higher temperatures clearly benefits the production of fruit bodies as stated by Mossberg et al. (1987). Additionally, Dreisbach et al. (2002) points out that neither precipitation nor temperature have so far been directly linked to the occurrences of fungi fruit bodies. All in all, climate is to be considered a complex factor when included in any SDM.

So, to generate a model applicable for several locations that is not dependent on any specific levels of temperature and precipitation, the climatological factors were not entered into the Maxent model. Instead, as mentioned in the methodology the recorded values for air temperature and precipitation were illustrated in graphs (Figures. 16 and 17).

To put this into context, the first couple of the sample points were recorded in the Svedala study area one week after a recorded peak in temperature which was preceded a week before that by a high amount of precipitation. This trend continued throughout the entire mushroom season and in large parts followed the recordings of the periods of alterations in precipitation and temperature as noted by Mossberg et al. (1987).

In future studies, a follow up on the areas where the probability of species occurrences was developed by the models to see if the locations contained any *Cantharellales*, alternatively applying the models on other types of fungi species to see if they would also produce viable results.

6.2 Pros and cons of the model

As with all things, there are advantages and disadvantages. One advantage of using Maxent, is that it works with presence data only. Since it is very hard to make sure that no fruit bodies of *Cantharellales* in the study were overlooked, the added difficulties of recording absence data, and the effects it will have on the end results, further promotes the use of presence data only. There is also the higher likelihood that parameter estimates are biased and that any output produced by a model will reflect the surveyors' ability to locate a certain species, and not where said species is in a designated environment (MacKenzie, 2009). Thus, it is very hard to verify 100% that no samples have been overlooked. Add to that the above and below ground aspect, when it comes to the width of the mycorrhizae, versus visible fruit bodies. Another detail is the lack of data that describes the locations where absence of a species is unquestionably known as explained by Jiménez-Valverde (2011). As a result, as Jiménez-Valverde (2011) puts it, using presence data is often the only data that modellers have for a given species. Arguably, presence data is the only reliable dataset for modelling fungi. Moreover, given the set time constraints of approximately two and a half months to complete the thesis in combination with the scope of the bachelor thesis, a SDM such as Maxent, felt more than adequate. Alternatively, a rules-based modelling could have been used if more time was available and the extent of the study was expanded (Dreisbach et al., 2002).

The importance of correcting for sampling bias in a Maxent SDM could be considered both as a pro and a con in the development of the method. Because of the underlying assumption that the entire area of interest in a Maxent model has been methodically sampled, it is sometimes easy to forget that Maxent models in general are based on occurrence data that are spatially biassed towards areas that are more easily accessed (Kramer-Schadt et al., 2013; Sillero et al., 2021). One of the options to overcome this, is to apply a type of background manipulation, such as a bias layer, which was the option applied in this thesis. Kramer-Schadt et al. (2013) describes it as "using environmental data that has the same spatial bias as the occurrence data". As pointed out by Phillips et al. (2009), the approach of adding a bias layer has been found to enhance the performance of a number of SDM.

One of the drawbacks is that both land covers are from 2018, which means that they are 6 years old, and a lot can happen with vegetation cover during that time. This could undoubtedly affect the predicted presence of *Cantharellales* in any area that is modelled. The most likely reasons that the vegetation has changed in a forest is logging. Either the trees are thinned out or the entire area is clear cut. Any of these actions will undoubtedly affect the mycorrhizae. Other factors could be forest fires, sickness of the trees caused by stress from insects or wood decay fungi (fungal brackets and toadstools).

In a larger study one way to mitigate this issue would be to digitise aerial orthophotos if those that are available for the chosen area and of a more recent date than the available land cover layer. Of course, this will have its restrictions if the given study area is very large.

Another thing that can cause issues is the importance of keeping track of the extent of all the layers applied in the Maxent model. Maxent's sensitivity made it abundantly clear and, in some ways, helped develop a deeper understanding of the functionality of extent and snap raster in ArcGIS Pro. Subsequently issues like this, which could cause a large time delay while figuring out the issue, should always be considered when using any kind of modelling program or GIS system. Besides, as with any new program that someone has not used prior to starting on a new project or study, there is an understanding that extra time needs to be set aside to learn the ins and outs of that specific system.

6.3 Limitations

There are a set of given limitations when any type of dataset is involved. For one, the quality of the data and its resolution. A second, is the need for collecting the sampling points before knowing how a given model might act depending on the sampling method that was used. Combine this with the fact that mushroom season varies, and had it been a bad mushroom year instead of a good one (which 2023 was), the number of sample points could have been severely compromised. Which could have led to a situation where two few samples would have made the model useless.

Also, given the fact that the external sample points were gathered by several different individuals, and as such provided no control over which sampling method might have been applied, could have influenced the end results. Mostly, the samples were recorded by many different individuals at random. Despite this fact, they were still deemed valid as they would provide an indication of the applicability, if the model worked.

Any disturbance, such as harvesting of trees, will cause a change in fungal diversity and productivity, which is pointed out by Pilz et al. (2003). Moreover, forest stand age and disturbances such as fire will also cause a change in the fungal productivity (Crites & Dale, 1998; III et al., 1999; Stendell et al., 1999). But in many records of decision (ROD), a document that is associated with an environmental impact statement (EIS), assumes that listed fungi are what they call old-growth dependant, which Dreisbach et al. (2002) points out is not always the case as chanterelles can also occur in younger stands.

6.4 Other aspects

The combination of fungi, biodiversity and ecosystem indicator might not be the first thought that comes to mind when these topics are brought up, but as Heilmann-Clausen et al. (2015) notes, fungal red lists have been widely applied in terms of management and conservation activities across Europe. For instance, Sweden is a country that has launched specific action plans that aims to protect fungal habitats (Dahlberg et al., 2010).

It has also been documented that changes in fungal communities can provide an early warning of changes in the local ecosystem and the surrounding environments (Heilmann-Clausen et al., 2015). This includes detection of pollutants and other elements that might be harmful to both humans and wildlife alike, for instance heavy metal pollution (Ediriweera et al., 2022). As such, mycophiles could contribute a lot to the scientific community, with recording and documenting of fungi occurrences in the wild.

Speaking of mycophiles, there are many different competitors for those who like to pick *Cantharellales*. Animals such as wild boar, roe deer, and hares, aside from other mushroom pickers, all benefit from the access of different forest fungi (Elliott et al., 2022; Heilmann-Clausen et al., 2015). For example, Elliott et al. (2022) mentions that fungi are probably one of the primary sources of selenium for mammals in the wild, making it a highly desirable mineral source.

Fungi and especially the ECM (ectomycorrhizal) are affected by agroforestry, especially clear-cutting since the fungi requires the roots of the trees to be alive (Hasby, 2022). Therefore, the complete removal of trees in sections of the forest kills the mycorrhizal and could be considered as habitat loss (Sterkenburg et al., 2019). Thus, the Fennoscandia (a geographical region that includes the Scandinavian peninsula, the Kola Peninsula, mainland Finland and Karelia) tradition of clear-cutting alters the forest ecosystems (Hasby, 2022). Overall, it affects the flora and fauna alike. If the ECM is not completely wiped out it might take decades to return (Varenius et al., 2016; Varenius et al., 2017). In terms of modelling, a new clear-cut area will not be reported or included in data used in a model if not included in the latest land cover. Naturally, this will affect the results and possibly show areas of false positives, that is areas indicating an occurrence probability, where there is none.

Lastly, the Swedish University of Agricultural Sciences is currently developing a web application where the agroforestry community and forest owners will be able to view where the edible- and ECM mushrooms can be found on their land (Swedish University of Agricultural Sciences, 2023). The app is also intended to show how the effects of different forest practices affects the mushrooms. The leader of the project is a researcher by the name of Anders Dahlberg who is cooperating with a former student, Jacob Bertilsson a science illustrator, who will help him develop the application based on data from Swedish Forest Soil Inventory [Markinventeringen] and SLU Swedish Species Information Centre [Artdatabanken]. According to him, there is a big demand for easily accessible information that displays environmental worth that might be within a specific population. He also explains that there is a desire from the forest owners to know how their choice of management could influence the ECM. The overall aim of the project is to contribute to the development of digitally available information on nature conservation for landowners (Swedish University of Agricultural Sciences, 2023).

7 Conclusions

With the main aim of this thesis being that of mapping locations of possible sites where *Cantharellales* mushrooms could be found with the use of SDM and GIS, the following conclusions can be found:

- The Svedala Municipality model was found to be a good model fit. As such, the model should in theory be applicable for use on other fungi species that might be of interest to apply to a SDM such as Maxent. Of course, consideration should be given to the number of sample points, the types of environmental factors available, and the specific characteristics that might apply to the area intended for the model. All three will undoubtedly have an influence on the end results. This cannot be more evident than what the results for the Scania County model shoved.

- In terms of which land cover to choose, the prediction of the model indicated that the use of national, or local land cover, if possible, is to be preferred. Something that would follow a common-sense reasoning. Realistically, a national or local land cover would in most cases contain more detail and perhaps if species specific information, which a global or "international" land cover would not have. Especially if the land cover used in the model was digitised based on photographic imagery of the same year that the sample points were obtained. As with most things however, whichever land cover (or any other environmental factor) is chosen for the model, its availability as well as the time and money spent to acquire it, will ultimately decide which one is chosen.

- Specific local characteristics of a selected area will have an impact on the relevancy of individual environmental factors. It appeared that the SLU classified soil moisture layer with its categorical values suited the "smaller" area of the Svedala Municipality model, whereas the Scania County model produced better AUC values with the continuous SLU soil moisture layer. It was also established that the topography of the Svedala area, where the sample points were collected was favoured and influenced the model to a higher degree than anticipated.

- It is possible to conclude that a valid model can be achieved using open-source data. This makes the idea that other mycophiles and GIS enthusiasts could use this thesis as a template for how to go about finding sites with a high probability of a certain fungi occurrence even greater. As stated in the introduction, this was one of the intended goals with the thesis.

- Continuing on that note, the model also showed that a small set of sample data did produce results that are considered fair to good. Consequently, a person would not be required to scavenge across a huge area during a limited time period to secure enough sample points for a model to work. As this thesis has shown, 50 to 60 sample points is more than enough in an area that covers the same size as that of the Svedala Municipality. Even the number of sample points for the Scania County model proved to work, even though it only produces what is considered a fair model. But most likely, a study with a bigger and more structured sample technique would likely produce an even better model fit.

- Lastly, both models confirmed the statements of Dreisbach et al. (2002), climatological factors are not necessary for the Maxent model to work. They will however remain a complex element for any that intends to add or exclude it as a factor in an SDM.

8 References

Aina, O. O. (2022). *Land suitability analysis for harvest species using a multi-criteria and GIS approach in northern Saskatchewan* University of Saskatchewan].

Alemayehu, B., Suarez Minguez, J., & Rosette, J. (2024). Modeling the Spatial Distribution of Acacia decurrens Plantation Forests Using PlanetScope Images and Environmental Variables in the Northwestern Highlands of Ethiopia. *Forests*, *15*, 18. https://doi.org/10.3390/f15020277

Araújo, M. B., Pearson, R. G., Thuiller, W., & Erhard, M. (2005). Validation of species–climate impact models under climate change. *Global Change Biology*, *11*(9), 1504-1513. https://doi.org/10.1111/j.1365-2486.2005.01000.x

Austin, M. (1980). Searching for a model for use in vegetation analysis. *Vegetatio*, *42*, 11-21.

Baldwin, R. A. (2009). Use of Maximum Entropy Modeling in Wildlife Research [Review]. *Entropy*, *11*(4), 854-866. https://doi.org/10.3390/e11040854

Brunet, J. (2005). *Skånes skogar – historia, mångfald och skydd*. Länsstyrelsen i Skånes Län

Cikarge, G. P., & Arifin, F. (2018). Oyster Mushrooms Humidity Control Based On Fuzzy Logic By Using Arduino ATMega238 Microcontroller. *Journal of Physics: Conference Series*, *1140*(1), 012002. https://doi.org/10.1088/1742-6596/1140/1/012002

Crites, S., & Dale, M. R. (1998). Diversity and abundance of bryophytes, lichens, and fungi in relation to woody substrate and successional stage in aspen mixedwood boreal forests. *Canadian Journal of Botany*, *76*(4), 641-651. https://doi.org/10.1139/b98-030

Dahlberg, A., Genney, D. R., & Heilmann-Clausen, J. (2010). Developing a comprehensive strategy for fungal conservation in Europe: current status and future needs. *Fungal Ecology*, *3*(2), 50-64. https://doi.org/10.1016/j.funeco.2009.10.004

Dramani, R., Gouwakinnou, G. N., Houdanon, R. D., De Kesel, A., Minter, D., & Yorou, N. S. (2022). Ecological niche modelling of Cantharellus species in Benin, and revision of their conservation status. *Fungal Ecology*, *60*. https://doi.org/10.1016/j.funeco.2022.101174

Dreisbach, T. A., Smith, J. E., & Molina, R. (2002). Challenges of modeling fungal habitat: when and where do you find chanterelles. *Predicting Species Occurrence: Issues of Scale and Accuracy*, 475-481.

Dyshko, V., Hilszczańska, D., Davydenko, K., Matić, S., Moser, W. K., Borowik, P., & Oszako, T. (2024). An Overview of Mycorrhiza in Pines: Research, Species, and Applications. *Plants (Basel)*, *13*(4). https://doi.org/10.3390/plants13040506

Ediriweera, A. N., Karunarathna, S. C., Yapa, P. N., Schaefer, D. A., Ranasinghe, A. K., Suwannarach, N., & Xu, J. (2022). Ectomycorrhizal Mushrooms as a Natural Bio-Indicator for Assessment of Heavy Metal Pollution. *Agronomy*, *12*(5), 1041. https://doi.org/10.3390/agronomy12051041

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and distributions*, *17*(1), 43-57.

Elliott, T., Truong, C., Jackson, S., Zúñiga, C., Trappe, J., & Vernes, K. (2022). Mammalian Mycophagy: a Global Review of Ecosystem Interactions Between Mammals and Fungi. *Fungal Systematics and Evolution*, *9*, 99-159. https://doi.org/10.3114/fuse.2022.09.07

Engler, R., Guisan, A., & Rechsteiner, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, *41*(2), 263-274. https://doi.org/10.1111/j.0021-8901.2004.00881.x

Esri. (n.d.). *Presence-only Prediction (MaxEnt)*. Retrieved 25th of April from https://pro.arcgis.com/en/pro-app/3.1/tool-reference/spatial-statistics/how-presence-only-prediction-works.htm

European Environment Agency. (2023, Febuary 28). *The Natura 2000 protected areas network*. Retrieved 4th of April from https://www.eea.europa.eu/themes/biodiversity/natura-2000/the-natura-2000-protected-areas-network

Faizollahzadeh Ardabili, S., Mahmoudi, A., Mesri Gundoshmian, T., & Roshanianfard, A. (2016). Modeling and comparison of fuzzy and on/off controller in a mushroom growing hall. *Measurement*, *90*, 127-134. https://doi.org/10.1016/j.measurement.2016.04.050

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*(1), 38-49. http://www.jstor.org.ludwig.lub.lu.se/stable/44519240

Franklin, J. (2010). *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511810602

GBIF. (2000-2023). *Occurrences* https://www.gbif.org/occurrence/search

GBIF. (n.d.). *Free and open access to biodiversity data*. Retrieved 12th of June from https://www.gbif.org/

Gry, J., & Petersen, J. H. (2012). *Mushrooms traded as food. Vol II section 1 : Nordic Risk assessments and background on edible mushrooms, suitable for commercial marketing and background lists. For industry, trade and food inspection. Background information and guidance lists on mushrooms Jørn Gry [and eight others] ; cover photo, Jens H. Petersen ; photo, Jens H. Petersen*. Nordic Council of Ministers. https://doi.org/10.6027/TN2012-543

Hall, I. R., Yun, W., & Amicucci, A. (2003). Cultivation of edible ectomycorrhizal mushrooms. *TRENDS in Biotechnology*, *21*(10), 433-438. https://doi.org/10.1016/S0167-7799(03)00204-X

Hall, L. S., Krausman, P. R., & Morrison, M. L. (1997). The habitat concept and a plea for standard terminology. *Wildlife Society Bulletin*, *25*(1), 173-182. http://www.jstor.org/stable/3783301

Hasby, F. (2022). *Impacts of clear-cutting on soil fungal communities and their activities in boreal forests : a metatranscriptomic approach* [Theses]. Swedish University of Agricultural Sciences. https://res.slu.se/id/publ/115772

Heilmann-Clausen, J., Barron, E. S., Boddy, L., Dahlberg, A., Griffith, G. W., Nordén, J., Ovaskainen, O., Perini, C., Senn-Irlet, B., & Halme, P. (2015). A fungal perspective on conservation biology. *Conservation Biology*, *29*(1), 61-68. https://doi.org/10.1111/cobi.12388

Holmberg, B., & Näslund, G. K. (1988). *VÅR SVAMPBOK* (2 ed.). Raben & Sjögren. (1982)

Holmberg, P., & Holmberg, I. (2021). *SVAMPBOK*. Bokforlaget Semic. (2007)

III, W. C., Carey, A. B., Trappe, J. M., Molina, R., & Thysell, D. (1999). Diversity and productivity of hypogeous fungal sporocarps in a variably thinned Douglas-fir forest. *Canadian Journal of Forest Research*, *29*(8), 1259-1268. https://doi.org/10.1139/x99-082

INaturalist. (2000-2023). *Sweden* https://www.inaturalist.org/places/sweden#/places/sweden=

INaturalist. (2022, 2022, May 4). *What is it*. Retrieved 12th of June from https://www.inaturalist.org/pages/what+is+it

Institutionen för skoglig ekologi och skötsel (SLU). (2020). *SLU Markfuktighetskarta* ([Data set]. https://www.skogsstyrelsen.se/globalassets/sjalvservice/karttjanster/geodatatjanster/produktbeskrivningar/markfuktighetskarta-slu---produktbeskrivning.pdf

Jansson, E. (2006). *Landscape Character Assessment - a Swedish trial* [Swedish University of Agricultural Sciences]. https://stud.epsilon.slu.se/11073/1/jansson_e_170929.pdf

Jiménez-Valverde, A. (2011). Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, *21*(4), 498-507. https://doi.org/10.1111/j.1466-8238.2011.00683.x

Khaund, P., & Joshi, S. R. (2016). Lentinula edodes based GIS mapping, biometabolites and antiinflamatory activity of wild edible mushrooms from tropical 'sacred grove' forests of Meghalaya, India. *Revista de Biología Tropical*, *64*(1), 247–257. https://doi.org/10.15517/rbt.v64i1.20349

Kozarski, M., Klaus, A., Vunduk, J., Zizak, Z., Niksic, M., Jakovljevic, D., Vrvic, M. M., & Van Griensven, L. J. L. D. (2015). Nutraceutical properties of the methanolic extract of edible mushroom Cantharellus cibarius (Fries): primary mechanisms [10.1039/C5FO00312A]. *Food & Function*, *6*(6), 1875-1886. https://doi.org/10.1039/C5FO00312A

Kramer-Schadt, S., Niedballa, J., Pilgrim, J. D., Schröder, B., Lindenborn, J., Reinfelder, V., Stillfried, M., Heckmann, I., Scharf, A. K., Augeri, D. M., Cheyne, S. M., Hearn, A. J., Ross, J., Macdonald, D. W., Mathai, J., Eaton, J., Marshall, A. J., Semiadi, G., Rustam, R., . . . Wilting, A. (2013). The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and distributions*, *19*(11), 1366-1379. https://doi.org/https://doi.org/10.1111/ddi.12096

Lantmäteriet. (2019a). *GSD-Elevation data, Grid 2+* ([Laser data]. https://www.lantmateriet.se/sv/geodata/vara-produkter/geodata-for-forskning-utbildning-och-kulturverksamheter/

Lantmäteriet. (2019b). *PRODUCT DESCRIPTION - CORINE Land Cover* (Version (3.0) [Data set]. https://www.lantmateriet.se/contentassets/703ae721445f4398a2fd3c890472bbf8/e_clcshmi.pdf

Länsstyrelsen Skåne. *Administration och förvaltning*. Retrieved 4th of March from https://www.lansstyrelsen.se/skane/besoksmal/kulturmiljoprogram/kulturmiljoprogram-skanes-historia-och-utveckling/kulturmiljoprogram-administrativa-landskap/administration-och-forvaltning.html

MacKenzie, D. (2009). What are the issues with Presence-Absence data for wildlife managers? *Journal of Wildlife Management*, *69*, 849-860. https://doi.org/10.2193/0022-541X(2005)069[0849:WATIWP]2.0.CO;2

Malmö Stad. (2023, October 30). *Bokskogen i Torup - Djur och natur*. Retrieved 27th of March from https://malmo.se/Uppleva-och-gora/Natur-och-parker/Bokskogen-i-Torup/Djur-och-natur.html

Margulis, L., & Chapman, M. J. (2009). Chapter Four - KINGDOM FUNGI. In L. Margulis & M. J. Chapman (Eds.), *Kingdoms and Domains (Fourth Edition)* (pp. 379-409). Academic Press. https://doi.org/10.1016/B978-0-12-373621-5.00004-0

Methakullachat, D., & Witchayangkoon, B. (2019). COORDINATES COMPARISON OF GOOGLE® MAPS AND ORTHOPHOTO MAPS IN THAILAND [Article]. *International Transaction Journal of Engineering Management & Applied Sciences & Technologies*, *10*(17), 8, Article 10a170. https://doi.org/10.14456/itjemast.2019.241

Miller, J. (2010). Species Distribution Modeling. *Geography Compass*, *4*(6), 490-509. https://doi.org/10.1111/j.1749-8198.2010.00351.x

Mossberg, B., Nilsson, S., & Persson, O. (1987). *Svampar i naturen*. Wahlström & Widstrand. (1977)

Naturvårdsverket. (2020). *Nationella marktäckedata 2018 basskikt* (Version (2.2) [Data set]. https://www.naturvardsverket.se/4a43ca/contentassets/37e8b38528774982b5840554f02a1f81/produktbeskrivning-nmd-2018-basskikt-v2-2.pdf

Peintner, U., Schwarz, S., Mešić, A., Moreau, P.-A., Moreno, G., & Saviuc, P. (2013). Mycophilic or Mycophobic? Legislation and Guidelines on Wild Mushroom Commerce Reveal Different Consumption Behaviour in European Countries. *PloS one*, *8*(5), e63926. https://doi.org/10.1371/journal.pone.0063926

Pent, M., Poldmaa, K., & Bahram, M. (2017). Bacterial Communities in Boreal Forest Mushrooms Are Shaped Both by Soil Parameters and Host Identity [Article]. *Frontiers in Microbiology*, *8*, 13, Article 836. https://doi.org/10.3389/fmicb.2017.00836

Phillips, S. J. (2005). A brief tutorial on Maxent. *At&t Research*, *190*(4), 231-259. http://biodiversityinformatics.amnh.org/open_source/maxent/

Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., & Blair, M. E. (2017). Opening the black box: an open-source release of Maxent. *Ecography*, *40*(7), 887-893. https://doi.org/10.1111/ecog.03049

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological modelling*, *190*(3-4), 231-259.

Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, *31*(2), 161-175. https://doi.org/10.1111/j.0906-7590.2008.5203.x

Phillips, S. J., Dudik, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample Selection Bias and Presence-Only Distribution Models: Implications for Background and Pseudo-Absence Data. *Ecological applications*, *19*(1), 181-197. https://www.jstor.org/stable/27645958

Pilz, D., Norvell, L., Danell, E., & Molina, R. (2003). *Ecology and management of commercially harvested chanterelle mushrooms*. http://dx.doi.org/10.2737/PNW-GTR-576

Region Skåne. (2024). *Discover Skåneleden*. Retrieved 4th of April from https://www.skaneleden.se/en

Ribeiro, B., Guedes de Pinho, P., Andrade, P. B., Baptista, P., & Valentão, P. (2009). Fatty acid composition of wild edible mushrooms species: A comparative study. *Microchemical Journal*, *93*(1), 29-35. https://doi.org/10.1016/j.microc.2009.04.005

Sillero, N., Arenas-Castro, S., Enriquez-Urzelai, U., Vale, C. G., Sousa-Guedes, D., Martínez-Freiría, F., Real, R., & Barbosa, A. M. (2021). Want to model a species niche? A step-by-step guideline on correlative ecological niche modelling. *Ecological modelling*, *456*, 109671. https://doi.org/10.1016/j.ecolmodel.2021.109671

Smith, S. E., & Read, D. J. (2008). *Mycorrhizal symbiosis*. Academic press.

Stendell, E. R., Horton, T. R., & Bruns, T. D. (1999). Early effects of prescribed fire on the structure of the ectomycorrhizal fungus community in a Sierra Nevada ponderosa pine forest. *Mycological Research*, *103*(10), 1353-1359. https://doi.org/10.1017/S0953756299008618

Sterkenburg, E., Clemmensen, K. E., Lindahl, B. D., & Dahlberg, A. (2019). The significance of retention trees for survival of ectomycorrhizal fungi in clear-cut Scots pine forests. *Journal of Applied Ecology*, *56*(6), 1367-1378. https://doi.org/10.1111/1365-2664.13363

Sveriges geologiska undersökning. (2018). *PRODUKT: JORDARTER 1:25 000-1:100 000* ([Data set]. https://www.sgu.se/globalassets/produkter/publikationer/produktbeskrivningar/jordarter-25-100000-beskrivning.pdf

Swedish Meteorological and Hydrological Institute. (2023a). *Ladda ner meteorologiska observationer - Lufttemperatur (månad) och Lufttemperatur (dygn)* ([Data set]. https://www.smhi.se/data/meteorologi/ladda-ner-meteorologiska-observationer#param=airTemperatureMeanMonth,stations=core,stationid=53300

Swedish Meteorological and Hydrological Institute. (2023b). *Ladda ner meteorologiska observationer - Nederbördsmängd (dygn)* ([Data set]. https://www.smhi.se/data/meteorologi/ladda-ner-meteorologiska-observationer#param=precipitation24HourSum,stations=core,stationid=52350

Swedish University of Agricultural Sciences. (2023). *Se svampen i din skog!* Retrieved 22nd of March from https://www.slu.se/ew-nyheter/2023/3/se-svampen-i-din-skog/

Syfert, M. M., Smith, M. J., & Coomes, D. A. (2013). The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. *PloS one*, *8*(2), e55158. https://dx.doi.org/10.1371/journal.pone.0055158

Söderström, B. (1992). The ecological potential of the ectomycorrhizal mycelium.

Varenius, K., Kårén, O., Lindahl, B., & Dahlberg, A. (2016). Long-term effects of tree harvesting on ectomycorrhizal fungal communities in boreal Scots pine forests. *Forest Ecology and Management*, *380*, 41-49. https://doi.org/10.1016/j.foreco.2016.08.006

Varenius, K., Lindahl, B., & Dahlberg, A. (2017). Retention of seed trees fails to lifeboat ectomycorrhizal fungal diversity in harvested Scots pine forests. *FEMS microbiology ecology*, *93*. https://doi.org/10.1093/femsec/fix105

Wang, M., Chen, H., & Lei, M. (2022). Identifying potentially contaminated areas with MaxEnt model for petrochemical industry in China. *Environmental Science and Pollution Research*, *29*(36), 54421-54431. https://doi.org/10.1007/s11356-022-19697-8

Wing, M. G., Eklund, A., & Kellogg, L. D. (2005). Consumer-grade global positioning system (GPS) accuracy and reliability [Article]. *Journal of Forestry*, *103*(4), 169-173. https://doi.org/10.1093/jof/103.4.169

Yadav, S., & Shukla, S. (2016). *Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification*. https://doi.org/10.1109/IACC.2016.25

Young, N., Carter, L., & Evangelista, P. (2011). A MaxEnt model v3. 3.3 e tutorial (ArcGIS v10). *Natural Resource Ecology Laboratory, Colorado State University and the National Institute of Invasive Species Science*. http://ibis.colostate.edu/webcontent/ws/coloradoview/tutorialsdownloads/a_maxent_model_v7.pdf

Zandbergen, P. A., & Barbeau, S. J. (2011). Positional Accuracy of Assisted GPS Data from High-Sensitivity GPS-enabled Mobile Phones [Article]. *Journal of Navigation*, *64*(3), 381-399. https://doi.org/10.1017/s0373463311000051

9 Appendices

Table A1: Overview of data used for the production of maps and models.

Data folder	Content	Source	Projection	Link
CORINE land cover Lantmäteriet (Skane)	Land cover Vector	European Union (EU) - provided by The Land Survey ([Lantmäteriet])	SWEREF99TM	https://www.lantmateriet.se/sv/geodata/vara-produkter/internationell-samverkan/corine-land-cover/#anchor-4
GSD Höjddata g2+	DEM Elevation data 2 m resolution Raster	The Land Survey ([Lantmäteriet])	SWEREF99TM RH2000	e-service ”Geodata Extraction Tool” (GET) https://www.lantmateriet.se/sv/geodata/vara-produkter/geodata-for-forskning-utbildning-och-kulturverksamheter/
Jordarter_25_100k	Soil types ([JORDARTER]) 1:25 000-1:100 000 (JG2 layer) Soil layer Vector	Geological survey of Sweden ([Sveriges geologisk undersökning – SGU])	SWEREF99TM	e-service ”Geodata Extraction Tool” (GET) https://www.lantmateriet.se/sv/geodata/vara-produkter/geodata-for-forskning-utbildning-och-kulturverksamheter/
SLU Markfuktighetskarta	SLU Soil moisture layer 2 m resolution Raster	Swedish University of Agricultural Sciences ([Sveriges lantbruksuniversitet -SLU])	SWEREF99TM	e-service ”Geodata Extraction Tool” (GET) https://www.lantmateriet.se/sv/geodata/vara-produkter/geodata-

				for-forskning-utbildning-och-kulturverksamheter/
National_landcover_data_2018_Scania	National land cover database (NMD) 10 m resolution Raster	The Swedish Environmental Protection Agency ([Naturvårdsverket])	SWEREF99TM	https://geodata.naturvardsverket.se/nedladdning/marktacke/NMD2018/bas_lan_ogen/
Orthogrb_allres	Ortho photo 0.25 m resolution Raster	The Land Survey ([Lantmäteriet])	SWEREF99TM RH2000	e-service "Geodata Extraction Tool" (GET) https://www.lantmateriet.se/sv/geodata/vara-produkter/geodata-for-forskning-utbildning-och-kulturverksamheter/
Points	Locations of confirmed sites with mushrooms in Bokskogen and Holmeja skogen in the Municipality of Svedala	Collected in the field during the 2023rd mushrooms season	WGS84	Author
SLU Markfuktighetskarta (MFKKlassad)	SLU classified Soil moisture layer 2 m resolution Raster	Swedish University of Agricultural Sciences ([Sveriges lantbruksuniversitet -SLU])	SWEREF99TM	e-service "Geodata Extraction Tool" (GET) https://www.lantmateriet.se/sv/geodata/vara-produkter/geodata-for-forskning-utbildning-och-kulturverksamheter/

Län, kommuner och LA-regioner	Shape file of county and municipalities of Sweden Vector	The Land Survey ([Lantmäteriet])	SWEREF99TM	e-service ”Geodata Extraction Tool” (GET) https://www.lantmateriet.se/sv/geodata/vara-produkter/geodata-for-forskning-utbildning-och-kulturverksamheter/
Sweden_shapefile	Point data	European Environment Agency	WGS84	https://www.eea.europa.eu/data-and-maps/data/eea-reference-grids-2/gis-files/sweden-shapefile
External sample points GBIF	Point data	Gfib.org	WGS84	https://www.gbif.org/occurrence/search
External sample points Inaturalist	Point data	Inaturalist.org	WGS84	https://www.inaturalist.org/places/sweden#/places/sweden=
Precipitation	Precipitation data Mean daily for 2023	SMHI – the Swedish Meteorological and Hydrological Institute ([Sveriges meteorologiska och hydrologiska institut])	SWEREF99 TM	https://www.smhi.se/data/meteorologi/ladda-ner-meteorologiska-observationer#param=precipitation24HourSum,stations=core,stationid=52350
Temperature	Air temperature Mean daily for August – October Mean monthly for 2023	SMHI – the Swedish Meteorological and Hydrological Institute	SWEREF99 TM	https://www.smhi.se/data/meteorologi/ladda-ner-meteorologiska-observationer#param=airTemperatureMeanMonth,stations=core,stationid=53300

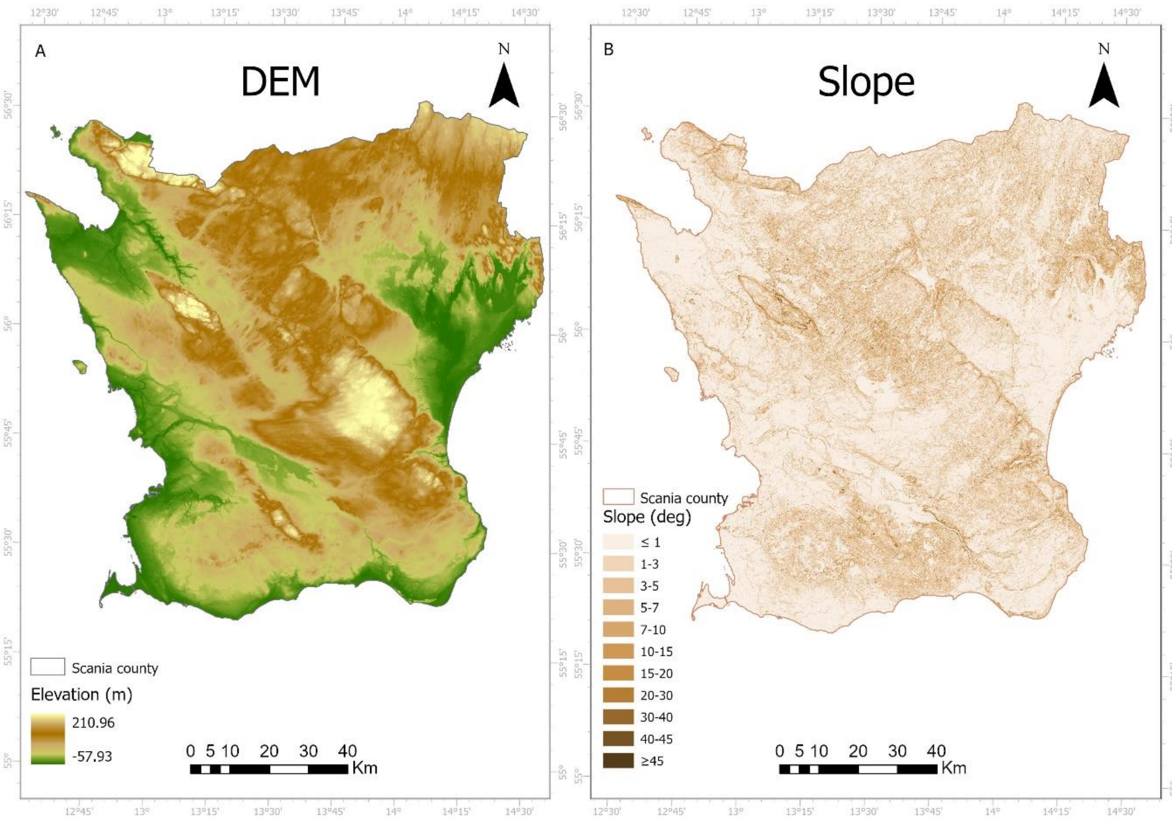

*Figure A1: **A** shows the DEM for the Scania County at a 10m spatial resolution, indicating a maximum elevation of around 211m and a minimum around -58m. **B** depicts the different slopes in degrees and their variations within the Municipality. From nearly level (≤1°) to very steep (≥45°).*

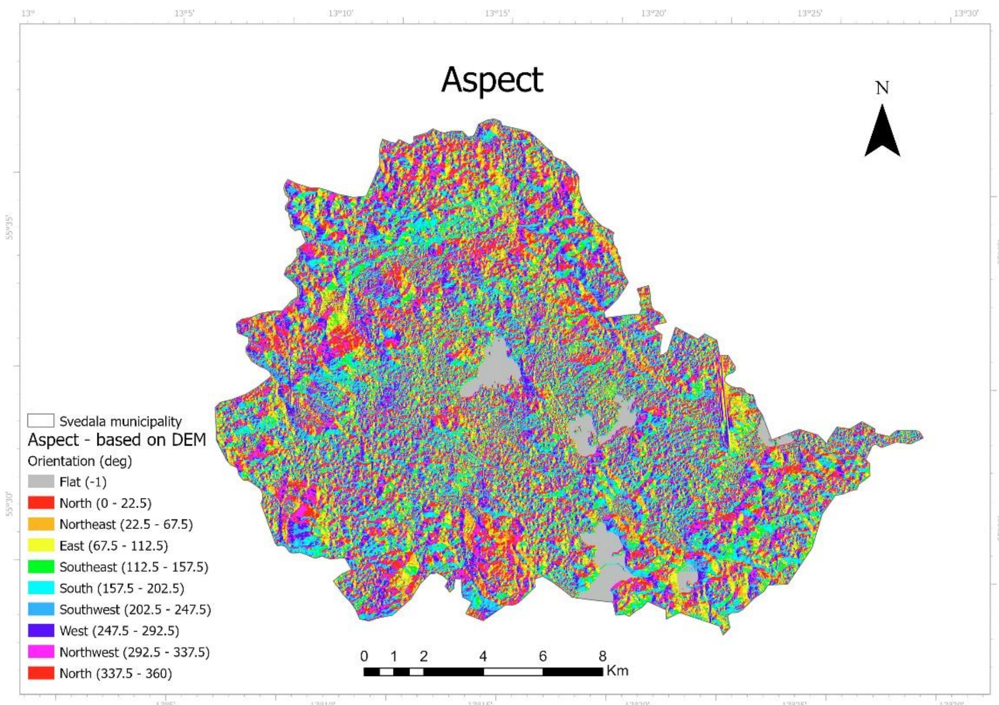

Figure A2: The aspect based on the DEM layer identifies the orientation (compass direction), in degrees, that the corresponding downhill slope faces for each location.

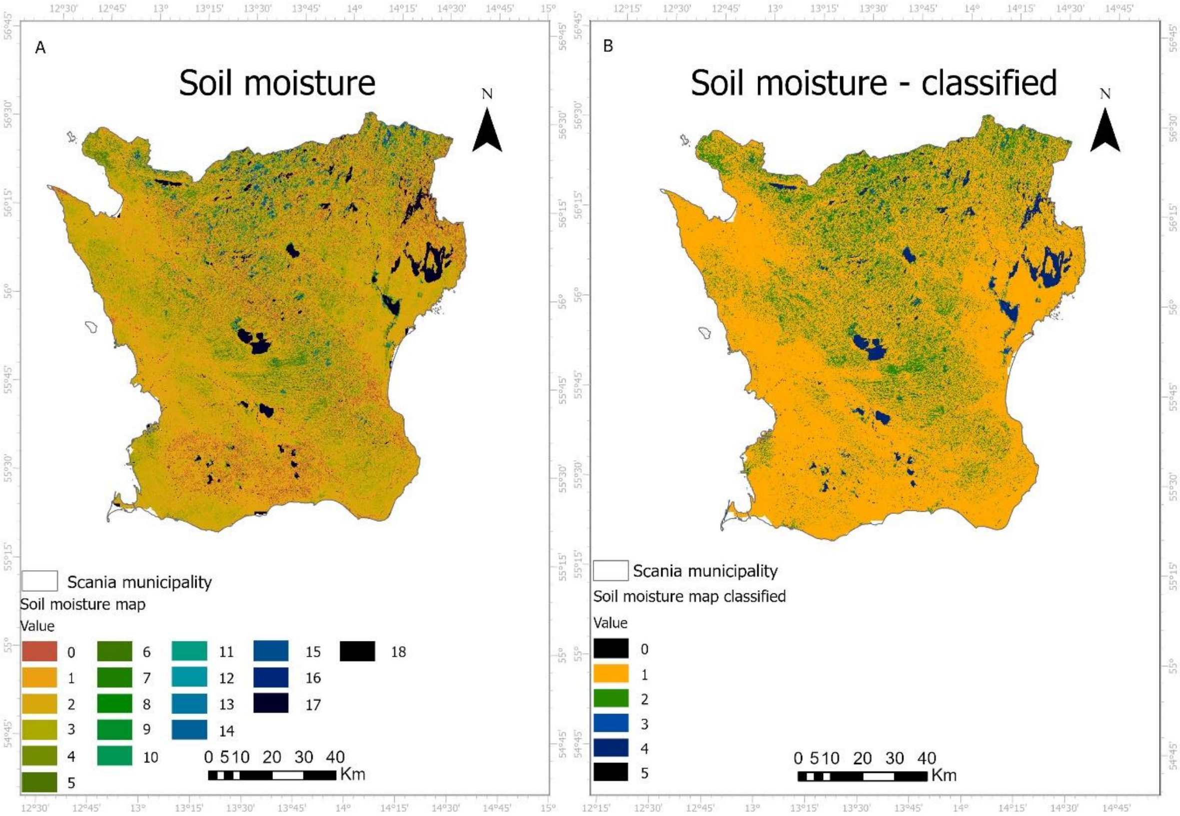

*Figure A3: The two SLU soil moisture layers at a 10m spatial resolution, where **A** has been developed to display continuous values (to create smother transitions between the classes) and **B** categorial values.*

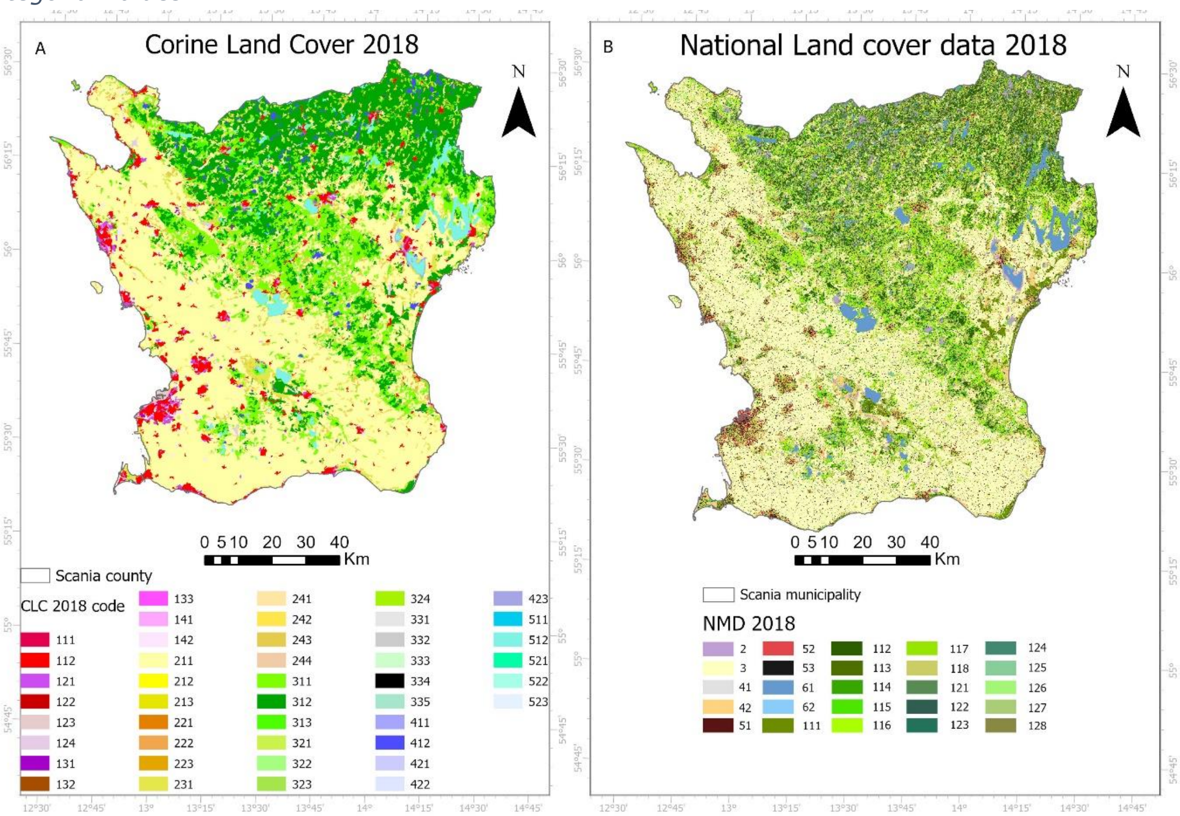

*Figure A4: The current CORINE land cover layer from 2018 (the layer is shown in its original vector format) (**A**) with its classifications ranging from red and purple colours for built-up areas, yellow and brown colours for agricultural lands, green colours for forests and blue colours for water bodies. The National land cover data from 2018 at a 10m spatial resolution (**B**), which is also the most current national land cover, displays dark red to black colours as built-up areas, yellow to orange as open ground and agricultural land, purple and blue colours as wetland and water bodies, and green colours as forest.*

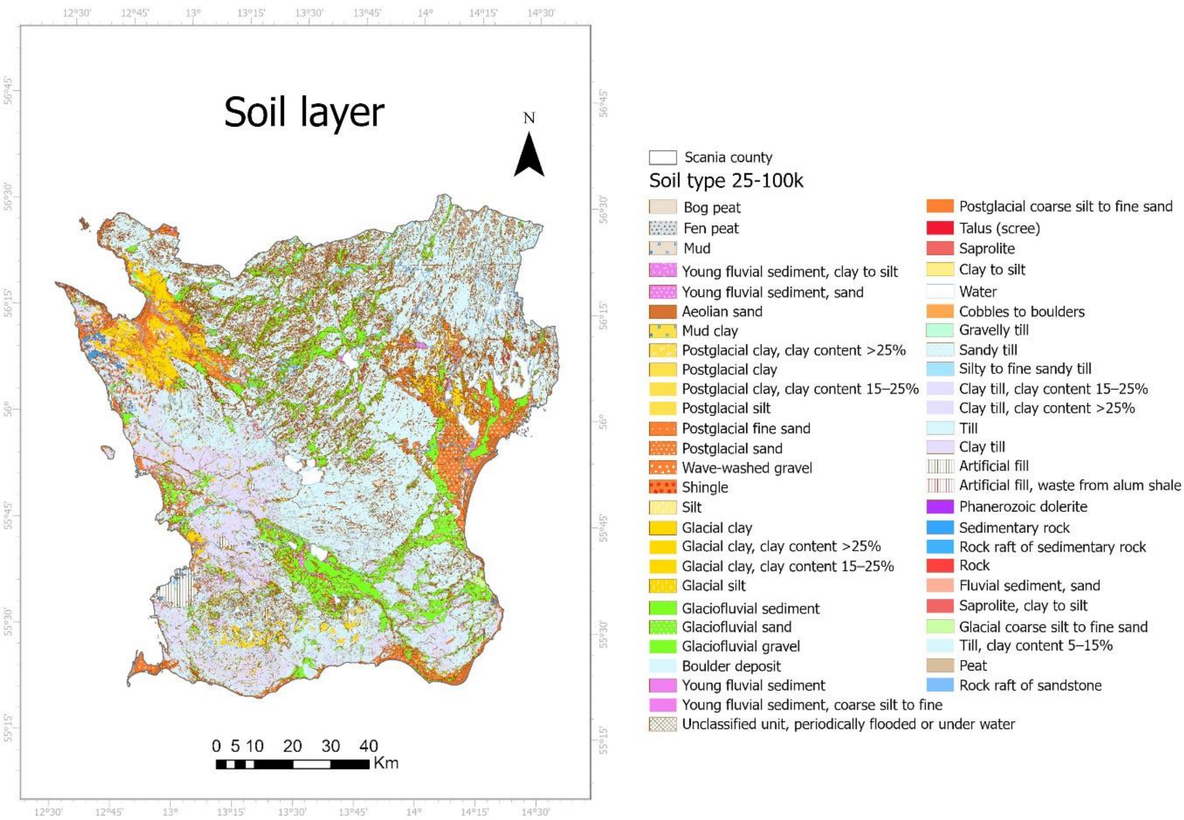

Figure A5: The soil map shows the extend of the soil types in and around the surface including any instance of larger rocks at ground level. The layer is shown in its original vector format.

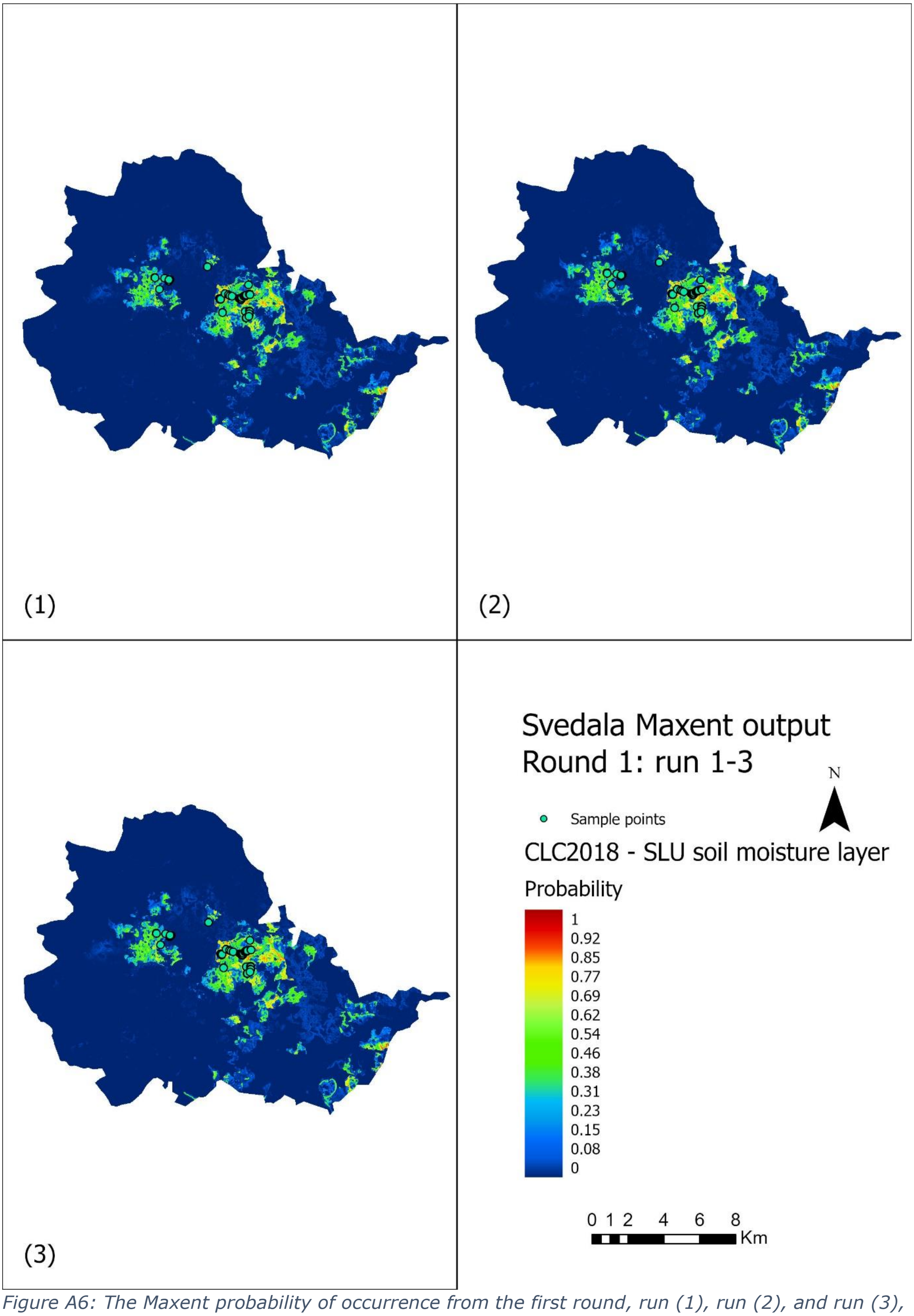

Figure A6: The Maxent probability of occurrence from the first round, run (1), run (2), and run (3), with the environmental factor layer combination of the CLC 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

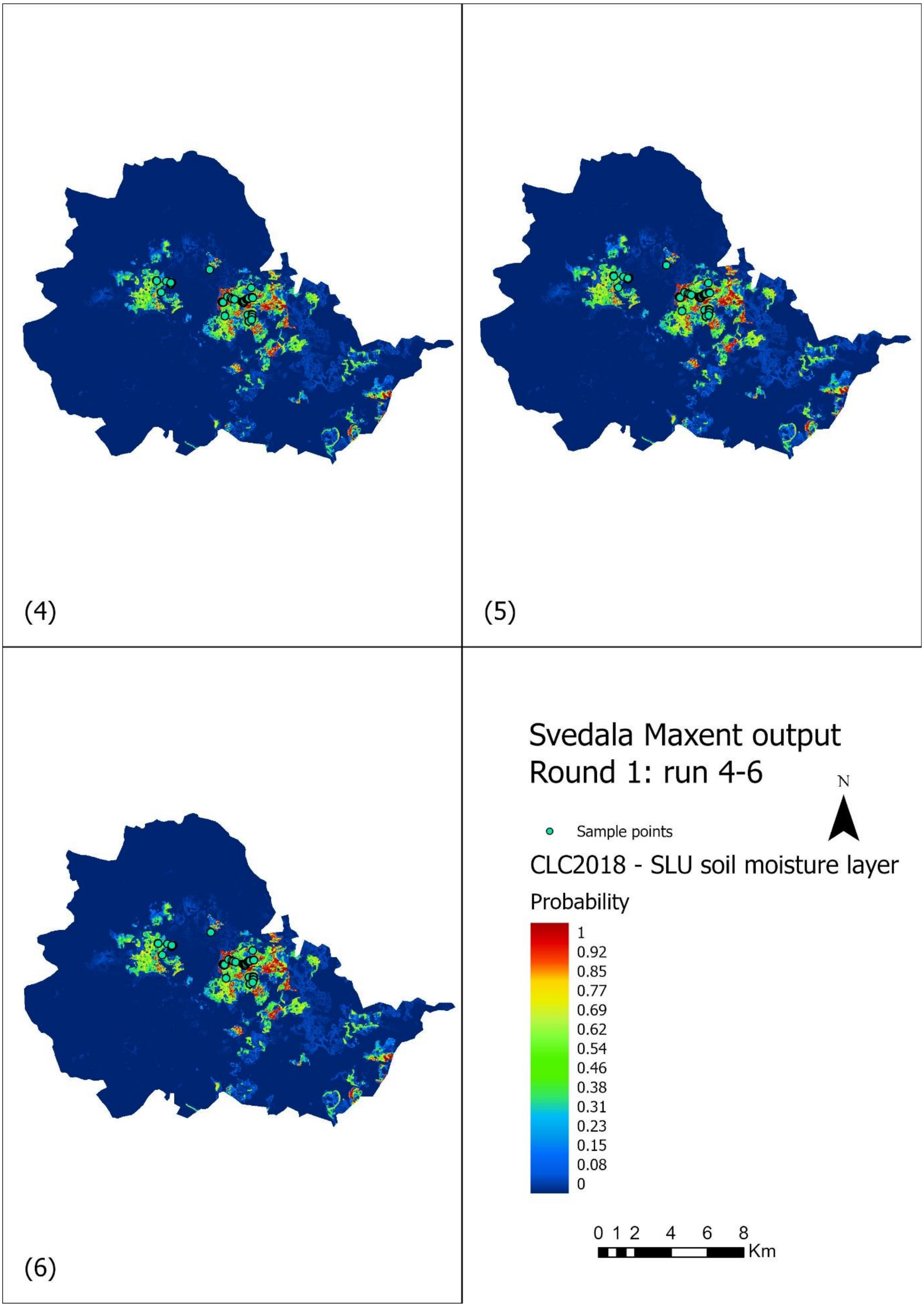

Figure A7: The Maxent probability of occurrence from the first round, run (4), run (5), and run (6), with the environmental factor layer combination of the CLC 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

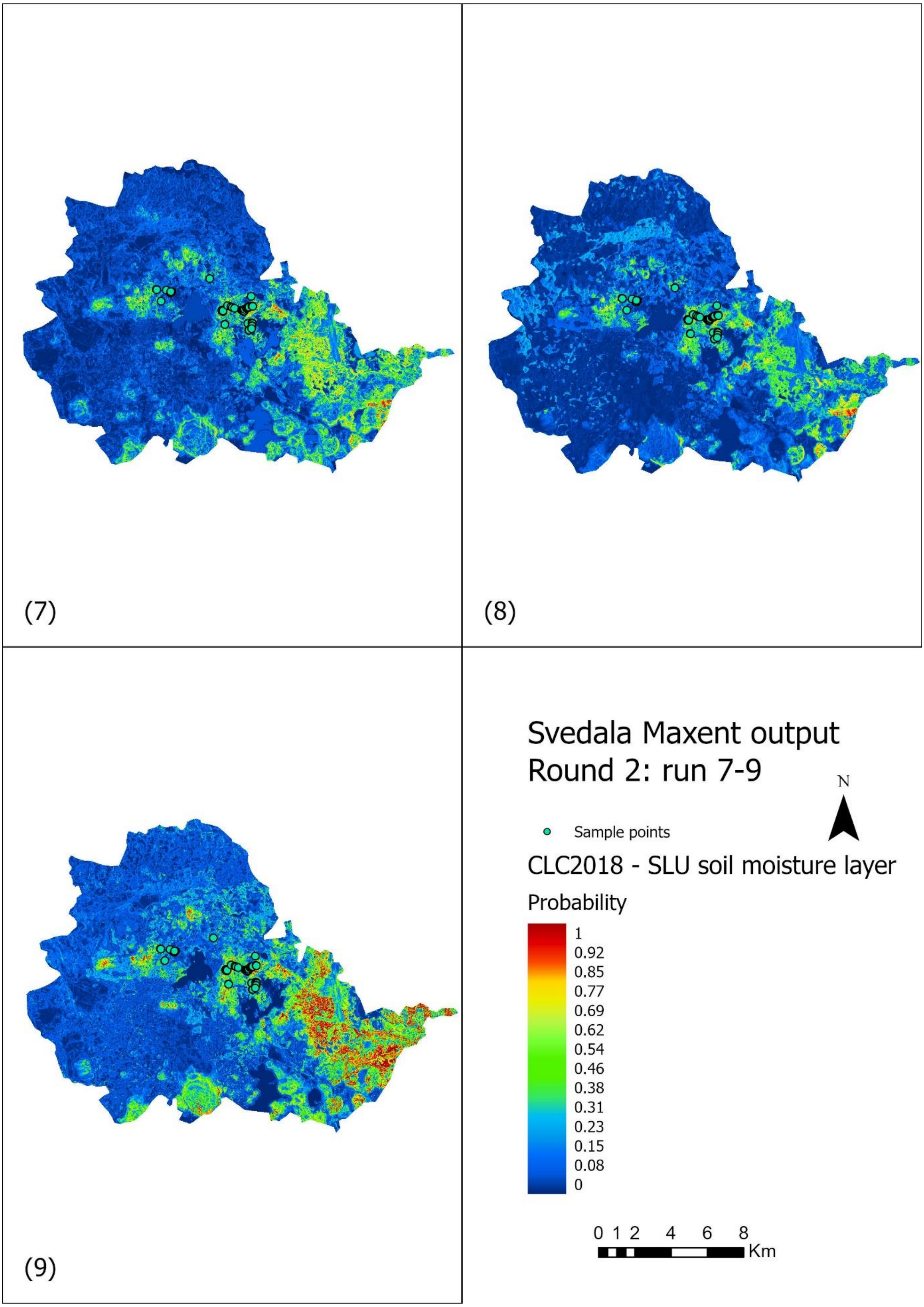

Figure A8: The Maxent probability of occurrence from the second round, run (7), run (8), and run (9), with the environmental factor layer combination of the CLC 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

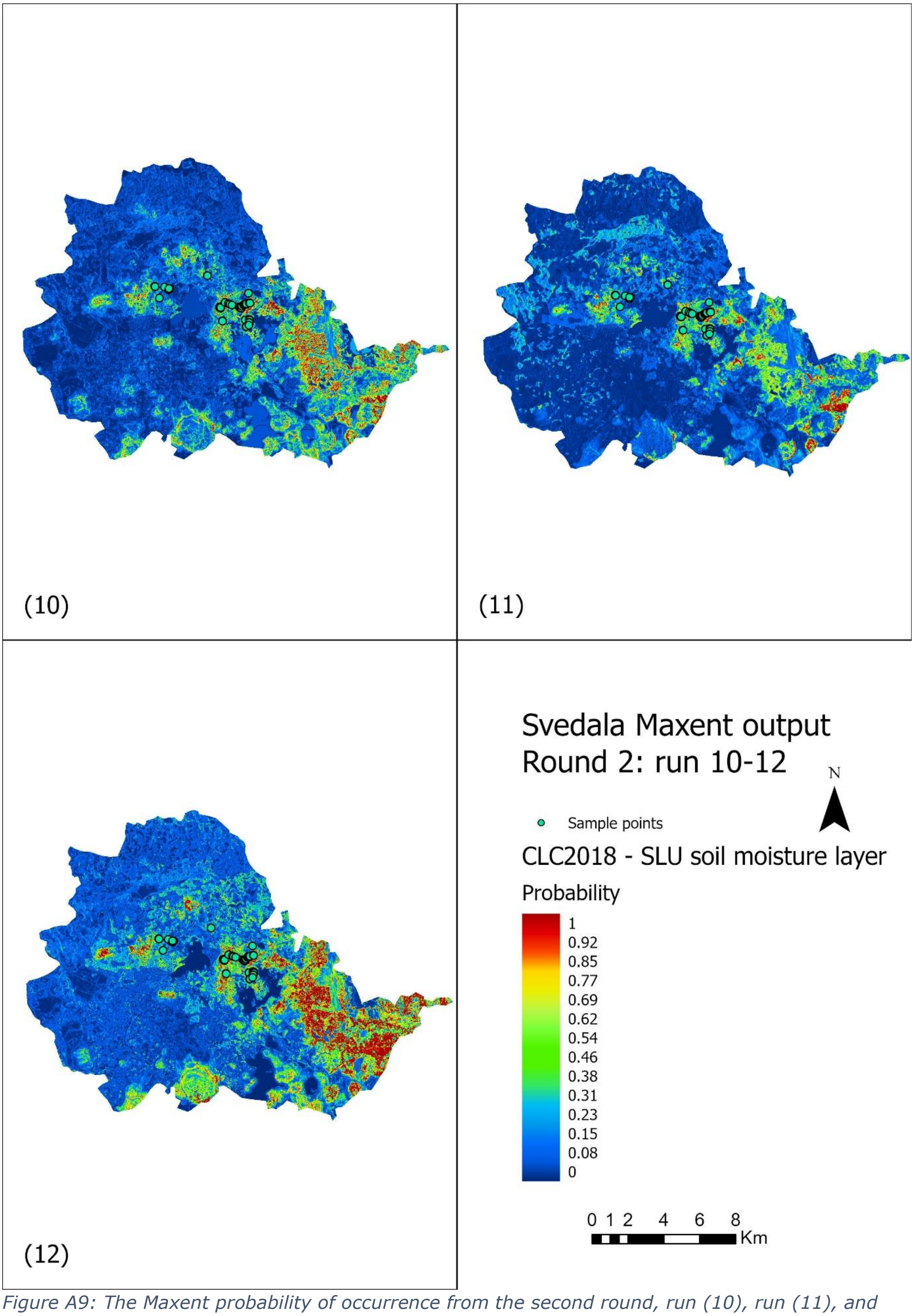

Figure A9: The Maxent probability of occurrence from the second round, run (10), run (11), and run (12), with the environmental factor layer combination of the CLC 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

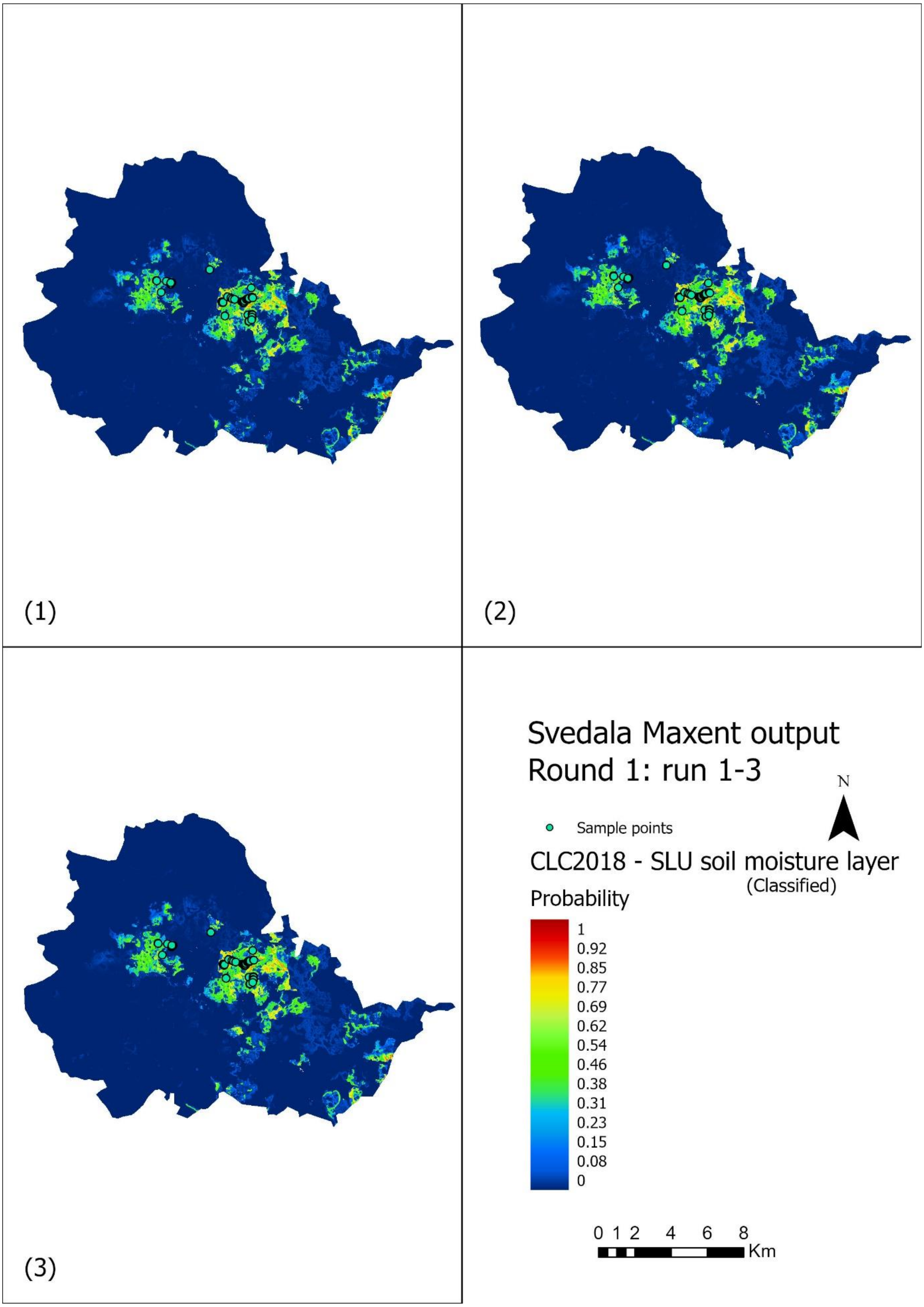

Figure A10: The Maxent probability of occurrence from the first round, run (1), run (2), and run (3), with the environmental factor layer combination of the CLC 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

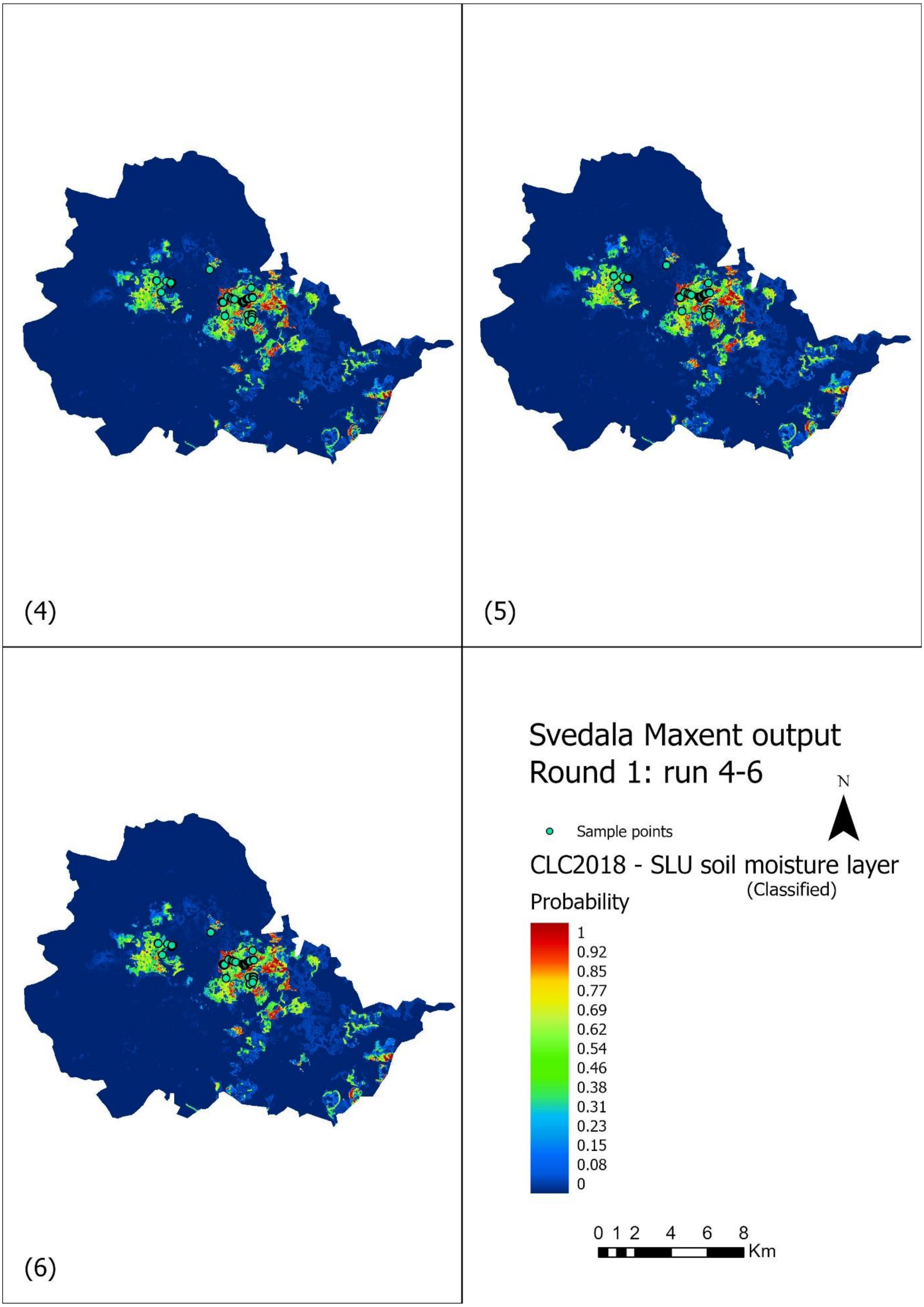

Figure A11: The Maxent probability of occurrence from the second round, run (4), run (5), and run (6), with the environmental factor layer combination of the CLC 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

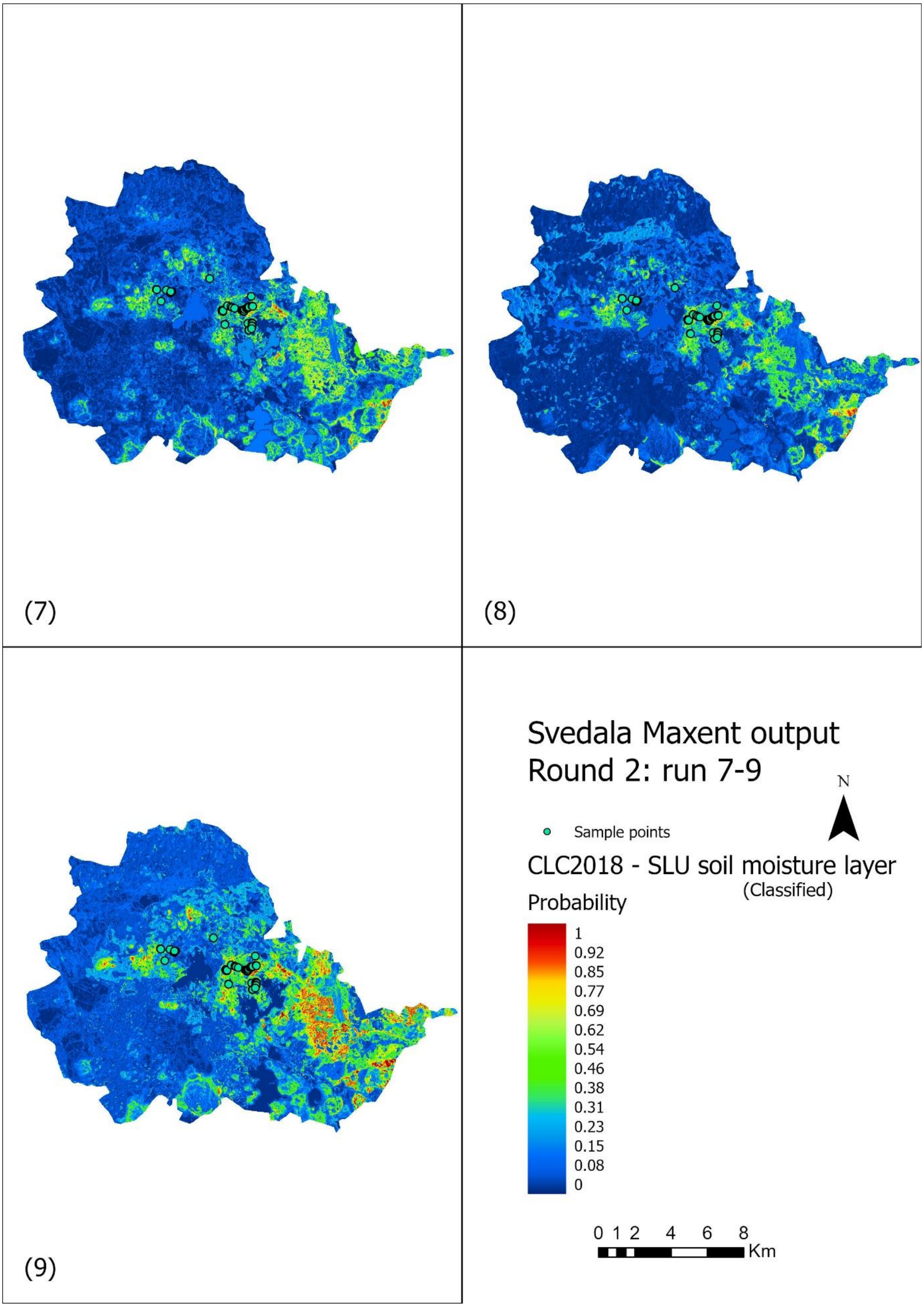

Figure A12: The Maxent probability of occurrence from the second round, run (7), run (8), and run (9), with the environmental factor layer combination of the CLC 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

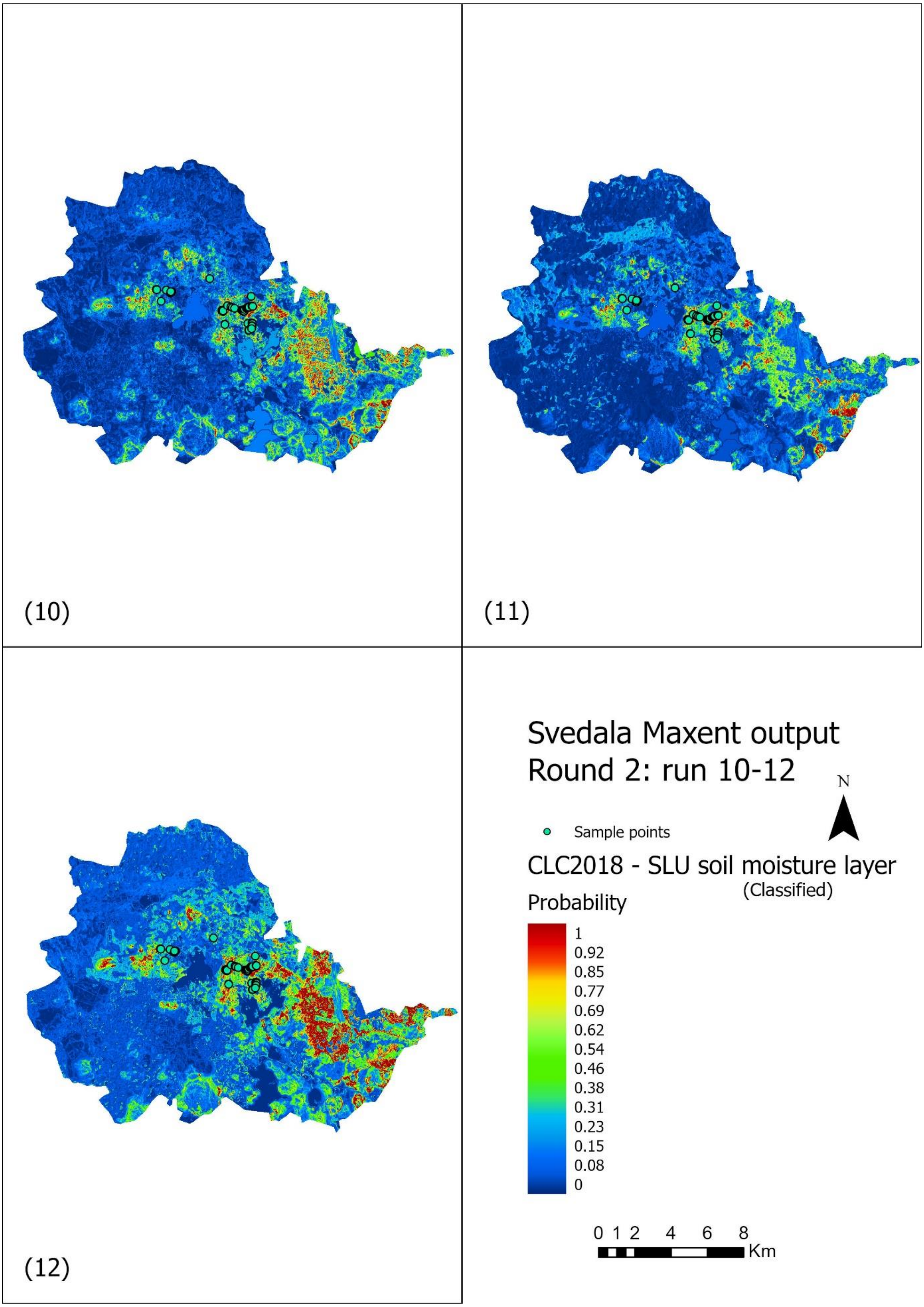

Figure A13: The Maxent probability of occurrence from the second round, run (10), run (11), and run (12), with the environmental factor layer combination of the CLC 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

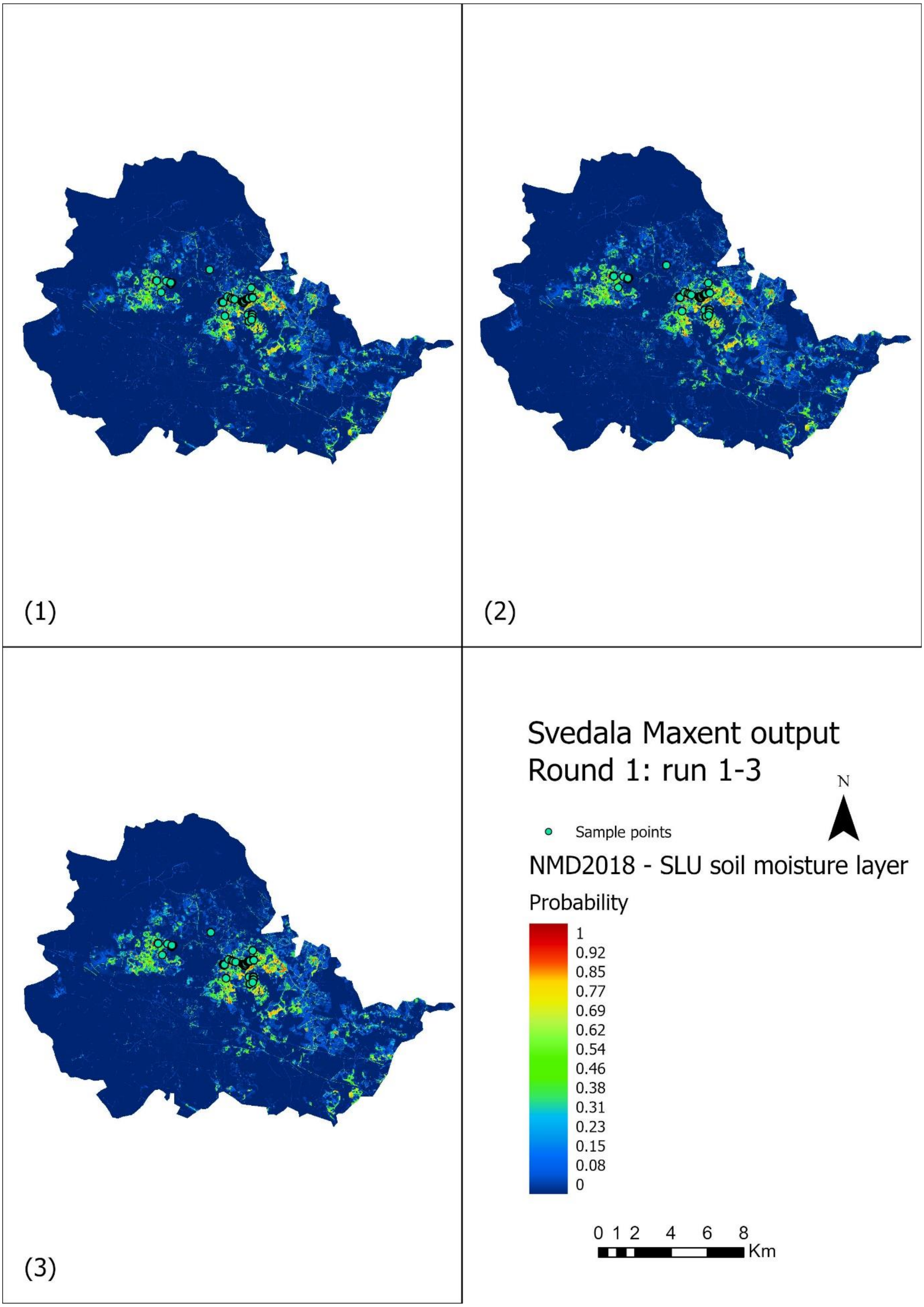

Figure A14: The Maxent probability of occurrence from the first round, run (1), run (2), and run (3), with the environmental factor layer combination of the NMD 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

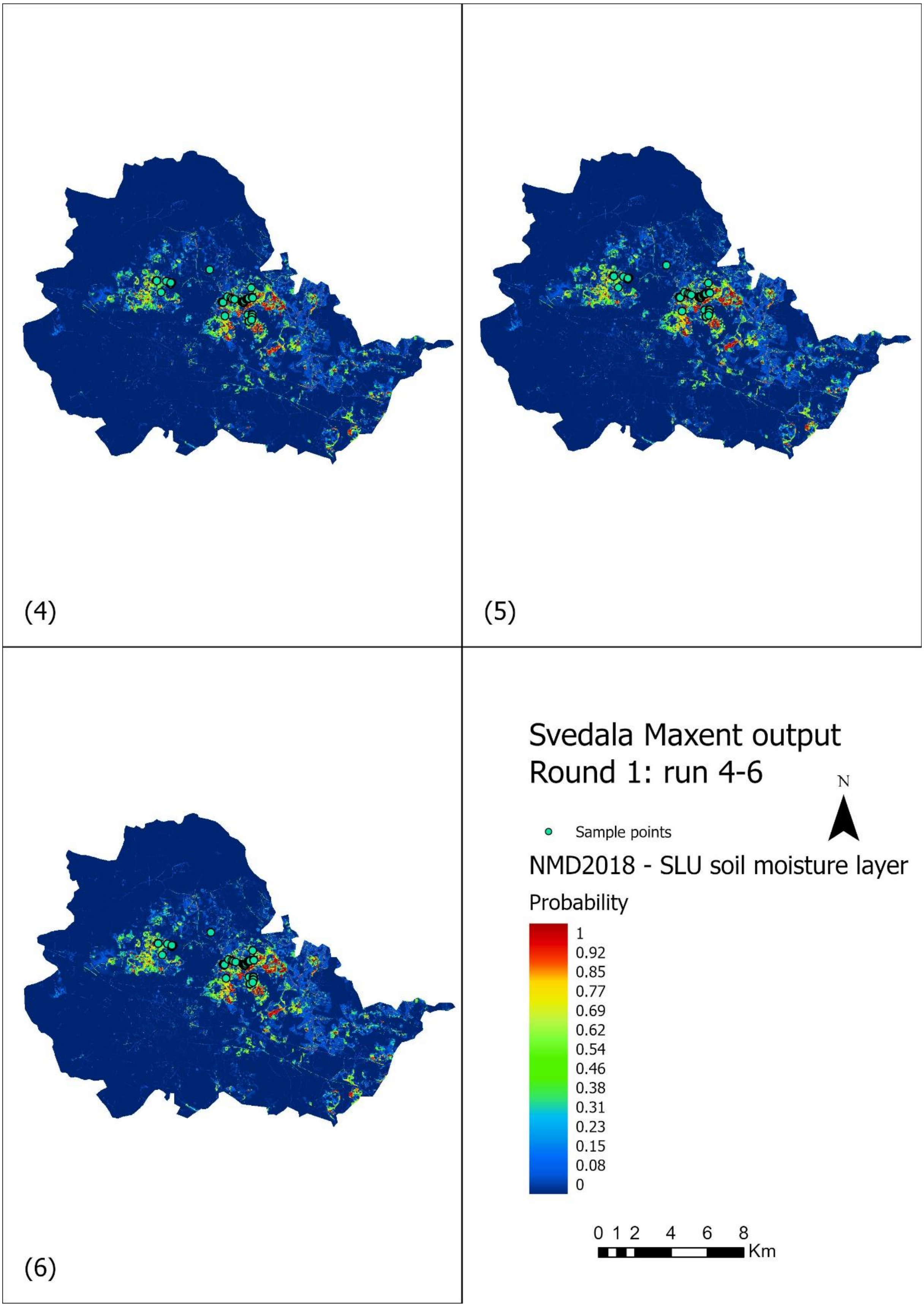

Figure A15: The Maxent probability of occurrence from the first round, run (4), run (5), and run (6), with the environmental factor layer combination of the NMD 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

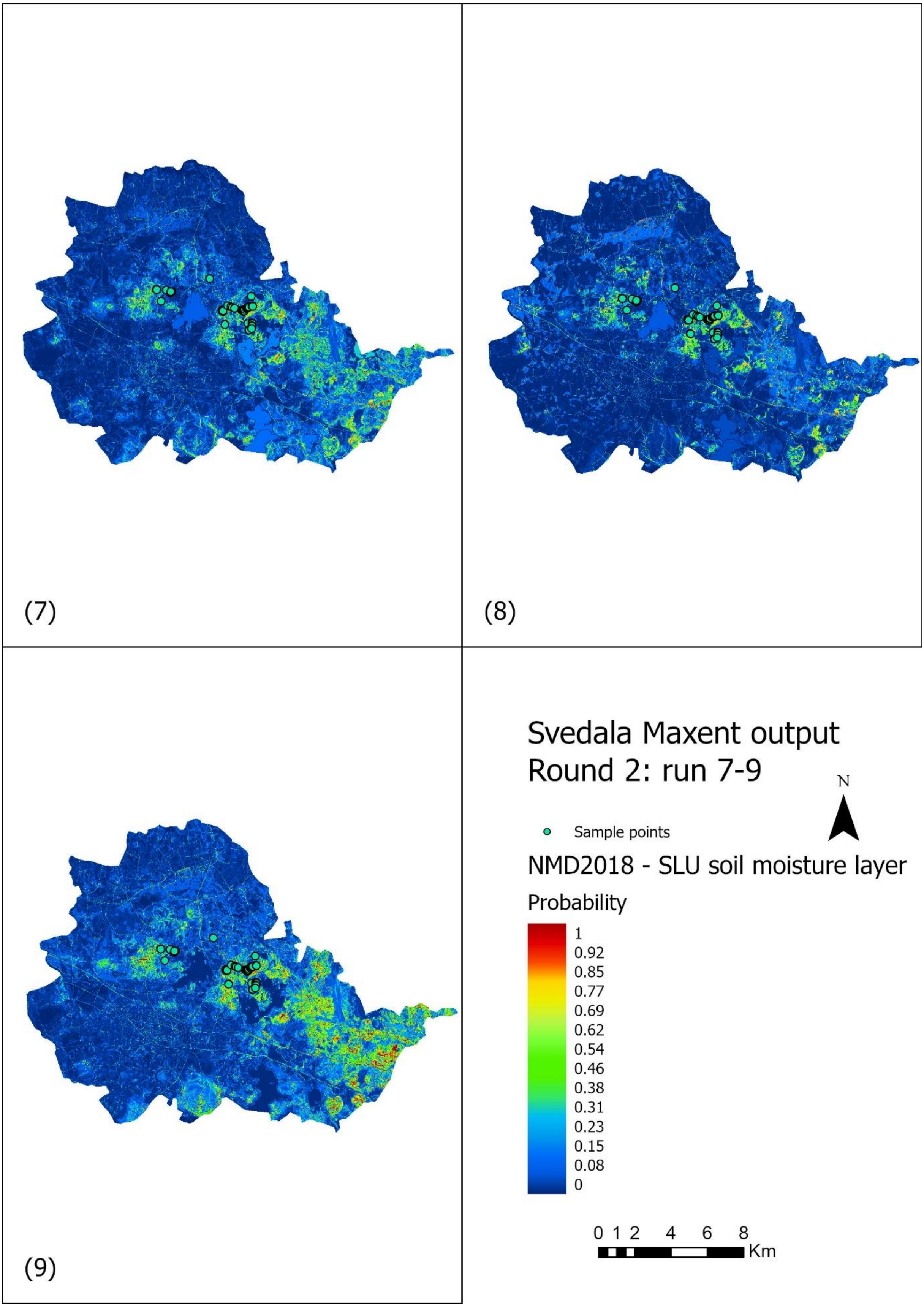

Figure A16: The Maxent probability of occurrence from the second round, run (7), run (8), and run (9), with the environmental factor layer combination of the NMD 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

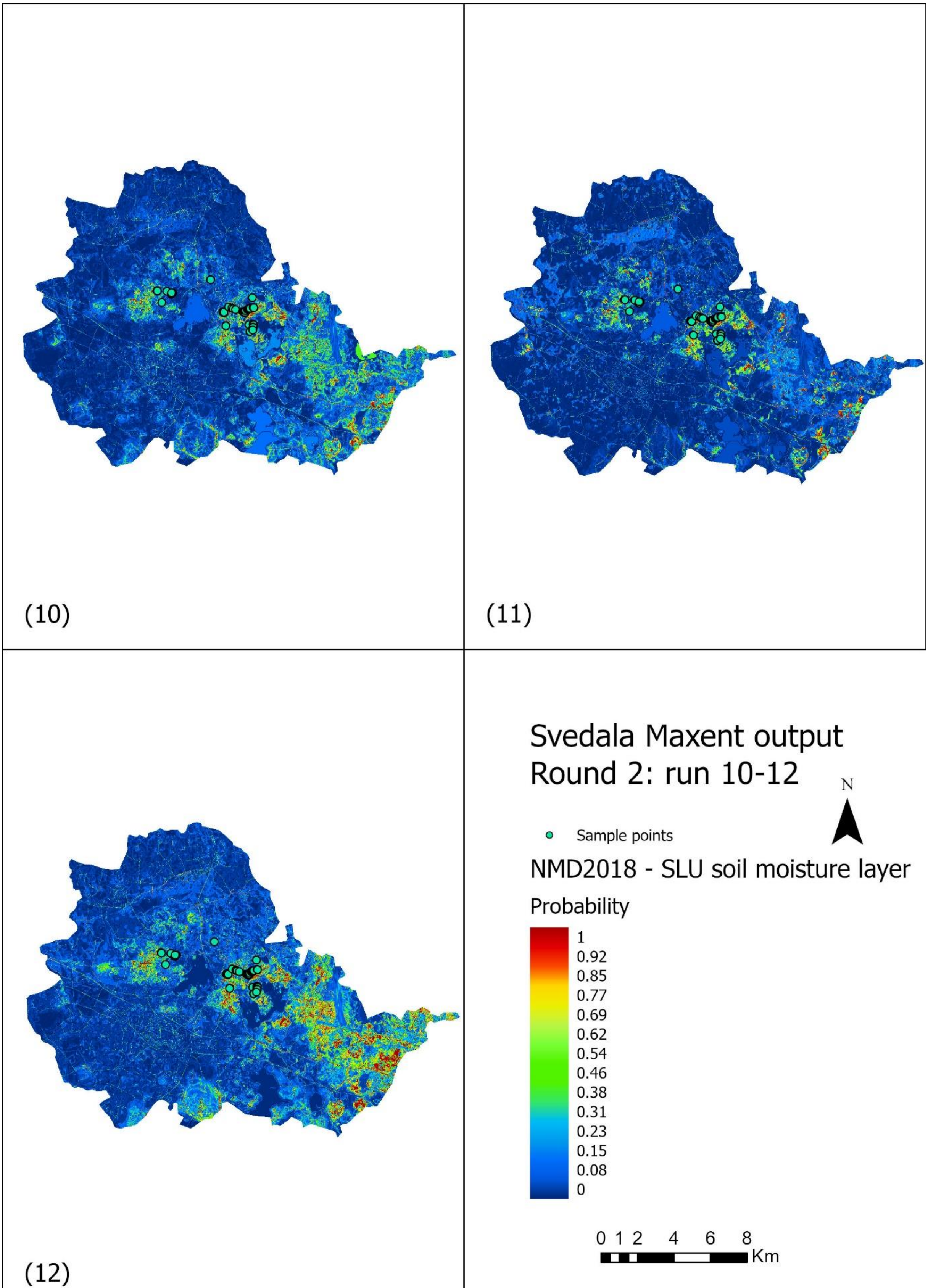

Figure A17: The Maxent probability of occurrence from the second round, run (10), run (11), and run (12), with the environmental factor layer combination of the NMD 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

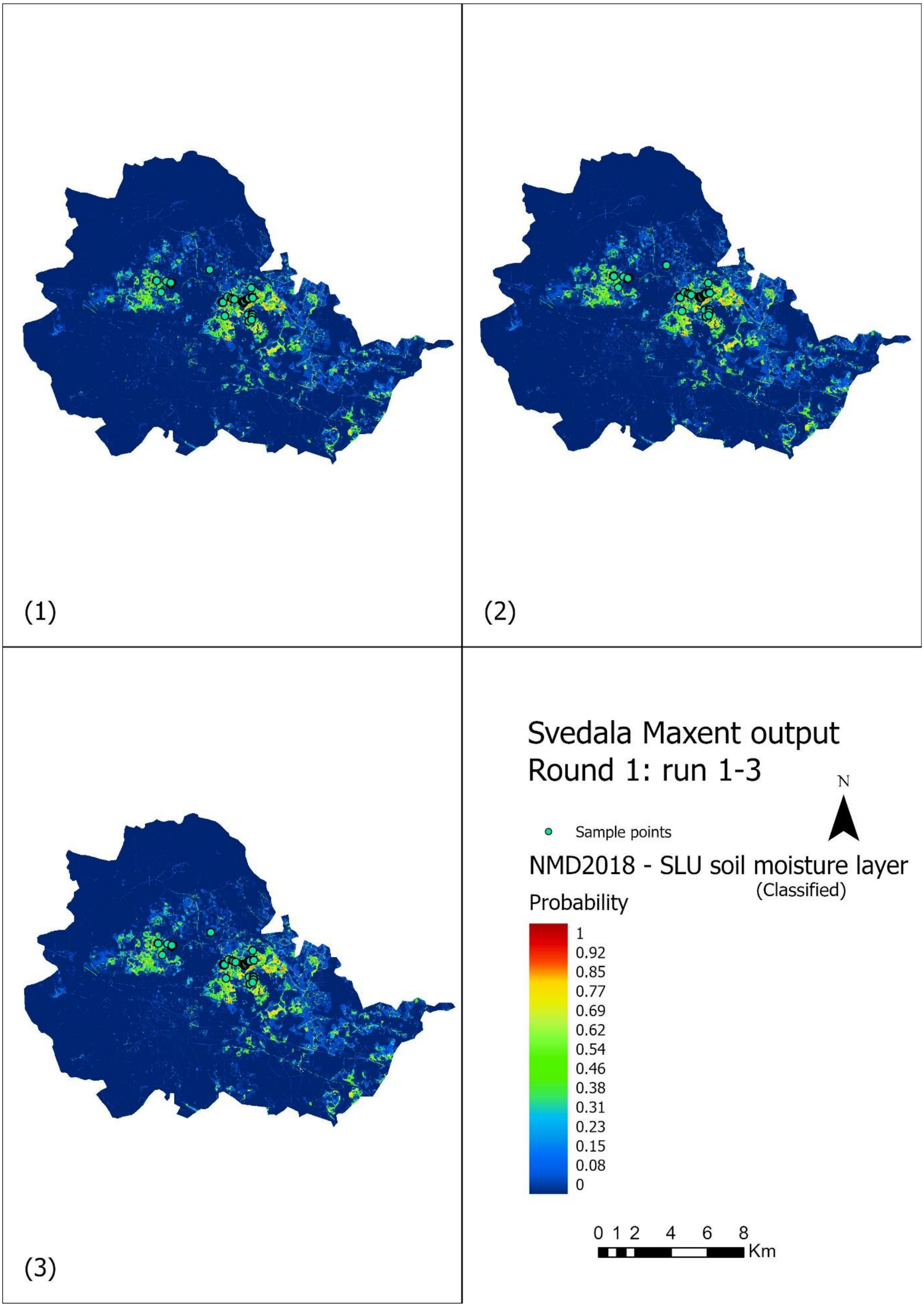

Figure A18: The Maxent probability of occurrence from the first round, run (1), run (2), and run (3), with the environmental factor layer combination of the NMD 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

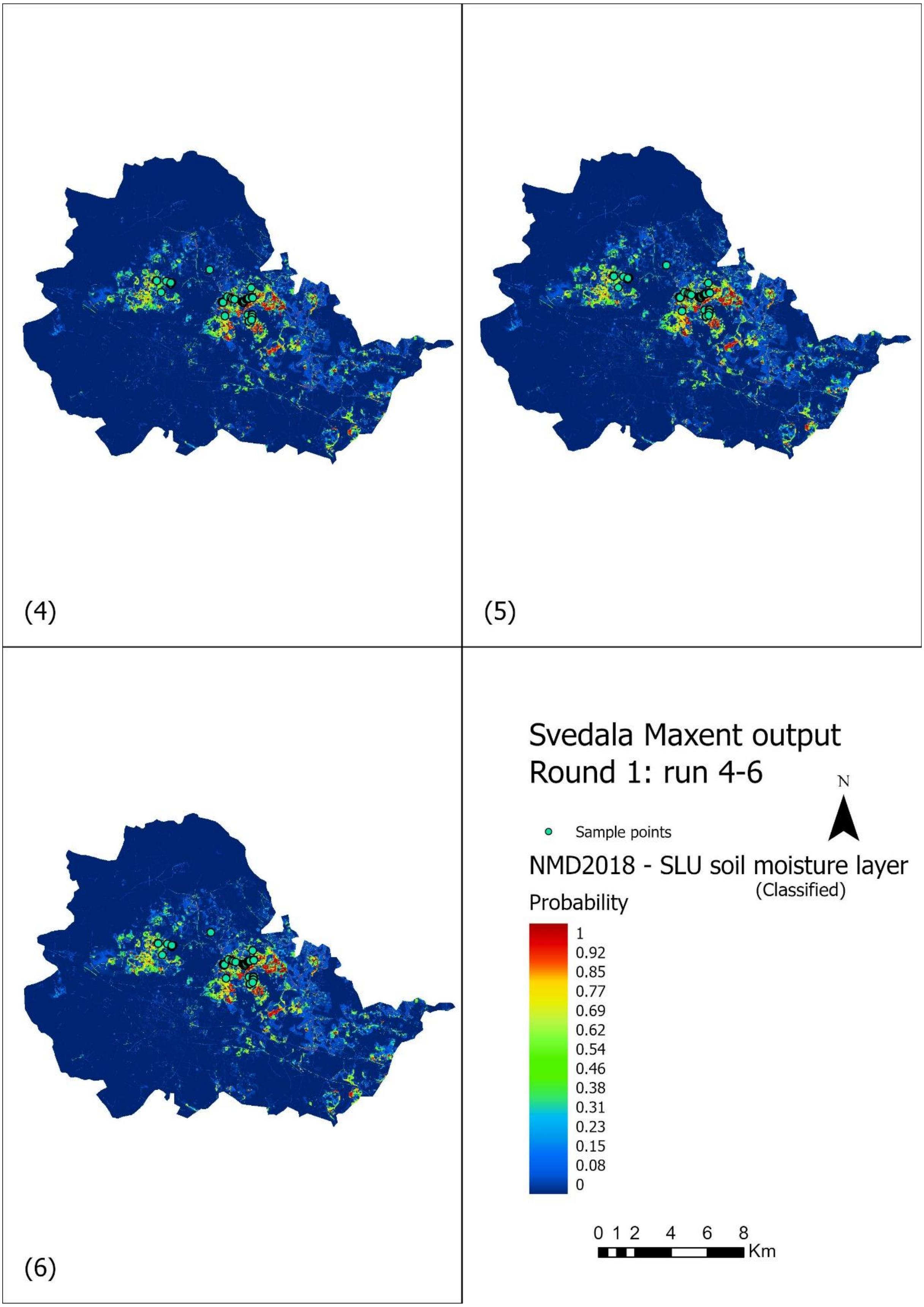

Figure A19: The Maxent probability of occurrence from the second round, run (4), run (5), and run (6), with the environmental factor layer combination of the NMD 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

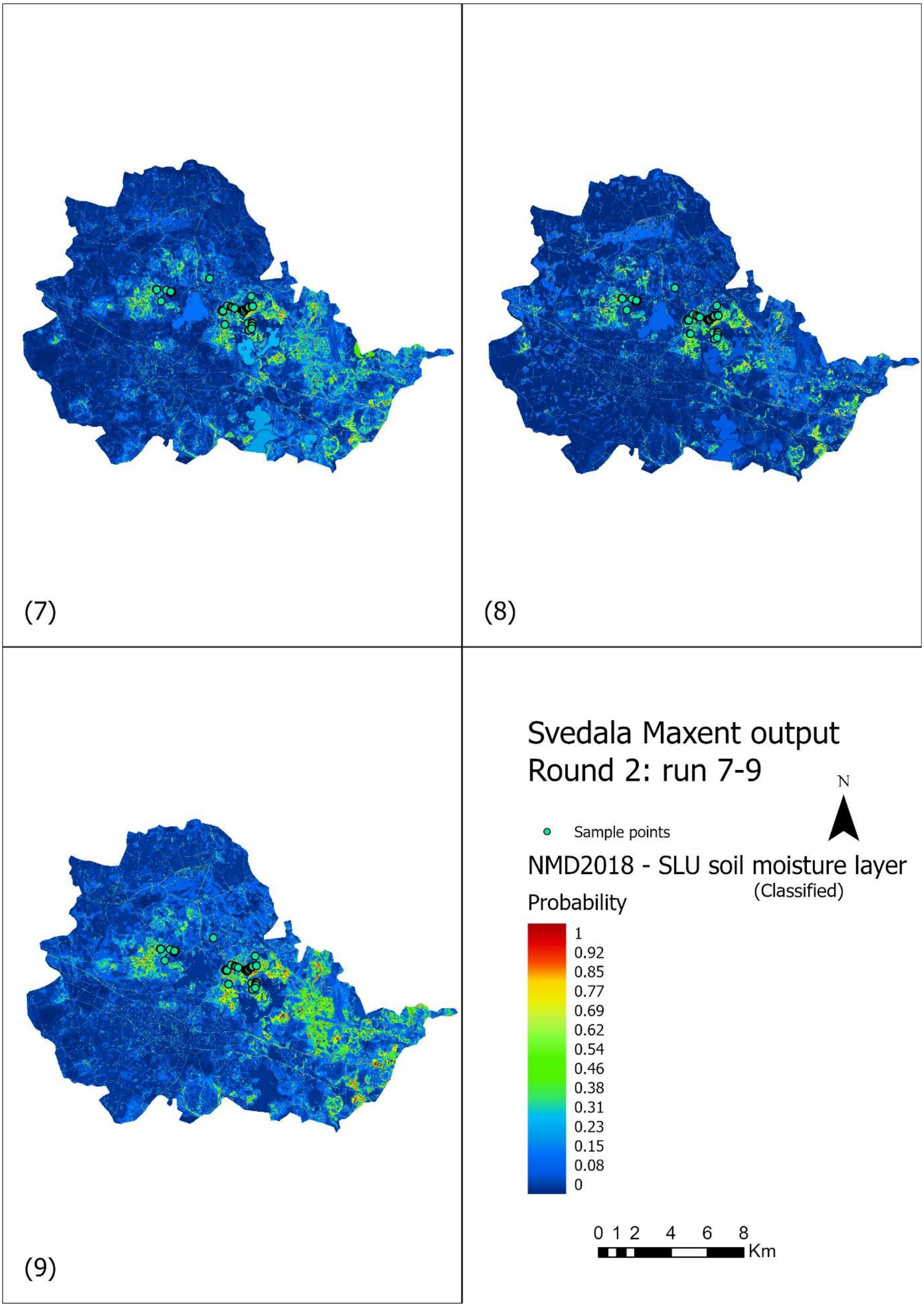

Figure A20: The Maxent probability of occurrence from the second round, run (7), run (8), and run (9), with the environmental factor layer combination of the NMD 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

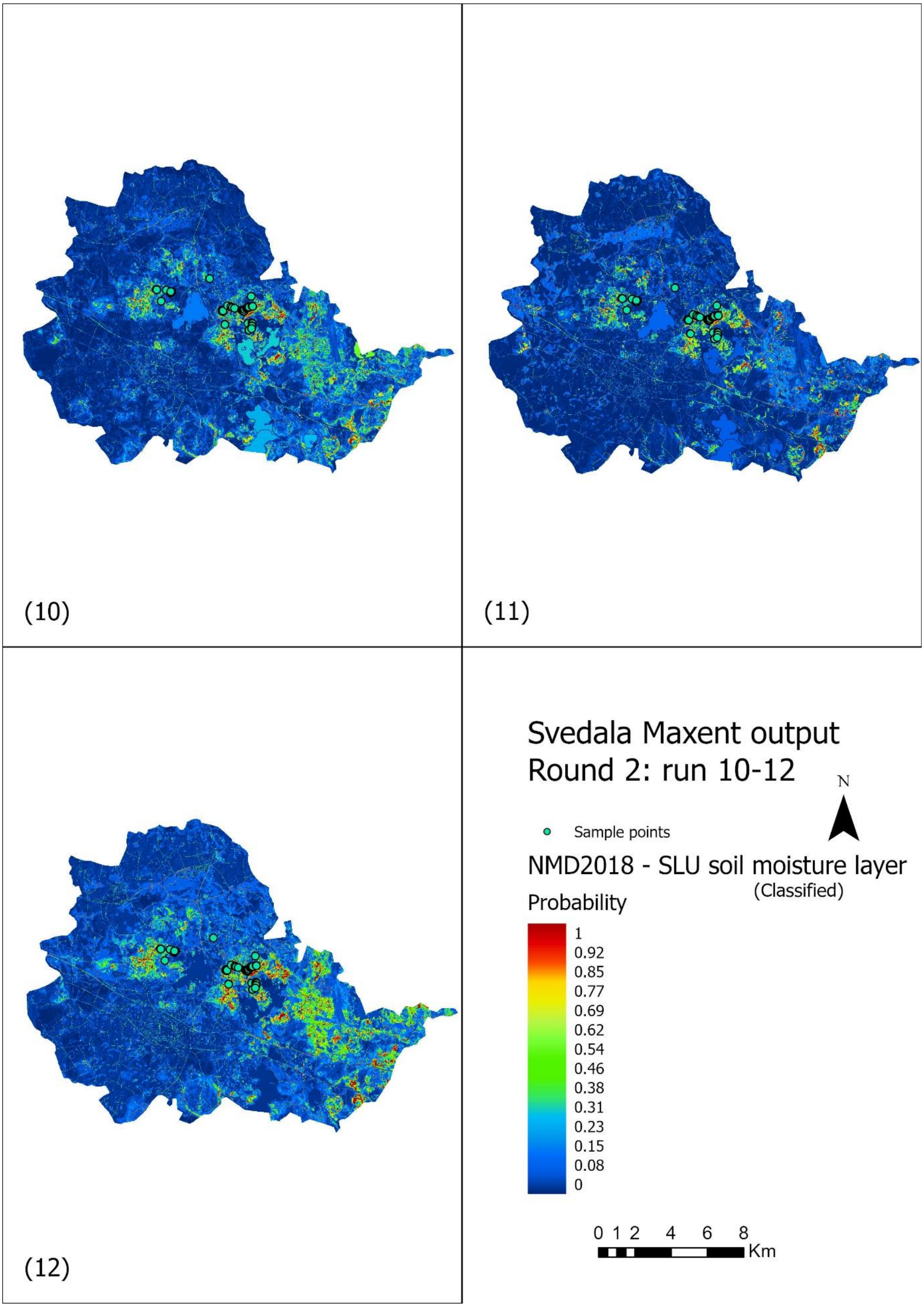

Figure A17: The Maxent probability of occurrence from the second round, run (10), run (11), and run (12), with the environmental factor layer combination of the NMD 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

Table A2: The training and test data AUC values for the runs, 1 to 12 of round 1and 2, for the CLC 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer).*

	Svedala					
	CLC 2018 and SLU soil moisture layer*					
Round 1	1	2	3	4	5	6
Training data (AUC)	0.976	0.974	0.975	0.976	0.974	0.975
Test data (AUC)	0.925	0.942	0.939	0.925	0.942	0.939
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5
Round 2	7	8	9	10	11	12
Training data (AUC)	0.836	0.856	0.841	0.836	0.856	0.841
Test data (AUC)	0.685	0.777	0.766	0.685	0.777	0.766
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5

Table A3: The training and test data AUC values for the runs, 1 to 12 of round 1and 2, for the CLC 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer).*

	Svedala					
	CLC 2018 and SLU classified soil moisture layer*					
Round 1	1	2	3	4	5	6
Training data (AUC)	0.976	0.974	0.975	0.976	0.974	0.975
Test data (AUC)	0.924	0.942	0.939	0.924	0.942	0.939
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5
Round 2	7	8	9	10	11	12
Training data (AUC)	0.827	0.858	0.818	0.827	0.858	0.818
Test data (AUC)	0.683	0.766	0.795	0.683	0.766	0.795
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5

Table A4: The training and test data AUC values for the runs, 1 to 12 of round 1and 2, for the NMD 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer).*

	Svedala					
	NMD 2018 and SLU soil moisture layer*					
Round 1	1	2	3	4	5	6
Training data (AUC)	0.978	0.977	0.977	0.978	0.977	0.977
Test data (AUC)	0.953	0.964	0.962	0.953	0.964	0.962
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5
Round 2	7	8	9	10	11	12
Training data (AUC)	0.876	0.903	0.882	0.876	0.903	0.882
Test data (AUC)	0.777	0.786	0.835	0.777	0.786	0.835
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5

Table A5: The training and test data AUC values for the runs, 1 to 12 of round 1and 2, for the NMD 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer).*

	Svedala					
	NMD 2018 and SLU classified soil moisture layer*					
Round 1	1	2	3	4	5	6
Training data (AUC)	0.978	0.976	0.976	0.978	0.976	0.976
Test data (AUC)	0.953	0.965	0.962	0.953	0.965	0.962
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5
Round 2	7	8	9	10	11	12
Training data (AUC)	0.871	0.904	0.870	0.871	0.904	0.870
Test data (AUC)	0.781	0.783	0.879	0.781	0.783	0.879
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5

Table A6: The percentage of each environmental factor the contributes to building the model, for the CLC 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Contribution	Svedala					
	CLC 2018 and SLU soil moisture layer*					
Round 1 (% Contribution)	1	2	3	4	5	6
CLC 2018	49.4	50.1	50.2	49.4	50.1	50.2
DEM	7.6	6.9	7.3	7.6	6.9	7.3
Slope	1.1	1	1	1.1	1	1
Aspect	1.9	0.9	0.9	1.9	0.9	0.9
Soil type	39.6	40.7	40.2	39.6	40.7	40.2
SLU soil moisture layer	0.3	0.3	0.3	0.3	0.3	0.3
Round 2 (% Contribution)	7	8	9	10	11	12
CLC 2018	22.6	21.4	18.2	22.6	21.4	18.2
DEM	27.6	35.9	35.4	27.6	35.9	35.4
Slope	25.3	1.4	12.4	25.3	1.4	12.4
Aspect	5.4	12.1	1.3	5.4	12.1	1.3
Soil type	13.5	22.4	18.2	13.5	22.4	18.2
SLU soil moisture layer	5.7	6.7	14.5	5.7	6.7	14.5

Table A7: The percentage of each environmental factor the contributes to building the model, for the CLC 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Contribution	Svedala					
	CLC 2018 and SLU classified soil moisture layer*					
Round 1 (% Contribution)	1	2	3	4	5	6
CLC 2018	49.9	51.3	50.7	49.9	51.3	50.7
DEM	7.5	5.2	6.4	7.5	5.2	6.4
Slope	0.8	0.7	0.9	0.8	0.7	0.9
Aspect	1.3	0.4	0.8	1.3	0.4	0.8
Soil type	39.4	41.4	40.3	39.4	41.4	40.3
SLU soil moisture layer	1.1	0.9	0.9	1.1	0.9	0.9
Round 2 (% Contribution)	7	8	9	10	11	12
CLC 2018	23.9	22.3	21.3	23.9	22.3	21.3
DEM	28.5	35.5	37.8	28.5	35.5	37.8
Slope	26.8	2.4	12.3	26.8	2.4	12.3
Aspect	6	13.5	2.3	6	13.5	2.3
Soil type	14.3	21.9	21.5	14.3	21.9	21.5
SLU soil moisture layer	0.6	4.5	4.8	0.6	4.5	4.8

Table A8: The percentage of each environmental factor the contributes to building the model, for the NMD 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Contribution	Svedala					
	NMD 2018 and SLU soil moisture layer*					
Round 1 (% Contribution)	1	2	3	4	5	6
NMD 2018	47.6	50.1	48.5	47.6	50.1	48.5
DEM	7.5	5.6	6.3	7.5	5.6	6.3
Slope	1	0.8	0.9	1	0.8	0.9
Aspect	1.5	1.1	1.2	1.5	1.1	1.2
Soil type	42.3	42.4	43	42.3	42.4	43
SLU soil moisture layer	0.1	0	0	0.1	0	0
Round 2 (% Contribution)	7	8	9	10	11	12
NMD 2018	46.1	52.5	44.1	46.1	52.5	44.1
DEM	14.7	19.7	24.9	14.7	19.7	24.9
Slope	21.7	1.7	9.6	21.7	1.7	9.6
Aspect	4.3	9.6	1	4.3	9.6	1
Soil type	10.6	15.8	11.6	10.6	15.8	11.6
SLU soil moisture layer	2.6	0.6	8.8	2.6	0.6	8.8

Table A9: The percentage of each environmental factor the contributes to building the model, for the NMD 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Contribution	Svedala					
	NMD 2018 and SLU classified soil moisture layer*					
Round 1 (% Contribution)	1	2	3	4	5	6
NMD 2018	47.7	50.6	49.1	47.7	50.6	49.1
DEM	7.1	5.1	6.1	7.1	5.1	6.1
Slope	0.8	0.3	0.4	0.8	0.3	0.4
Aspect	1.2	0.7	0.6	1.2	0.7	0.6
Soil type	42.7	42.7	43.3	42.7	42.7	43.3
SLU soil moisture layer	0.5	0.6	0.5	0.5	0.6	0.5
Round 2 (% Contribution)	7	8	9	10	11	12
NMD 2018	47.5	53	49.9	47.5	53	49.9
DEM	14.7	19.6	23.9	14.7	19.6	23.9
Slope	21.9	1.6	9	21.9	1.6	9
Aspect	5	9.7	2.3	5	9.7	2.3
Soil type	10.9	15.9	14	10.9	15.9	14
SLU soil moisture layer	0	0.2	1	0	0.2	1

Table A10: The percentage of permutation importance of each environmental factor in the established model, for the CLC 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Permutation importance	Svedala					
	CLC 2018 and SLU soil moisture layer*					
Round 1 (% Permutation importance)	1	2	3	4	5	6
CLC 2018	55.4	74.1	68.3	55.4	74.1	68.3
DEM	13	7.4	10	13	7.4	10
Slope	0.7	0.2	2.2	0.7	0.2	2.2
Aspect	4.9	2.4	3.1	4.9	2.4	3.1
Soil type	24.6	14.6	16.4	24.6	40.7	16.4
SLU soil moisture layer	1.3	1.2	0	1.3	1.2	0
Round 2 (% Permutation importance)	7	8	9	10	11	12
CLC 2018	17.1	35.7	9.4	17.1	35.7	9.4
DEM	25.7	23.6	33.7	25.7	23.6	33.7
Slope	27	5.6	4	27	5.6	4
Aspect	4.1	14.1	6.4	4.1	14.1	6.4
Soil type	10.2	21	23.4	10.2	21	23.4
SLU soil moisture layer	15.8	0	23.1	15.8	0	23.1

Table A11: The percentage of permutation importance of each environmental factor in the established model, for the CLC 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Permutation importance	Svedala					
	CLC 2018 and SLU classified soil moisture layer*					
Round 1 (% Permutation importance)	1	2	3	4	5	6
CLC 2018	58.1	72.7	58.5	58.1	72.7	58.5
DEM	13.5	8.6	13.2	13.5	8.6	13.2
Slope	0.6	0.6	0	0.6	0.6	0
Aspect	1.5	2	0.3	1.5	2	0.3
Soil type	24.7	14.6	24.7	24.7	14.6	24.7
SLU soil moisture layer	1.6	1.4	3.3	1.6	1.4	3.3
Round 2 (% Permutation importance)	7	8	9	10	11	12
CLC 2018	25.2	34.7	12.9	25.2	34.7	12.9
DEM	28.9	21.4	30	28.9	21.4	30
Slope	20.1	5.1	4.6	20.1	5.1	4.6
Aspect	7.3	13.5	8.2	7.3	13.5	8.2
Soil type	13.7	20.8	22.5	13.7	20.8	22.5
SLU soil moisture layer	4.9	4.5	21.8	4.9	4.5	21.8

Table A12: The percentage of permutation importance of each environmental factor in the established model, for the NMD 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Permutation importance	Svedala					
	NMD 2018 and SLU soil moisture layer*					
Round 1 (% Permutation importance)	1	2	3	4	5	6
NMD 2018	57.1	48	51.5	57.1	48	51.5
DEM	8	5.3	11.5	8	5.3	11.5
Slope	1.1	0.8	1.8	1.1	0.8	1.8
Aspect	4.4	1.4	2	4.4	1.4	2
Soil type	29.4	44.4	33	29.4	44.4	33
SLU soil moisture layer	0	0.1	0.1	0	0.1	0.1
Round 2 (% Permutation importance)	7	8	9	10	11	12
NMD 2018	53.2	64	28.8	53.2	64	28.8
DEM	13.7	8.8	33.8	13.7	8.8	33.8
Slope	19.2	2	2.6	19.2	2	2.6
Aspect	3.1	5.1	3.7	3.1	5.1	3.7
Soil type	6.7	19.4	19.3	6.7	19.4	19.3
SLU soil moisture layer	4.2	0.8	11.9	4.2	0.8	11.9

Table A13: The percentage of permutation importance of each environmental factor in the established model, for the NMD 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Permutation importance	Svedala					
	NMD 2018 and SLU classified soil moisture layer*					
Round 1 (% Permutation importance)	1	2	3	4	5	6
NMD 2018	53.7	51.8	52.6	53.7	51.8	52.6
DEM	9.5	6.2	11.7	9.5	6.2	11.7
Slope	0.1	0.2	0	0.1	0.2	0
Aspect	1.3	1.5	0.4	1.3	1.5	0.4
Soil type	34.1	39.3	33.2	34.1	39.3	33.2
SLU soil moisture layer	1.3	1	2.1	1.3	1	2.1
Round 2 (% Permutation importance)	7	8	9	10	11	12
NMD 2018	65.8	65.7	42.5	65.8	65.7	42.5
DEM	14.2	8.4	27.9	14.2	8.4	27.9
Slope	9	1.1	1.7	9	1.1	1.7
Aspect	3.6	5.3	5.2	3.6	5.3	5.2
Soil type	7.2	19.2	16.2	7.2	19.2	16.2
SLU soil moisture layer	0.2	0.2	6.6	0.2	0.2	6.6

Table A14: How each environmental factor layer was ranked based on their respective percentage values for the CLC 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Contribution	Svedala					
	CLC 2018 and SLU soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
CLC 2018	1	1	1	1	1	1
DEM	3	3	3	3	3	3
Slope	5	4	4	5	4	4
Aspect	4	5	5	4	5	5
Soil type	2	2	2	2	2	2
SLU soil moisture layer	6	6	6	6	6	6
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
CLC 2018	3	3	2	3	3	2
DEM	1	1	1	1	1	1
Slope	2	6	4	2	6	4
Aspect	6	4	5	6	4	5
Soil type	4	2	2	4	2	2
SLU soil moisture layer	5	5	3	5	5	3

Table A15: How each environmental factor layer was ranked based on their respective percentage values for the CLC 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Contribution	Svedala					
	CLC 2018 and SLU classified soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
CLC 2018	1	1	1	1	1	1
DEM	3	3	3	3	3	3
Slope	6	5	4	6	5	4
Aspect	4	6	5	4	6	5
Soil type	2	2	2	2	2	2
SLU soil moisture layer	5	4	4	5	4	4
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
CLC 2018	3	2	3	3	2	3
DEM	1	1	1	1	1	1
Slope	2	6	4	2	6	4
Aspect	5	4	6	5	4	6
Soil type	4	3	2	4	3	2
SLU soil moisture layer	6	5	5	6	5	5

Table A16: How each environmental factor layer was ranked based on their respective percentage values for the NMD 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Contribution	Svedala					
	NMD 2018 and SLU soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
NMD 2018	1	1	1	1	1	1
DEM	3	3	3	3	3	3
Slope	5	5	5	5	5	5
Aspect	4	4	4	4	4	4
Soil type	2	2	2	2	2	2
SLU soil moisture layer	6	6	6	6	6	6
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
NMD 2018	1	1	1	1	1	1
DEM	3	2	2	3	2	2
Slope	2	5	4	2	5	4
Aspect	5	4	6	5	4	6
Soil type	4	3	3	4	3	3
SLU soil moisture layer	6	6	5	6	6	5

Table A17: How each environmental factor layer was ranked based on their respective percentage values for the NMD 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Contribution	Svedala					
	NMD 2018 and SLU classified soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
NMD 2018	1	1	1	1	1	1
DEM	3	3	3	3	3	3
Slope	5	6	6	5	6	6
Aspect	4	4	4	4	4	4
Soil type	2	2	2	2	2	2
SLU soil moisture layer	6	5	5	6	5	5
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
NMD 2018	1	1	1	1	1	1
DEM	3	2	2	3	2	2
Slope	2	5	4	2	5	4
Aspect	5	4	5	5	4	5
Soil type	4	3	3	4	3	3
SLU soil moisture layer	6	6	6	6	6	6

Table A18: How each environmental factor layer was ranked based on their respective percentage values for the CLC 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Permutation importance	Svedala					
	CLC 2018 and SLU soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
CLC 2018	1	1	1	1	1	1
DEM	3	3	3	3	3	3
Slope	6	6	5	6	6	5
Aspect	4	4	4	4	4	4
Soil type	2	2	2	2	2	2
SLU soil moisture layer	5	5	6	5	5	6
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
CLC 2018	3	1	4	3	1	4
DEM	2	2	1	2	2	1
Slope	1	5	6	1	5	6
Aspect	6	4	5	6	4	5
Soil type	5	3	2	5	3	2
SLU soil moisture layer	4	6	3	4	6	3

Table A19: How each environmental factor layer was ranked based on their respective percentage values for the CLC 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Permutation importance	Svedala					
	CLC 2018 and SLU classified soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
CLC 2018	1	1	1	1	1	1
DEM	3	3	3	3	3	3
Slope	6	6	6	6	6	6
Aspect	5	4	5	5	4	5
Soil type	2	2	2	2	2	2
SLU soil moisture layer	4	5	4	4	5	4
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
CLC 2018	2	1	4	2	1	4
DEM	1	2	1	1	2	1
Slope	3	5	6	3	5	6
Aspect	5	4	5	5	4	5
Soil type	4	3	2	4	3	2
SLU soil moisture layer	6	6	3	6	6	3

Table A20: How each environmental factor layer was ranked based on their respective percentage values for the NMD 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Permutation importance	Svedala					
	NMD 2018 and SLU soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
NMD 2018	1	1	1	1	1	1
DEM	3	3	3	3	3	3
Slope	5	5	5	5	5	5
Aspect	4	4	4	4	4	4
Soil type	2	2	2	2	2	2
SLU soil moisture layer	6	6	6	6	6	6
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
NMD 2018	1	1	2	1	1	2
DEM	3	3	1	3	3	1
Slope	2	5	6	2	5	6
Aspect	6	4	5	6	4	5
Soil type	4	2	3	4	2	3
SLU soil moisture layer	5	6	4	5	6	4

Table A21: How each environmental factor layer was ranked based on their respective percentage values for the NMD 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Permutation importance	Svedala					
	NMD 2018 and SLU classified soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
NMD 2018	1	1	1	1	1	1
DEM	3	3	3	3	3	3
Slope	5	6	6	5	6	6
Aspect	4	4	5	4	4	5
Soil type	2	2	2	2	2	2
SLU soil moisture layer	4	5	4	4	5	4
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
NMD 2018	1	1	1	1	1	1
DEM	2	3	2	2	3	2
Slope	3	5	6	3	5	6
Aspect	5	4	5	5	4	5
Soil type	4	2	3	4	2	3
SLU soil moisture layer	6	6	4	6	6	4

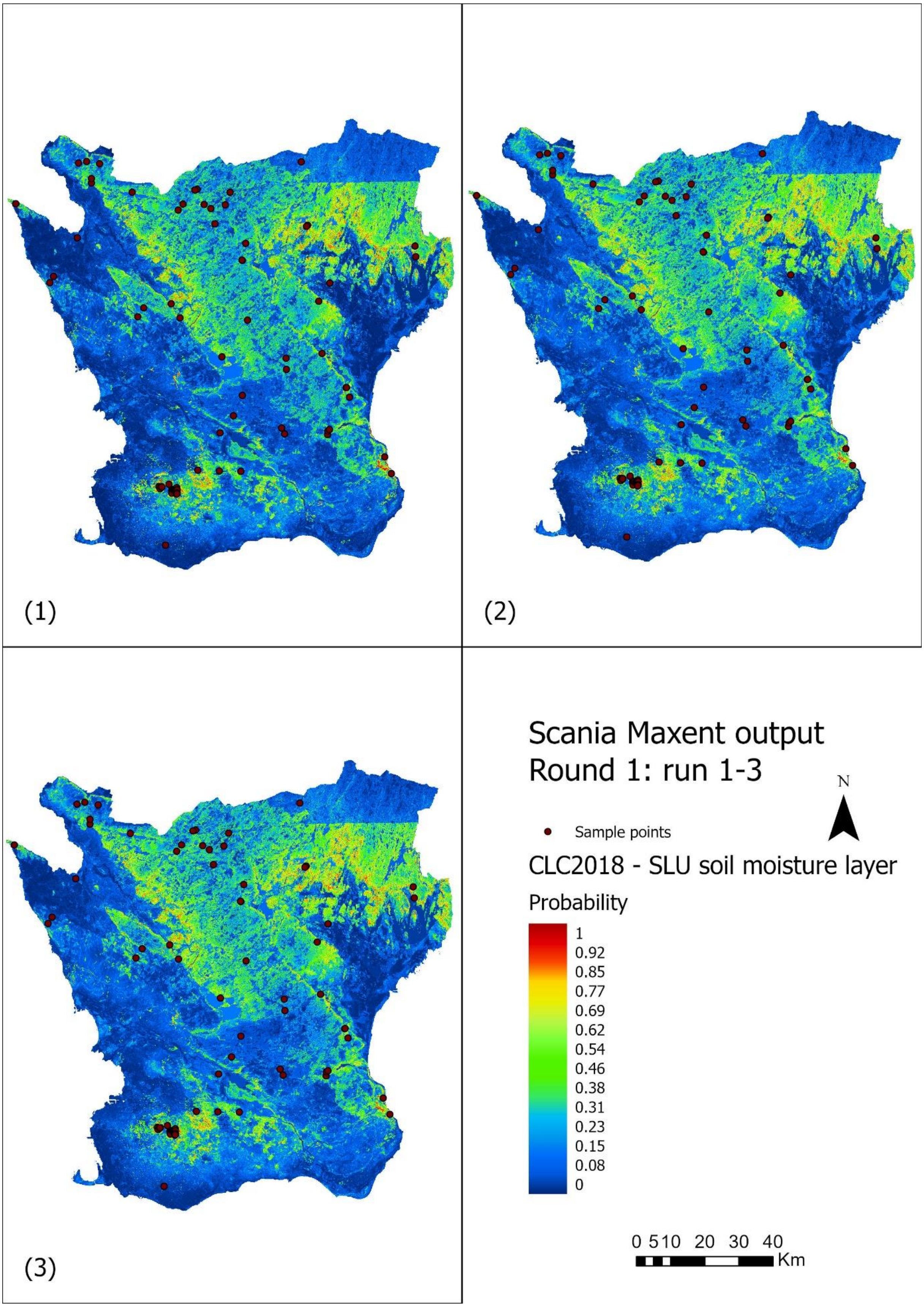

Figure A18: The Maxent probability of occurrence from the first round, run (1), run (2), and run (3), with the environmental factor layer combination of the CLC 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

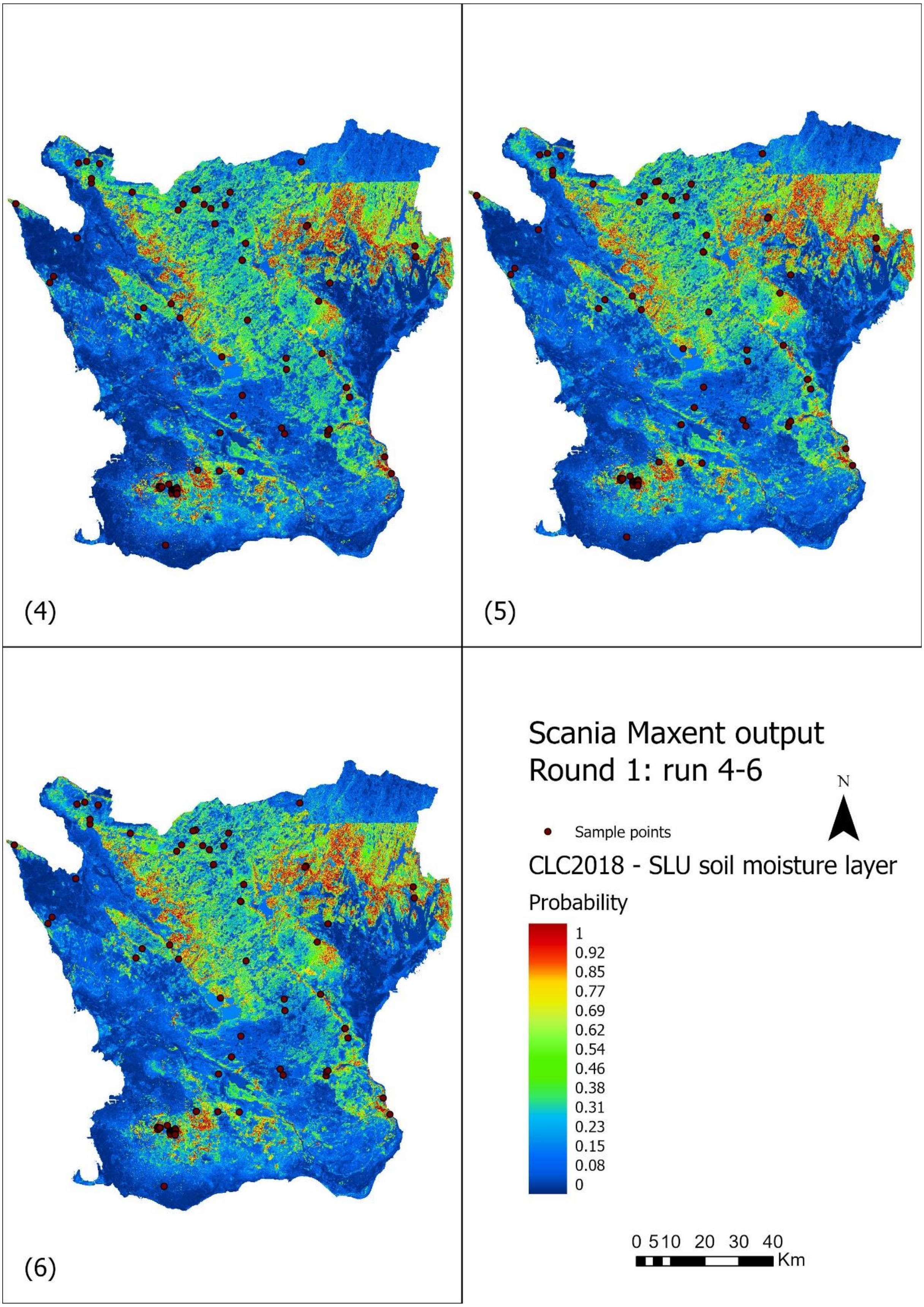

Figure A23: The Maxent probability of occurrence from the first round, run (4), run (5), and run (6), with the environmental factor layer combination of the CLC 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

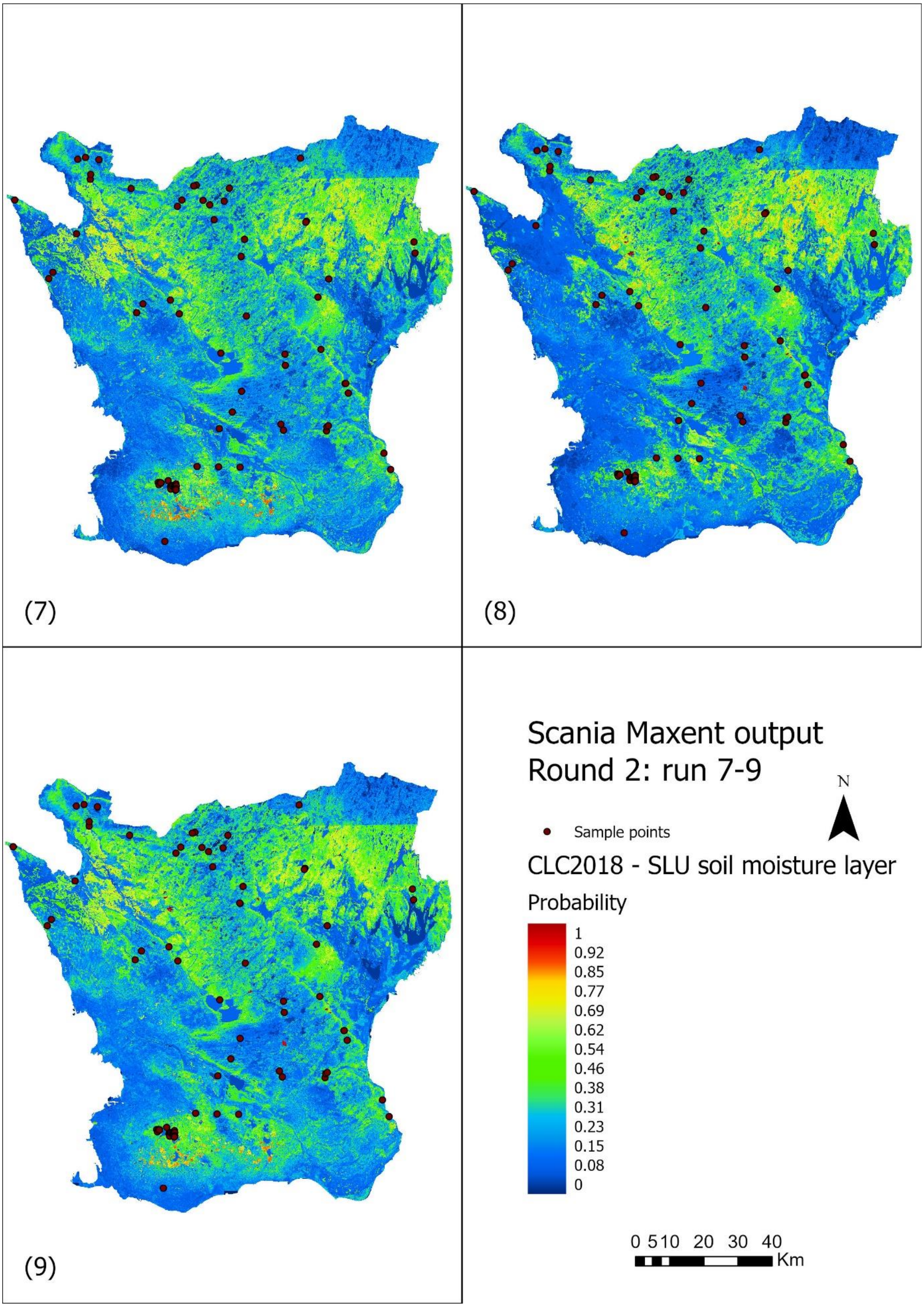

Figure A24: The Maxent probability of occurrence from the second round, run (7), run (8), and run (9), with the environmental factor layer combination of the CLC 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

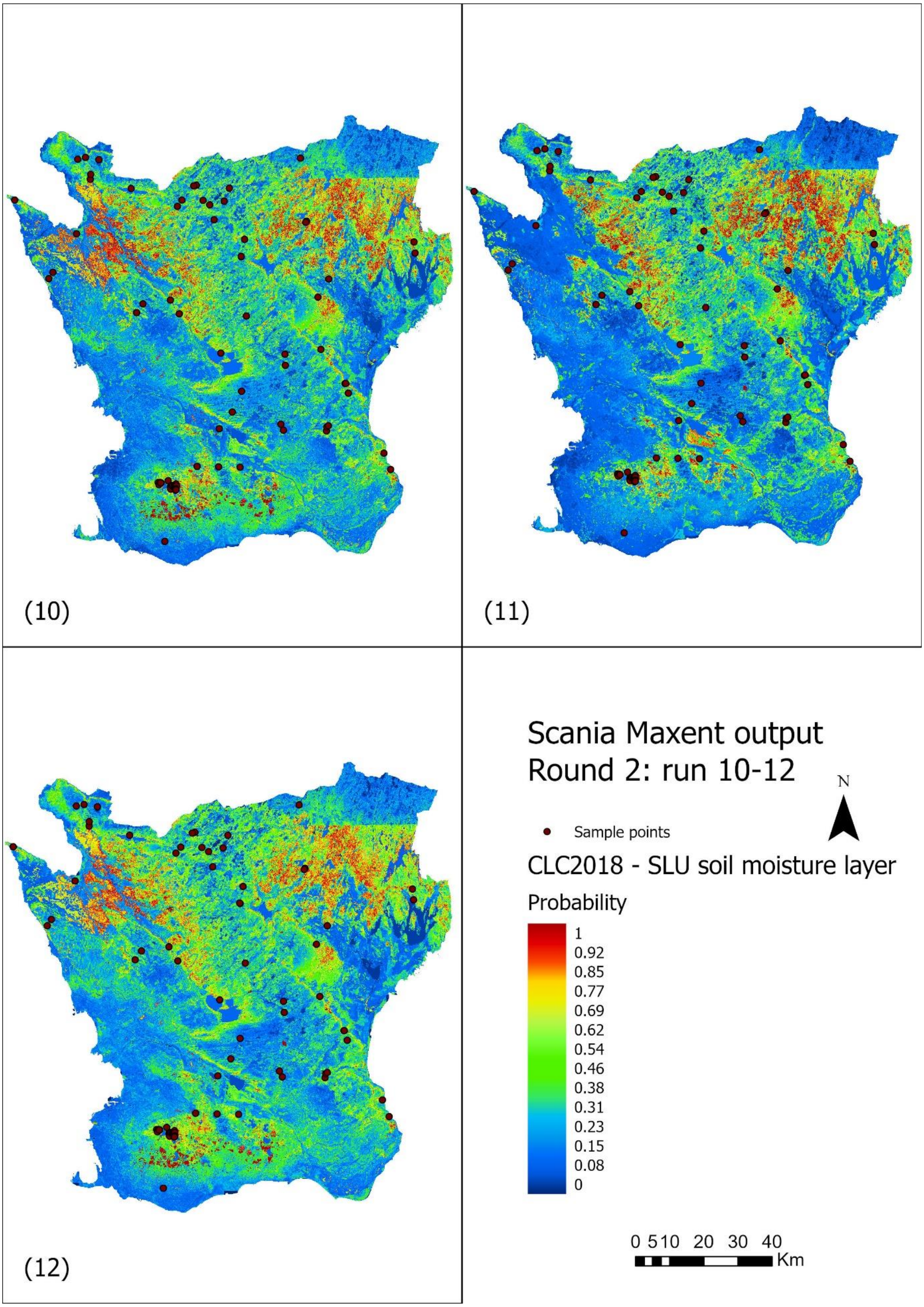

Figure A25: The Maxent probability of occurrence from the second round, run (10), run (11), and run (12), with the environmental factor layer combination of the CLC 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

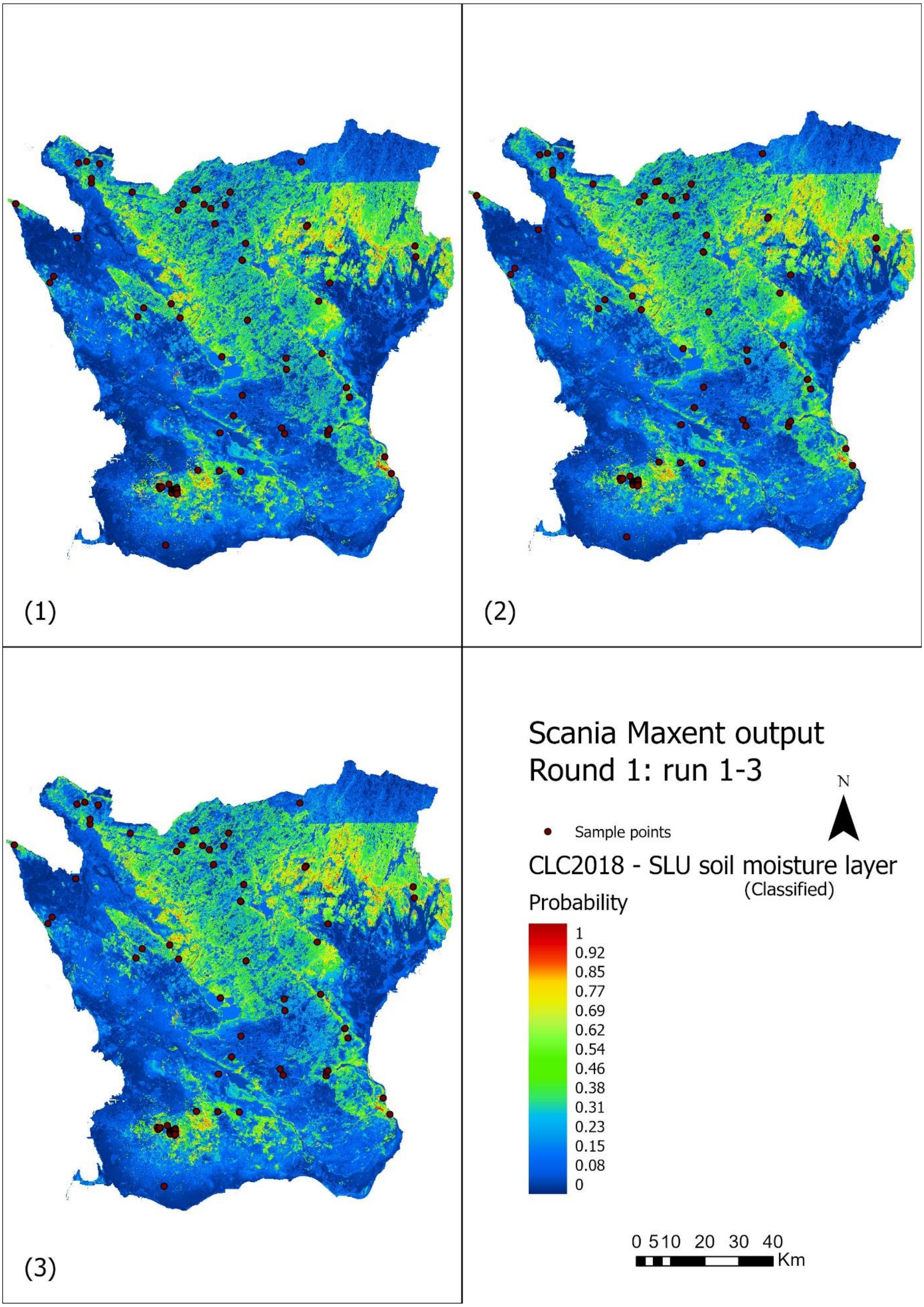

Figure A26: The Maxent probability of occurrence from the first round, run (1), run (2), and run (3), with the environmental factor layer combination of the CLC 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

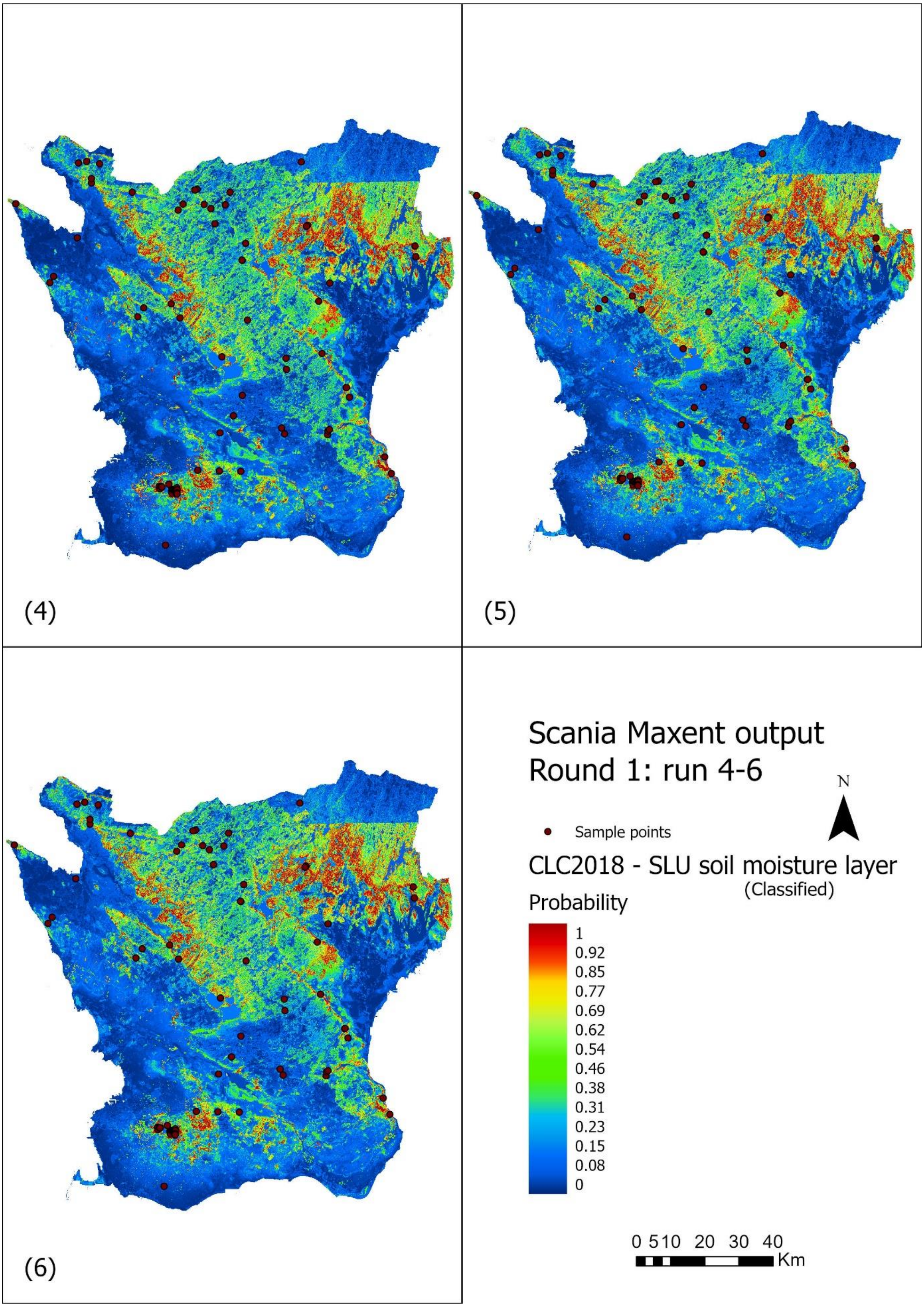

Figure A27: The Maxent probability of occurrence from the second round, run (4), run (5), and run (6), with the environmental factor layer combination of the CLC 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

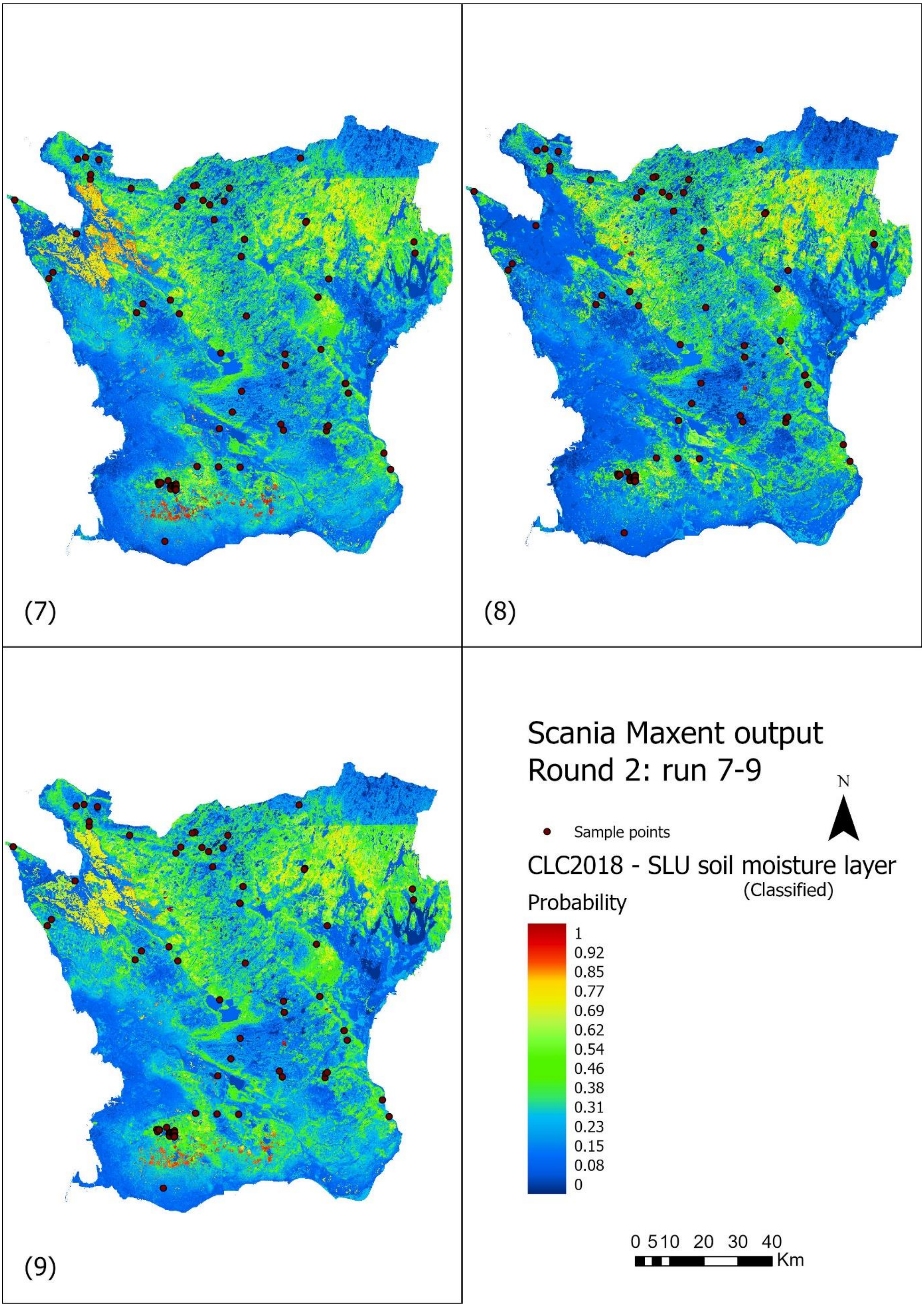

Figure A28: The Maxent probability of occurrence from the second round, run (7), run (8), and run (9), with the environmental factor layer combination of the CLC 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

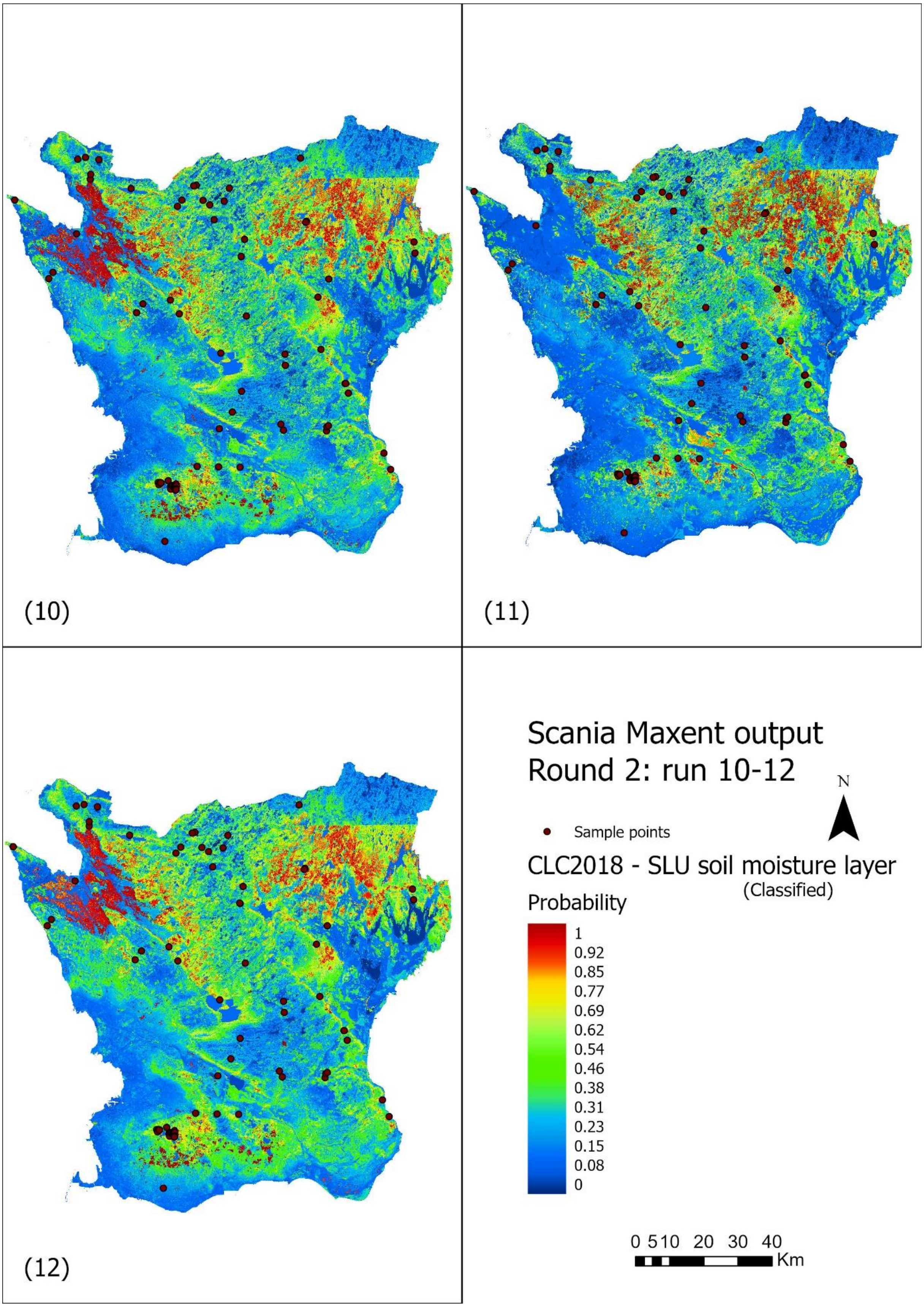

Figure A29: The Maxent probability of occurrence from the second round, run (10), run (11), and run (12), with the environmental factor layer combination of the CLC 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

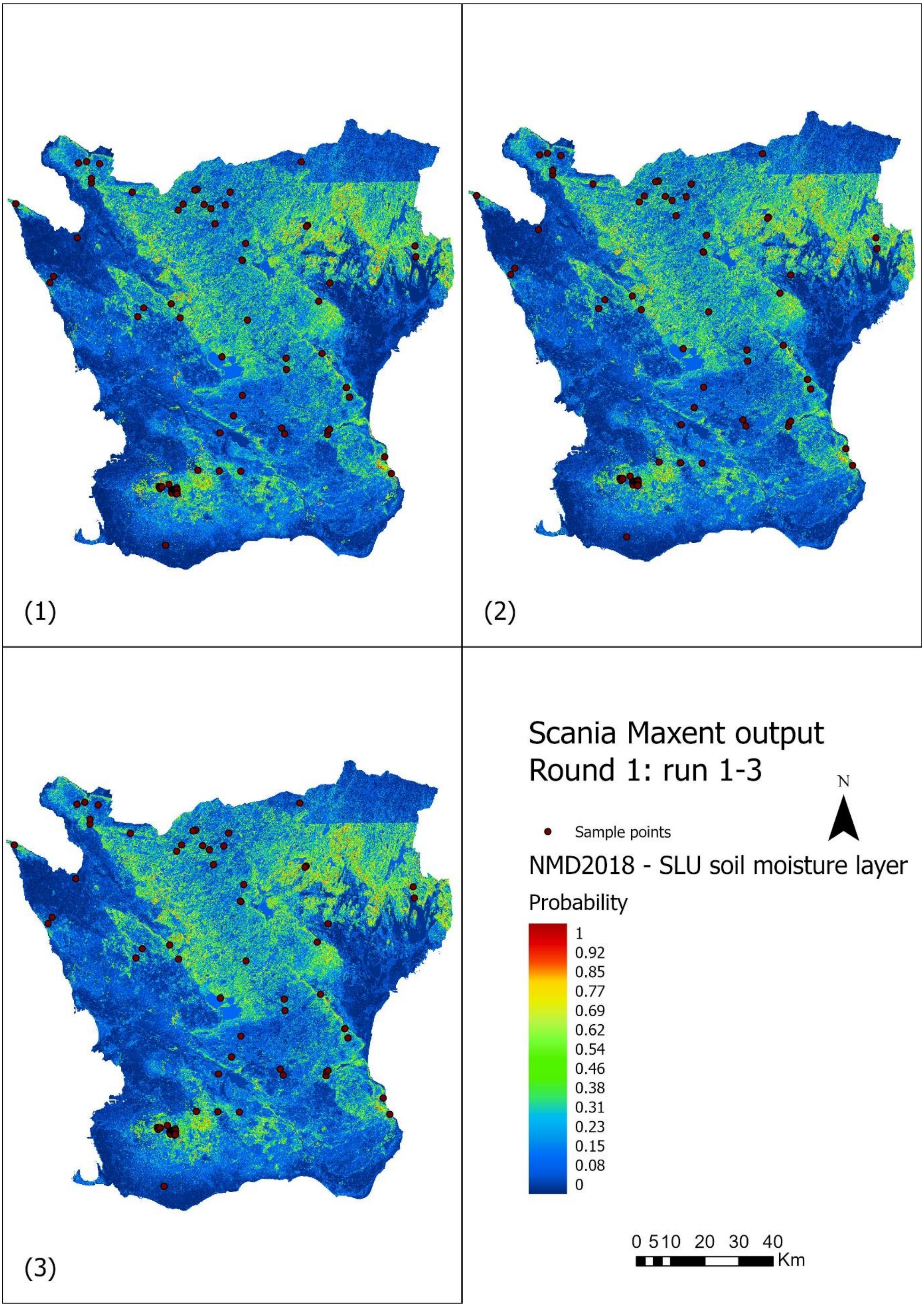

Figure A30: The Maxent probability of occurrence from the first round, run (1), run (2), and run (3), with the environmental factor layer combination of the NMD 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

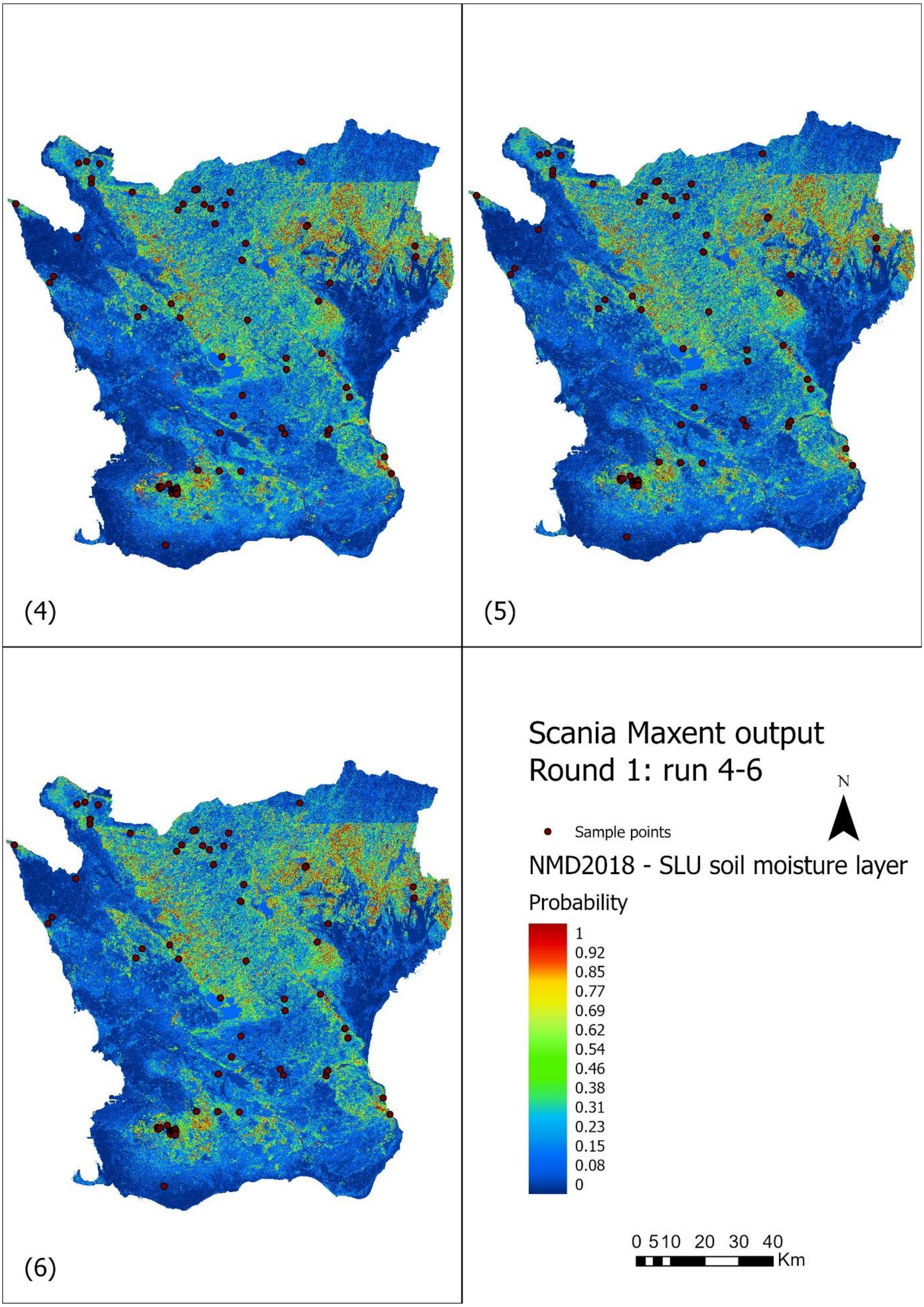

Figure A31: The Maxent probability of occurrence from the first round, run (4), run (5), and run (6), with the environmental factor layer combination of the NMD 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

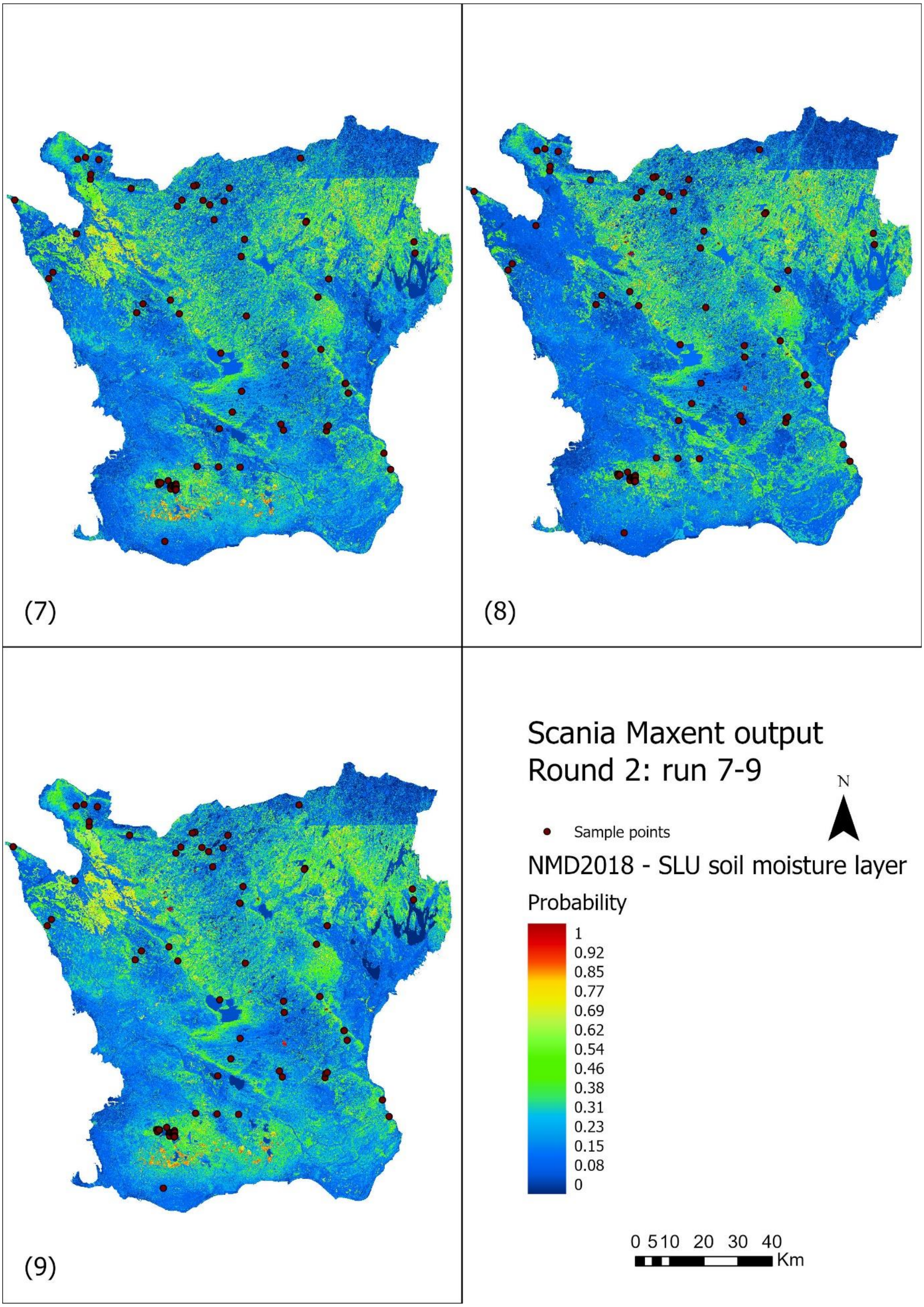

Figure A32: The Maxent probability of occurrence from the second round, run (7), run (8), and run (9), with the environmental factor layer combination of the NMD 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

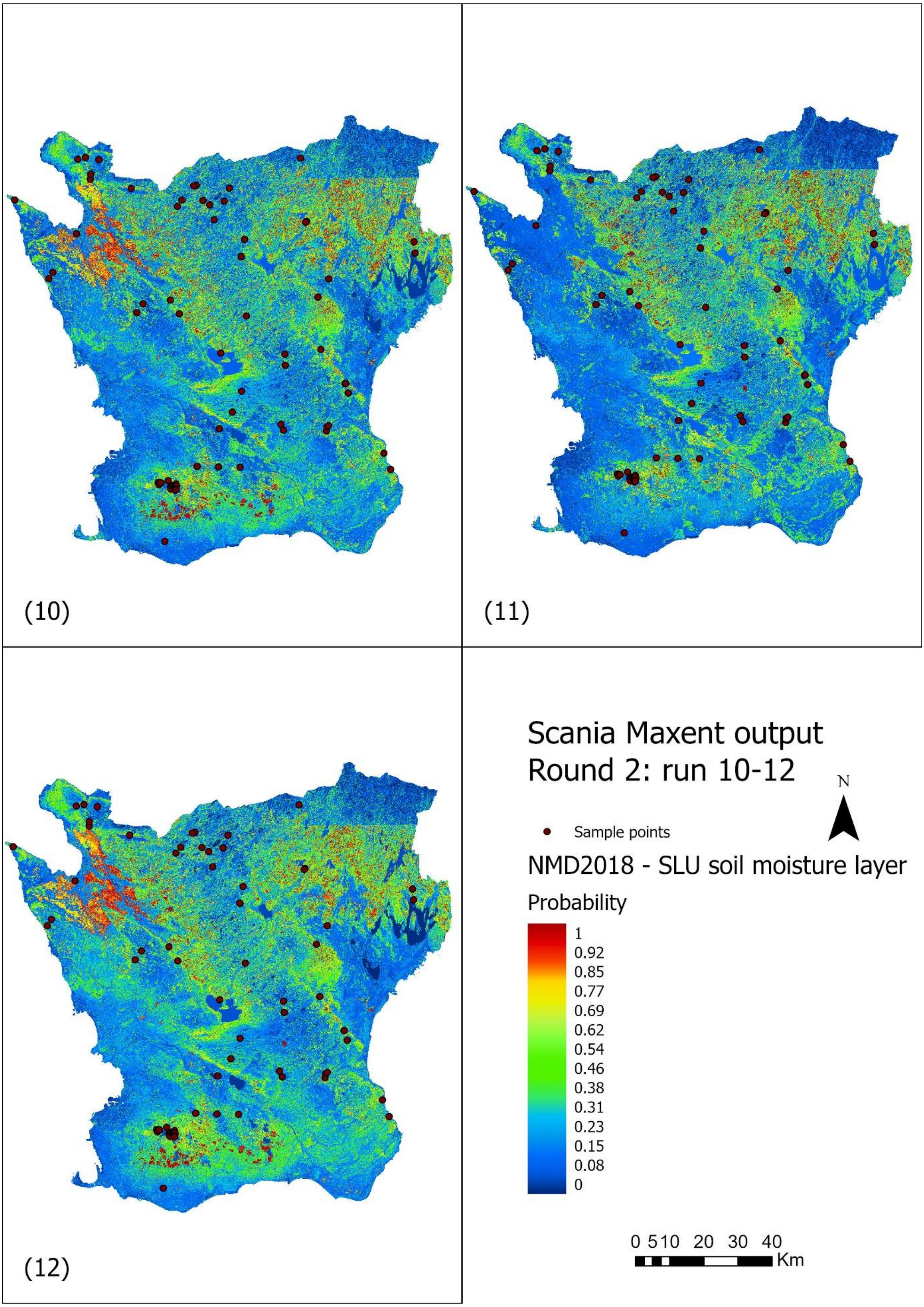

Figure A33: The Maxent probability of occurrence from the second round, run (10), run (11), and run (12), with the environmental factor layer combination of the NMD 2018 land cover- and SLU soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

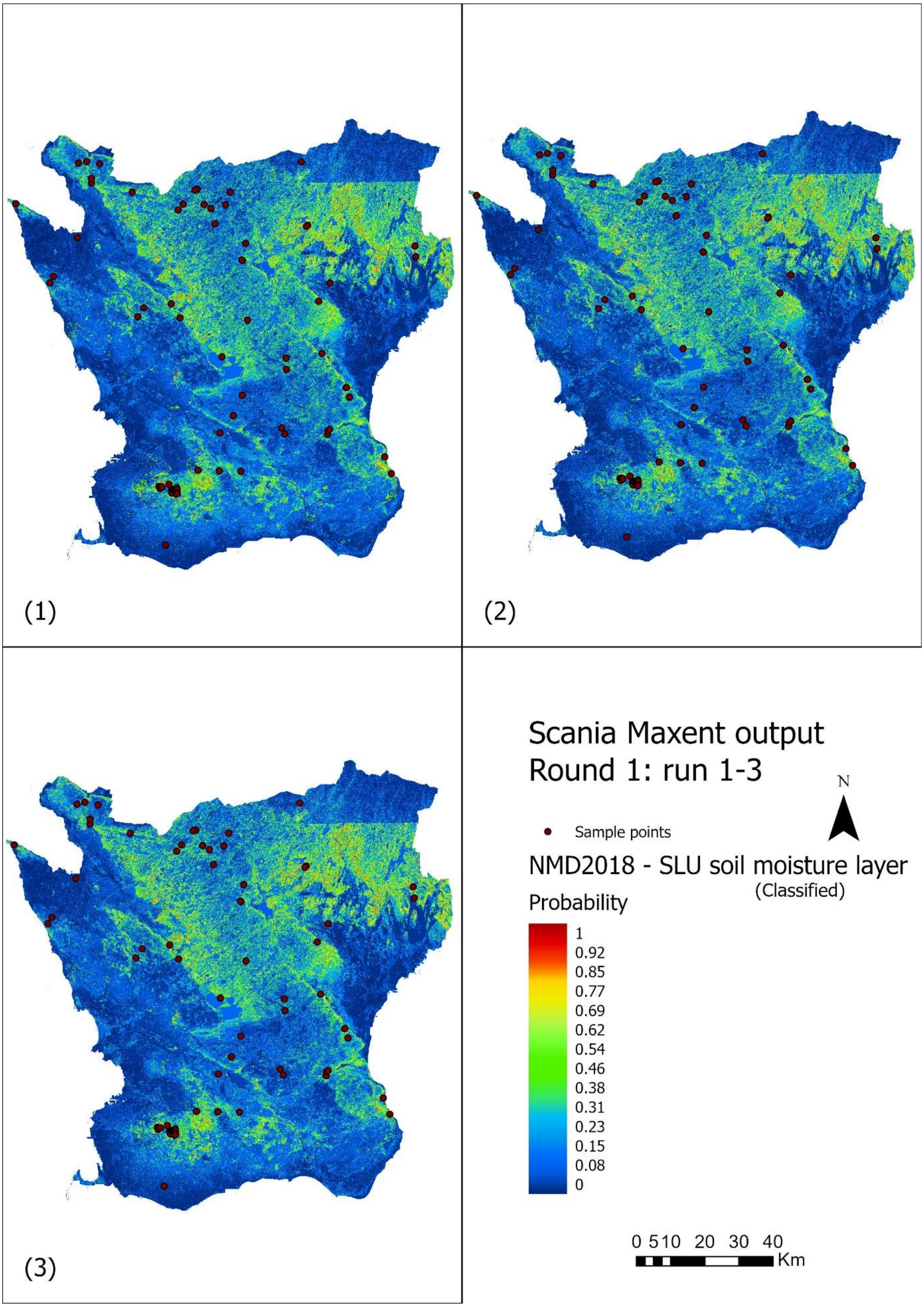

Figure A34: The Maxent probability of occurrence from the first round, run (1), run (2), and run (3), with the environmental factor layer combination of the NMD 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

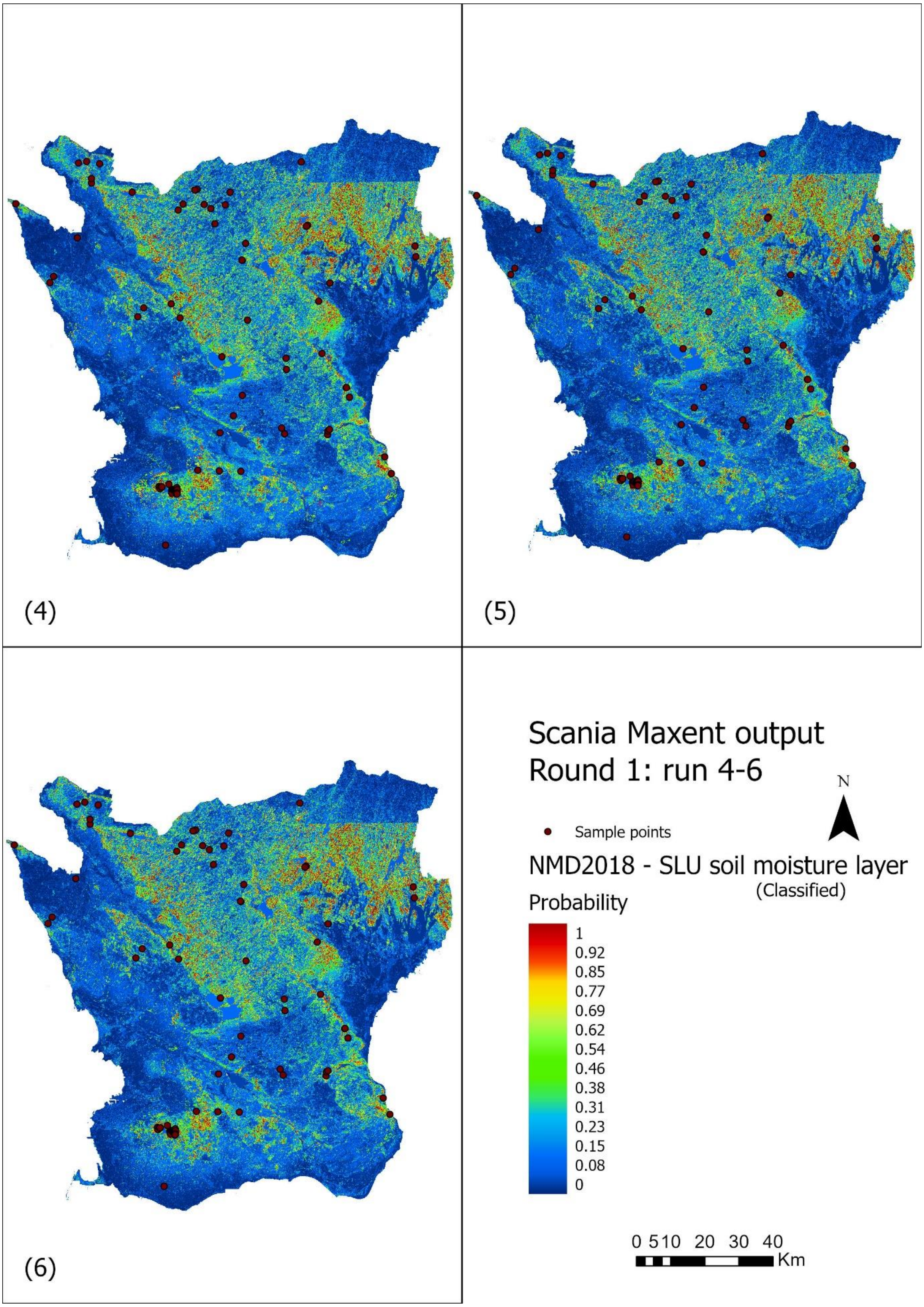

Figure A35: The Maxent probability of occurrence from the second round, run (4), run (5), and run (6), with the environmental factor layer combination of the NMD 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

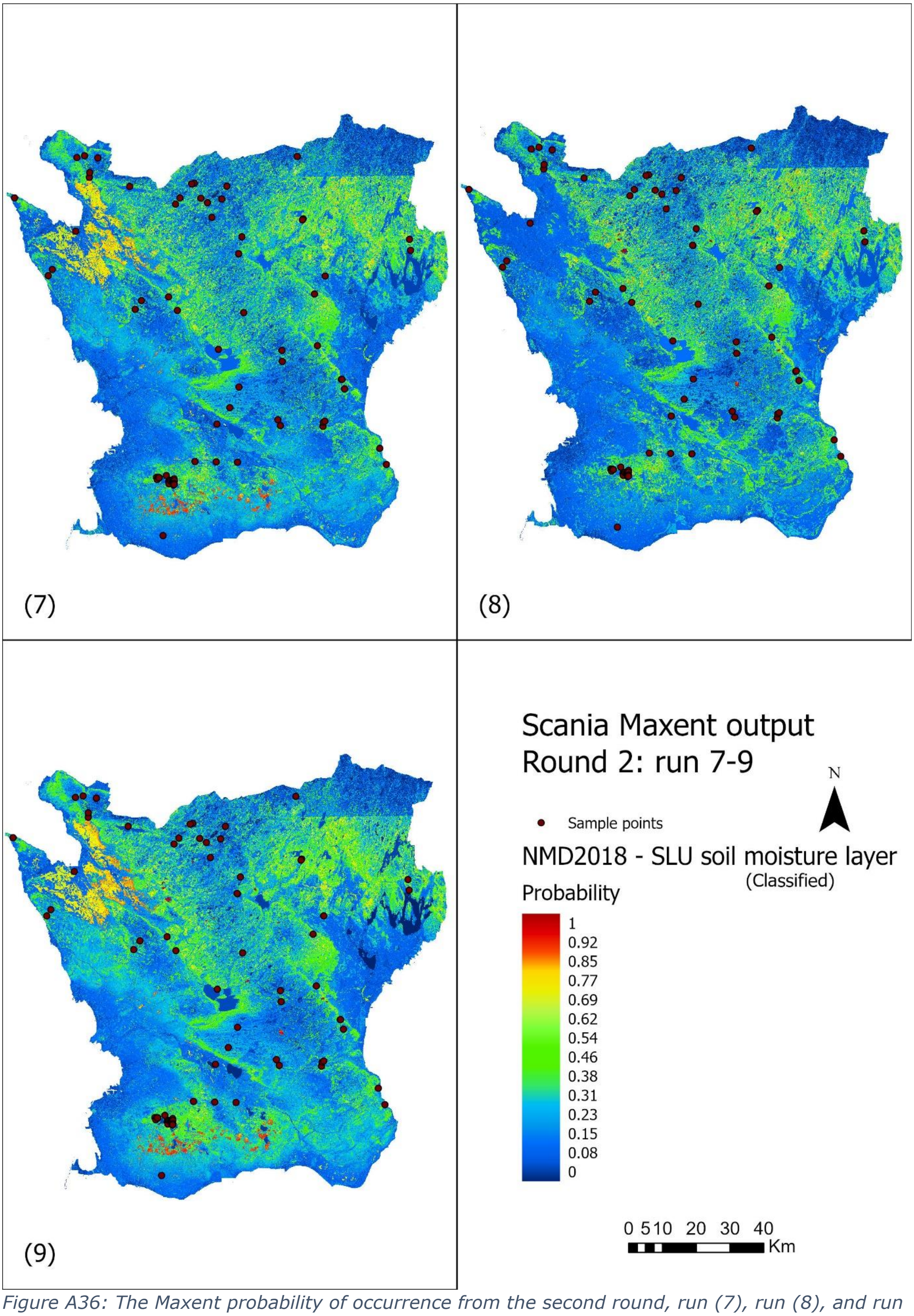

Figure A36: The Maxent probability of occurrence from the second round, run (7), run (8), and run (9), with the environmental factor layer combination of the NMD 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

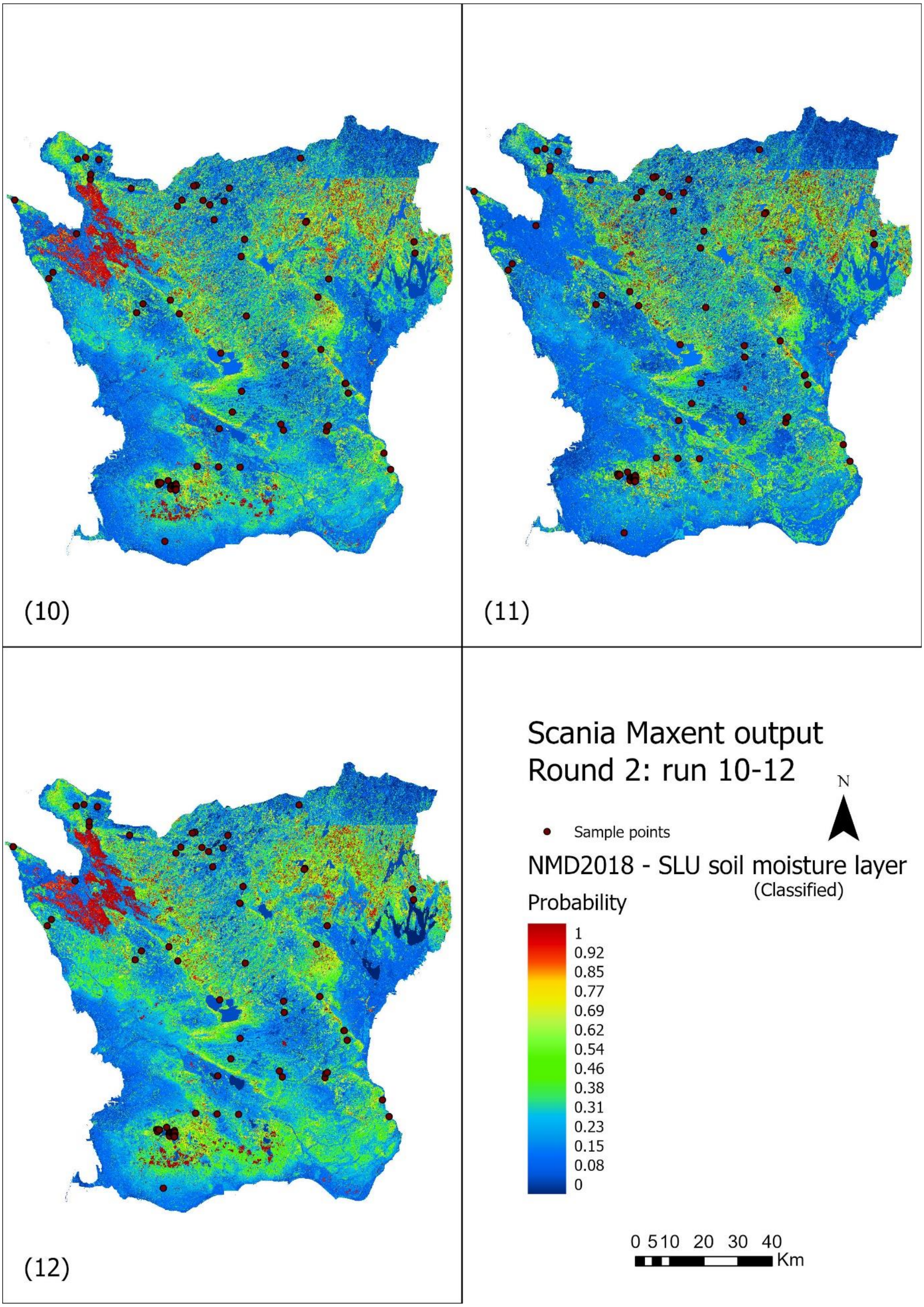

Figure A37: The Maxent probability of occurrence from the second round, run (10), run (11), and run (12), with the environmental factor layer combination of the NMD 2018 land cover- and SLU classified soil moisture layer that also includes the DEM, Slope, Aspect, and soil layer all with the spatial resolution of 10m.

Table A22: The training and test data AUC values for the runs, 1 to 12 of round 1and 2, for the CLC 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer).*

	Scania					
	CLC 2018 and SLU soil moisture layer*					
Round 1	1	2	3	4	5	6
Training data (AUC)	0.852	0.843	0.846	0.852	0.843	0.846
Test data (AUC)	0.794	0.807	0.796	0.794	0.807	0.796
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5
Round 2	7	8	9	10	11	12
Training data (AUC)	0.772	0.785	0.755	0.772	0.785	0.755
Test data (AUC)	0.670	0.636	0.710	0.670	0.636	0.710
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5

Table A23: The training and test data AUC values for the runs, 1 to 12 of round 1and 2, for the CLC 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer).*

	Scania					
	CLC 2018 and SLU classified soil moisture layer*					
Round 1	1	2	3	4	5	6
Training data (AUC)	0.845	0.839	0.842	0.845	0.839	0.842
Test data (AUC)	0.798	0.808	0.797	0.798	0.808	0.797
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5
Round 2	7	8	9	10	11	12
Training data (AUC)	0.764	0.788	0.746	0.764	0.788	0.746
Test data (AUC)	0.658	0.623	0.705	0.658	0.623	0.705
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5

Table A24: The training and test data AUC values for the runs, 1 to 12 of round 1and 2, for the NMD 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer).*

	Scania					
	NMD 2018 and SLU soil moisture layer*					
Round 1	1	2	3	4	5	6
Training data (AUC)	0.856	0.853	0.853	0.856	0.853	0.853
Test data (AUC)	0.806	0.805	0.796	0.806	0.805	0.796
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5
Round 2	7	8	9	10	11	12
Training data (AUC)	0.785	0.808	0.771	0.785	0.808	0.771
Test data (AUC)	0.705	0.676	0.753	0.705	0.676	0.753
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5

Table A25: The training and test data AUC values for the runs, 1 to 12 of round 1and 2, for the NMD 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer).*

	Scania					
	NMD 2018 and SLU classified soil moisture layer*					
Round 1	1	2	3	4	5	6
Training data (AUC)	0.858	0.857	0.856	0.858	0.857	0.856
Test data (AUC)	0.811	0.804	0.795	0.811	0.804	0.795
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5
Round 2	7	8	9	10	11	12
Training data (AUC)	0.775	0.809	0.764	0.775	0.809	0.764
Test data (AUC)	0.697	0.662	0.737	0.697	0.662	0.737
Random Prediction (AUC=0.5)	0.5	0.5	0.5	0.5	0.5	0.5

Table A26: The percentage of each environmental factor the contributes to building the model, for the CLC 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Contribution	Scania					
	CLC 2018 and SLU soil moisture layer*					
Round 1 (% Contribution)	1	2	3	4	5	6
CLC 2018	23.5	24.7	24.8	23.5	24.7	24.8
DEM	18.6	15.8	17.2	18.6	15.8	17.2
Slope	5	7.3	7.1	5	7.3	7.1
Aspect	0.4	1.6	0.8	0.4	1.6	0.8
Soil type	33.3	32	31.5	33.3	32	31.5
SLU soil moisture layer	19.1	18.6	18.6	19.1	18.6	18.6
Round 2 (% Contribution)	7	8	9	10	11	12
CLC 2018	15.7	20.6	14.7	15.7	20.6	14.7
DEM	18.7	18.5	18.5	18.7	18.5	18.5
Slope	1.7	0	1.3	1.7	0	1.3
Aspect	0.7	0.7	0.2	0.7	0.7	0.2
Soil type	36	38.2	40.2	36	38.2	40.2
SLU soil moisture layer	27.2	22.1	25.1	27.2	22.1	25.1

Table A27: The percentage of each environmental factor the contributes to building the model, for the CLC 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Contribution	Scania					
	CLC 2018 and SLU classified soil moisture layer*					
Round 1 (% Contribution)	1	2	3	4	5	6
CLC 2018	24.5	24	24.7	24.5	24	24.7
DEM	20	18.7	18.8	20	18.7	18.8
Slope	13.9	14.8	15.7	13.9	14.8	15.7
Aspect	0.8	2	0.8	0.8	2	0.8
Soil type	37.6	36.1	35.8	37.6	36.1	35.8
SLU soil moisture layer	3.3	4.4	4.2	3.3	4.4	4.2
Round 2 (% Contribution)	7	8	9	10	11	12
CLC 2018	19.4	24.4	18.1	19.4	24.4	18.1
DEM	23.7	20.9	24	23.7	20.9	24
Slope	7.1	0.3	4.9	7.1	0.3	4.9
Aspect	0.5	0.8	0.1	0.5	0.8	0.1
Soil type	38.1	34.7	42.7	38.1	34.7	42.7
SLU soil moisture layer	11.1	18.9	10.2	11.1	18.9	10.2

Table A28: The percentage of each environmental factor the contributes to building the model, for the NMD 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Contribution	Scania					
	NMD 2018 and SLU soil moisture layer*					
Round 1 (% Contribution)	1	2	3	4	5	6
NMD 2018	37.5	42	40.6	37.5	42	40.6
DEM	14.3	11.9	12	14.3	11.9	12
Slope	1.6	4.8	3.7	1.6	4.8	3.7
Aspect	0.4	1.2	0.6	0.4	1.2	0.6
Soil type	31	26.7	28.4	31	26.7	28.4
SLU soil moisture layer	15.1	13.4	14.8	15.1	13.4	14.8
Round 2 (% Contribution)	7	8	9	10	11	12
NMD 2018	38.2	38.8	36	38.2	38.8	36
DEM	11.5	11.1	12.5	11.5	11.1	12.5
Slope	1.2	0	0.6	1.2	0	0.6
Aspect	0.4	0.6	0.2	0.4	0.6	0.2
Soil type	28.2	32.4	32	28.2	32.4	32
SLU soil moisture layer	20.5	17	18.8	20.5	17	18.8

Table A29: The percentage of each environmental factor the contributes to building the model, for the NMD 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Contribution	Scania					
	NMD 2018 and SLU classified soil moisture layer*					
Round 1 (% Contribution)	1	2	3	4	5	6
NMD 2018	37.3	43	41.1	37.3	43	41.1
DEM	16.3	10	11.7	16.3	10	11.7
Slope	11.6	14	14.9	11.6	14	14.9
Aspect	0.5	1.9	0.9	0.5	1.9	0.9
Soil type	32.4	28.5	28.7	32.4	28.5	28.7
SLU soil moisture layer	2	2.5	2.6	2	2.5	2.6
Round 2 (% Contribution)	7	8	9	10	11	12
NMD 2018	39.4	39.8	39.5	39.4	39.8	39.5
DEM	15.1	13.5	16.1	15.1	13.5	16.1
Slope	7.6	0.5	3.2	7.6	0.5	3.2
Aspect	0.4	0.7	0.1	0.4	0.7	0.1
Soil type	30.6	31.3	34	30.6	31.3	34
SLU soil moisture layer	7	14.2	7.1	7	14.2	7.1

Table A30: The percentage of permutation importance of each environmental factor in the established model, for the CLC 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Permutation importance	Scania					
	CLC 2018 and SLU soil moisture layer*					
Round 1 (% Permutation importance)	1	2	3	4	5	6
CLC 2018	22.2	37.3	20.5	22.2	37.3	20.5
DEM	26.5	20.8	13.8	26.5	20.8	13.8
Slope	5.5	9.6	11.8	5.5	9.6	11.8
Aspect	3.5	3	4	3.5	3	4
Soil type	36.8	27	39	36.8	27	39
SLU soil moisture layer	5.6	2.3	10.9	5.6	2.3	10.9
Round 2 (% Permutation importance)	7	8	9	10	11	12
CLC 2018	20.5	24.3	26.1	20.5	24.3	26.1
DEM	32.4	28.1	24.6	32.4	28.1	24.6
Slope	13	0	5.2	13	0	5.2
Aspect	6.3	0	0.9	6.3	0	0.9
Soil type	14.3	34.3	22.6	14.3	34.3	22.6
SLU soil moisture layer	13.5	13.3	20.7	13.5	13.3	20.7

Table A31: The percentage of permutation importance of each environmental factor in the established model, for the CLC 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Permutation importance	Scania					
	CLC 2018 and SLU classified soil moisture layer*					
Round 1 (% Permutation importance)	1	2	3	4	5	6
CLC 2018	23.7	22.3	36	23.7	22.3	36
DEM	25.5	23.9	20.4	25.5	23.9	20.4
Slope	10.7	9	10.3	10.7	9	10.3
Aspect	3.5	3.8	2.3	3.5	3.8	2.3
Soil type	34.6	39.3	29.1	34.6	39.3	29.1
SLU soil moisture layer	1.9	1.8	2	1.9	1.8	2
Round 2 (% Permutation importance)	7	8	9	10	11	12
CLC 2018	20.4	21.6	29.5	20.4	21.6	29.5
DEM	38.6	29.5	26.2	38.6	29.5	26.2
Slope	9.7	0.8	10	9.7	0.8	10
Aspect	2.8	1.6	0	2.8	1.6	0
Soil type	22.8	44.4	26.9	22.8	44.4	26.9
SLU soil moisture layer	5.6	2.1	7.4	5.6	2.1	7.4

Table A32: The percentage of permutation importance of each environmental factor in the established model, for the NMD 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Permutation importance	Scania					
	NMD 2018 and SLU soil moisture layer*					
Round 1 (% Permutation importance)	1	2	3	4	5	6
NMD 2018	41.5	49.1	42.7	41.5	49.1	42.7
DEM	15.2	6.2	8.3	15.2	6.2	8.3
Slope	10.5	3.2	7.1	10.5	3.2	7.1
Aspect	6.3	2.9	2.9	6.3	2.9	2.9
Soil type	25.6	31.5	31.4	25.6	31.5	31.4
SLU soil moisture layer	0.9	7.1	7.6	0.9	7.1	7.6
Round 2 (% Permutation importance)	7	8	9	10	11	12
NMD 2018	29.7	47.6	27.7	29.7	47.6	27.7
DEM	11.1	14.4	23.5	11.1	14.4	23.5
Slope	6	0.2	3.9	6	0.2	3.9
Aspect	3.5	0	0.9	3.5	0	0.9
Soil type	34.8	24.3	25	34.8	24.3	25
SLU soil moisture layer	14.9	13.5	19	14.9	13.5	19

Table A33: The percentage of permutation importance of each environmental factor in the established model, for the NMD 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

% Permutation importance	Scania					
	NMD 2018 and SLU classified soil moisture layer*					
Round 1 (% Permutation importance)	1	2	3	4	5	6
NMD 2018	27.4	34	48.7	27.4	34	48.7
DEM	15.7	5.5	16.3	15.7	5.5	16.3
Slope	8.6	15.1	7.8	8.6	15.1	7.8
Aspect	5.6	6	5	5.6	6	5
Soil type	37.7	39.3	21.1	37.7	39.3	21.1
SLU soil moisture layer	5.2	0	1	5.2	0	1
Round 2 (% Permutation importance)	7	8	9	10	11	12
NMD 2018	22.3	47.7	40.1	22.3	47.7	40.1
DEM	18.5	12.3	26.5	18.5	12.3	26.5
Slope	10.2	1.5	3.4	10.2	1.5	3.4
Aspect	1.7	1.2	0	1.7	1.2	0
Soil type	42.2	35.4	26.7	42.2	35.4	26.7
SLU soil moisture layer	5.1	1.9	3.3	5.1	1.9	3.3

Table A34: How each environmental factor layer was ranked based on their respective percentage values for the CLC 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Contribution	Scania					
	CLC 2018 and SLU soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
CLC 2018	2	2	2	2	2	2
DEM	4	4	4	4	4	4
Slope	5	5	5	5	5	5
Aspect	6	6	6	6	6	6
Soil type	1	1	1	1	1	1
SLU soil moisture layer	3	3	3	3	3	3
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
CLC 2018	4	3	4	4	3	4
DEM	3	4	3	3	4	3
Slope	5	6	5	5	6	5
Aspect	6	5	6	6	5	6
Soil type	1	1	1	1	1	1
SLU soil moisture layer	2	2	2	2	2	2

Table A35: How each environmental factor layer was ranked based on their respective percentage values for the CLC 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Contribution	Scania					
	CLC 2018 and SLU classified soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
CLC 2018	2	2	2	2	2	2
DEM	3	3	3	3	3	3
Slope	4	4	4	4	4	4
Aspect	6	6	6	6	6	6
Soil type	1	1	1	1	1	1
SLU soil moisture layer	5	5	5	5	5	5
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
CLC 2018	3	2	3	3	2	3
DEM	2	3	2	2	3	2
Slope	5	6	5	5	6	5
Aspect	6	5	6	6	5	6
Soil type	1	1	1	1	1	1
SLU soil moisture layer	4	4	4	4	4	4

Table A36: How each environmental factor layer was ranked based on their respective percentage values for the NMD 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Contribution	Scania					
	NMD 2018 and SLU soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
NMD 2018	1	1	1	1	1	1
DEM	4	4	4	4	4	4
Slope	5	5	5	5	5	5
Aspect	6	6	6	6	6	6
Soil type	2	2	2	2	2	2
SLU soil moisture layer	3	3	3	3	3	3
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
NMD 2018	1	1	1	1	1	1
DEM	4	4	4	4	4	4
Slope	5	6	5	5	6	5
Aspect	6	5	6	6	5	6
Soil type	2	2	2	2	2	2
SLU soil moisture layer	3	3	3	3	3	3

Table A37: How each environmental factor layer was ranked based on their respective percentage values for the NMD 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Contribution	Scania					
	NMD 2018 and SLU classified soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
NMD 2018	1	1	1	1	1	1
DEM	3	4	4	3	4	4
Slope	4	3	3	4	3	3
Aspect	6	6	6	6	6	6
Soil type	2	2	2	2	2	2
SLU soil moisture layer	5	5	5	5	5	5
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
NMD 2018	1	1	1	1	1	1
DEM	3	4	3	3	4	3
Slope	4	6	5	4	6	5
Aspect	6	5	6	6	5	6
Soil type	2	2	2	2	2	2
SLU soil moisture layer	5	3	4	5	3	4

Table A38: How each environmental factor layer was ranked based on their respective percentage values for the CLC 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Permutation importance	Scania					
	CLC 2018 and SLU soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
CLC 2018	3	1	2	3	1	2
DEM	2	3	3	2	3	3
Slope	5	4	4	5	4	4
Aspect	6	5	6	6	5	6
Soil type	1	2	1	1	2	1
SLU soil moisture layer	4	6	5	4	6	5
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
CLC 2018	2	3	1	2	3	1
DEM	1	2	2	1	2	2
Slope	5	5	5	5	5	5
Aspect	6	5	6	6	5	6
Soil type	3	1	3	3	1	3
SLU soil moisture layer	4	4	4	4	4	4

Table A39: How each environmental factor layer was ranked based on their respective percentage values for the CLC 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Permutation importance	Scania					
	CLC 2018 and SLU classified soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
CLC 2018	3	3	1	3	3	1
DEM	2	2	3	2	2	3
Slope	4	4	4	4	4	4
Aspect	5	5	5	5	5	5
Soil type	1	1	2	1	1	2
SLU soil moisture layer	6	6	6	6	6	6
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
CLC 2018	3	3	2	3	3	2
DEM	1	2	3	1	2	3
Slope	4	6	4	4	6	4
Aspect	6	5	6	6	5	6
Soil type	2	1	1	2	1	1
SLU soil moisture layer	5	4	5	5	4	5

Table A40: How each environmental factor layer was ranked based on their respective percentage values for the NMD 2018 land cover- and SLU soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Permutation importance	Scania					
	NMD 2018 and SLU soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
NMD 2018	1	1	1	1	1	1
DEM	3	4	3	3	4	3
Slope	4	5	5	4	5	5
Aspect	5	6	6	5	6	6
Soil type	2	2	2	2	2	2
SLU soil moisture layer	6	3	4	6	3	4
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
NMD 2018	2	1	1	2	1	1
DEM	4	3	3	4	3	3
Slope	5	5	5	5	5	5
Aspect	6	6	6	6	6	6
Soil type	1	2	2	1	2	2
SLU soil moisture layer	3	4	4	3	4	4

Table A41: How each environmental factor layer was ranked based on their respective percentage values for the NMD 2018 land cover- and SLU classified soil moisture layer (Including DEM, Slope, Aspect, and soil layer) of round 1 and 2 for the runs 1 to 12.*

Ranked Permutation importance	Scania					
	NMD 2018 and SLU classified soil moisture layer*					
Round 1	**1**	**2**	**3**	**4**	**5**	**6**
NMD 2018	2	2	1	2	2	1
DEM	3	5	3	3	5	3
Slope	4	3	4	4	3	4
Aspect	5	4	5	5	4	5
Soil type	1	1	2	1	1	2
SLU soil moisture layer	6	6	6	6	6	6
Round 2	**7**	**8**	**9**	**10**	**11**	**12**
NMD 2018	2	1	1	2	1	1
DEM	3	3	3	3	3	3
Slope	4	5	4	4	5	4
Aspect	6	6	6	6	6	6
Soil type	1	2	2	1	2	2
SLU soil moisture layer	5	4	5	5	4	5