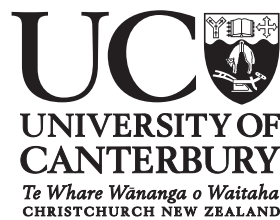

A New Approach to Adaptive Monitoring

A thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy in Statistics
at the University of Canterbury

by
Peter Jaksons



School of Mathematics and Statistics
University of Canterbury, Christchurch
New Zealand, 2014

“You have to be fast on your feet and adaptive or else a strategy is useless.”

Charles de Gaulle

Acknowledgements

First of all, I would like to thank my supervisors Prof. Jennifer Brown and Dr. Blair Robertson for all their hard work, help and support during my thesis. Without you I would probably be still looking at my river model. Additionally, I would like to thank Prof. Jennifer Brown, and the department of Mathematics and Statistics for awarding me a PhD scholarship, which gave me the chance to come to New Zealand. The last couple of years have been an amazing experience (including the earthquakes) and New Zealand is such a wonderful place that I do not think I will go home any time soon.

I would like to thank my mum, dad and brother for being so understanding and supportive while I was doing that “counting species thing” on the other side of the world. Geen paniek, ik ben echt wel van plan om terug thuis naar huis te komen.

I also would like to thank all my friends in New Zealand and overseas for being there for me. Special thank go to Shannon, Anna, Joe, Nick, Rachelle, Abdul and all the other postgrad students at the department who almost made studying statistics look cool.

I would like to thank Professor Brian McArdle and Professor Frank d’Amico for reviewing my thesis and providing many insightful suggestions.

Last but not least, I would like to thank Rodelyn for being so patient with me while I was working on my thesis. And oh yes... also for carrying our baby.

Contents

Abstract	i
List of Abbreviations	ii
List of Symbols	iv
1 Introduction	1
1.1 Motivation	1
1.2 Standing on the Shoulders of Giants: How can this Thesis make a Difference? . .	9
1.3 Thesis Outline	13
2 A New General Framework for Adaptive Monitoring	14
2.1 Adaptive Monitoring: Proposed Algorithm	14
2.2 Sampling	16
2.2.1 Probability Sampling	18
2.2.2 Spatial autocorrelation and Spatially Balanced Sampling	21
2.3 Adaptive Monitoring Using Probability Sampling	24
2.4 Discussion	25
3 The Selection of a Sampling Design for Adaptive Monitoring	27
3.1 Probability Sampling Designs	28
3.1.1 Simple Random Sampling	28
3.1.2 Stratified Random Sampling	29

3.1.3	Conditional Poisson Sampling	34
3.2	Spatially Balanced Sampling Designs	35
3.2.1	Systematic Sampling	35
3.2.2	StratRS to Improve the Spatial Balance	37
3.2.3	Generalized Randomized Tessellation Stratified Sampling (GRTS)	38
3.2.4	Balanced Acceptance Sampling (BAS)	43
3.2.5	Other Spatially Balanced Sampling Designs	49
3.2.6	Test for Spatial Balance	50
3.3	The Selection of Sampling Design	55
3.3.1	Detection Rate Initial Sample	56
3.3.2	Population Estimation	63
3.4	Practical aspects of BAS	74
3.4.1	Long Term Sampling	74
3.4.2	Premature Ending, Non-Response and Oversampling	77
3.4.3	Computational Efficiency of BAS	77
3.5	Discussion	80
4	Adaptive Monitoring using Ecological Information to update the inclusion probabilities	82
4.1	Introduction	82
4.2	Methodology	87
4.2.1	Setting the inclusion probabilities using species distribution models	87
4.2.2	Rescaling Approaches	89
4.3	Simulation Study	91
4.3.1	Asian tiger mosquito Data	91
4.3.2	Testing the use of different Sampling Designs and Rescaling Techniques	94
4.3.3	Test for Spatially Balanced Sampling Designs vs Non-spatially balanced Designs.	103
4.4	Discussion	106

5	Adaptive Monitoring Using Spatial Information: Nearest Unit Tessellation	109
5.1	Introduction	109
5.2	Methodology	113
5.2.1	Simple Nearest Unit Tessellation	115
5.2.2	Unequal Nearest Unit Tessellation	115
5.2.3	Presence Nearest Unit Tessellation	117
5.2.4	Presence/Absence Nearest Unit Tessellation	120
5.2.5	Count Nearest Unit Tessellation	121
5.2.6	Adaptive Monitoring Using the NUT Methods	123
5.3	Simulation Study	124
5.3.1	Details of the Species Populations	125
5.3.2	Results	127
5.4	Discussion	135
6	Adaptive Monitoring using Ecological and Spatial Information: Eradication of the Great White Butterfly	138
6.1	Methodology Adaptive Monitoring using Ecological and Spatial Information . . .	140
6.2	The Great White Butterfly Data	142
6.3	Simulation study: Stationary GWB population	144
6.3.1	Results	147
6.4	Adaptive Monitoring to set up an Eradication Strategy	151
6.5	Simulation study: Non-stationary GWB Populations	152
6.5.1	Results	153
6.5.2	Alternative Eradication Method	154
6.6	Discussion	156
7	Discussion and Future Work	158
7.1	Future Work	163

8	Appendix	165
8.1	Appendix A: Data case studies	165
8.1.1	Ecological Covariates	165
8.1.2	Distribution maps Asian tiger mosquito	170
8.2	Appendix B: R code	171

Abstract

Invasive species are an increasing problem, costing several billion dollars to the global economy. Monitoring methods are needed to provide scientists the necessary information to control populations of invasive species. Adaptive monitoring means that, for each additional survey, the monitoring design is updated based on the information obtained during the previous surveys. In this thesis we introduce a new general framework for adaptive monitoring. This method focuses on increasing the detection rate of invasive species over consecutive surveys. Compared with the existing monitoring methods, the proposed algorithm aims to improve adaptive monitoring by the following three points: (1) The use of a spatially balanced sampling design to select a sample in each survey, (2) by using both spatial and ecological information to update the monitoring strategy, and (3) by incorporating an eradication strategy to an adaptive monitoring design. In Chapter 1, we emphasise the need for the development of adaptive monitoring methods for invasive species. In Chapter 2, we outline the algorithm of our proposed method for adaptive monitoring and discuss the use of spatially balanced sampling designs. In Chapter 3, several sampling designs are introduced and their use for (adaptive) monitoring is evaluated. We give special attention to a new spatially balanced sampling design, named Balanced Acceptance Sampling, and compare it with a selection of existing (spatially balanced) sampling designs such as Generalized Random Tessellation Stratified sampling. In Chapter 4, we illustrate how ecological information can be used to update the monitoring strategy, which is demonstrated using a case study on the Asian tiger mosquito. In addition, several practical issues with probability sampling are discussed. In Chapter 5, the Nearest Unit Tessellation methods are introduced. These methods can model the observed spatial information of the species to update the monitoring strategy. Finally, we demonstrate how to combine ecological and spatial information to adjust a monitoring strategy. Chapter 6 also explores the use of a method to incorporate an eradication strategy to an adaptive monitoring strategy, using the Great White Butterfly data.

List of Abbreviations

Acronym	Definition
AMBAS	Adaptive monitoring using balanced acceptance sampling
AMStratRS	Adaptive monitoring using stratified random sampling
ARS	Acceptance/Rejection Random Sampling
BAS	Balanced Acceptance Sampling
CNUT	Count Nearest Unit Tessellation
CONS	Conservative π_i approach
CP	Conditional Poisson
GRTS	Generalized Randomized Tessellation Stratified
HT	Horvitz-Thompson
GWB	Great white butterfly
LIM.xx	Minimum suitability approach with a lower limit of 0.xx
LPM1	Local Pivotal Methods 1
LPM2	Local Pivotal Method 2
NUT	Nearest Unit Tessellation
PANUT	Presence/Absence Nearest Unit Tessellation
PNUT	Presence Nearest Unit Tessellation
PROG	Progressive π_i approach
RV	Relative variance
SDM	Species Distribution Model

Acronym	Definition
SNUT	Simple Nearest Unit Tessellation
SRS	Simple Random Sampling
StratRS	Stratified Random Sampling
UNUT	Unequal Nearest Unit Tessellation

List of Symbols

Symbol	Definition
D	Detection rate
\bar{D}	Mean detection rate
ϕ_p	Radical inverse function with base p
L	Number of strata
h	Strata index
$I_S(i)$	Indicator function for samples: One if unit i is sampled/included and zero otherwise
$I_D(i)$	Indicator function for detection: One if the species is present in unit i and zero otherwise
k	Survey number
m	The number of samples / Monte Carlo simulation
N	Population size
n	Sample size
p	Rescaled inclusion probability with maximum of one
π_i	First order inclusion probability
π_{ij}	Second order inclusion probability
π'_i	Ideal inclusion probability
$\hat{\pi}_i$	Observed inclusion probability
π_i^{eco}	Inclusion probability based on ecological information

Symbol	Definition
π_i^{spat}	Inclusion probability based on spatial information
π_i^{both}	Inclusion probability based on ecological and spatial information
ψ	Area Voronoi polygon
S	Species habitat suitability
\hat{S}	Estimated species habitat suitability
s^2	Sample variance
σ^2	Population variance
U	Set of all units in the population
V	Variance
\hat{V}	Estimated Variance
$\hat{V}_{\text{NBH}}(\hat{Y})$	Local mean variance estimator
ξ	Variance area Voronoi polygons for a sample size n
Y	Population total
\hat{Y}	Estimated population total
y	Number of individuals of a species (or presence/absence) in unit
\bar{y}	Sample mean number of individuals of a species per unit

Chapter 1

Introduction

1.1 Motivation

In this thesis, a new generic method for adaptive monitoring is introduced to improve the detection rate of invasive species. The motivation for such a method is threefold: (1) Any increase of invasive species is costly, (2) the need of high detection rates for eradication of such species, and (3) a lack of generic methods for adaptive monitoring to increase these detection rates.

(1) Invasive Species

In 1999 USA President Clinton signed Executive Order 13112 to define invasive species as follows:

“Invasive species are those plants, animals, and microbes that are non-native to an area and have caused or have the potential to cause economic, health or ecological damage.”

New invasions by species can occur naturally. However, the rate at which new introductions of invasive species occur has increased in recent decades as a by-product of international travel and trade (Aukema et al., 2010; Hulme, 2009). Because of this, the number of invasive species that have managed to survive and settle in their new host environment has increased as well. An example of an invasive species are the caterpillars from the Asian gypsy moth, *Lymantria dispar asiatica*. These caterpillars cause severe defoliation, reduced growth and accelerated

mortality of host trees throughout large parts of the northern hemisphere (Ranjan et al., 2008; Epanchin-Niell et al., 2012). Another, well-known example, is the invasion of the Nile perch, *Lates niloticus*, of Lake Victoria, Africa. Since the species introduction in 1954, it has contributed to the extinction of more than 200 endemic fish species through direct competition with, or predation of, other species (Strayer, 2010). A third example is *Caulerpa taxifolia*, a popular decorative aquarium alga. Since it was accidentally introduced into the Mediterranean Sea, it has spread over more than 13,000 hectares of the seabed (Ranjan et al., 2008; Glasby, 2012). Growth of *Caulerpa taxifolia* prevents the establishment of many native seaweeds. As a direct result of this, almost all marine life is absent in areas invaded by *Caulerpa taxifolia* (Glasby, 2012).

The effect of invasive species on the stability of many ecosystems worldwide causes direct and indirect economic costs. For example, invasive species may cause serious damage to agricultural productivity by using agricultural land as foraging grounds (Kean et al., 2008; Pimentel et al., 2001). These economical implications of new invasions, on top of increasing expenditures for controlling invasive species, have grown to a concerning cost for many effected parties (Aukema et al., 2010). An example is the cost breakdown of the invasion of the emerald ash borer in the United States which mainly effects ash trees. It is expected that for the next ten years alone, the total cost for home-owners and municipalities will be approximately 10.7 billion US dollars for (ash) tree treatment or removal (Kovacs et al., 2010). Pimentel et al. (2005) estimated that invasive plant species alone cause an environmental and economical cost of 34 US billion dollars. The total cost of invasive species in the United States is estimated around 120 billion US dollars. Giera and Bell (2008) estimated the total economic cost of invasive species and pests to New Zealand's primary agricultural sector to be approximately 2.128 billion NZ dollars per year. With the inclusion of secondary up and downstream services and industries, these costs could be as high as 3.291 billion NZ dollars, or 1.96 percent of New Zealand's GDP in 2008 (Giera and Bell, 2008).

Furthermore, some invasive species can cause a direct threat to human health. For example, the hairs of the caterpillars of the Asian gypsy moth can cause allergic skin reactions to some people (Ranjan et al., 2008). Brush-tail possums, which have invaded many parts of New Zealand,

are carriers of the Bovine Tuberculosis Myco bacterium, better known as bovine TB disease (Ramsey and Efford, 2010). Due to global warming, several mosquito species are settling in new host areas. Some of these mosquitoes can transmit diseases such as malaria and yellow fever (Medlock et al., 2012).

(2) Detection Rates

The first step in the management of invasive species is the prevention of new introductions. However, once an invasive species manages to invade and settle, the management strategy changes to controlling the newly established species population. This second step is aimed at the prevention of further dispersal of the species and/or by eradication of the species (Parnell et al., 2009; Epanchin-Niell et al., 2012; Burgman et al., 2013). Once an invasive species population reaches a certain threshold size, it is hard to control further growth and dispersal of the species population. Therefore, it is important to start eradication as soon as an invasive species is detected (Tobin et al., 2011; Suckling et al., 2012). Nevertheless, eradication will often only be considered if the potential environmental loss and economical costs are higher than the cost of the eradication method (Gaeta et al., 2012). Therefore, the availability of efficient and early eradication strategies are critical for decision makers to reduce the ecological and economical costs of new invasions (Lodge et al., 2006; Hulme, 2009; Epanchin-Niell et al., 2012).

An essential step for efficient eradication, especially at an early stage of the invasion, is to map the distribution of the invasive species population. This is important to target those areas that are invaded, or are likely to be invaded in the future, by the species. The prevalence of an invasive species is defined as the number of units that are occupied by the species in the population of sampling units that constitute to the sampling domain. When selecting a sample, the number of times the targeted species is present when visiting a site or unit over the total number of visited sites, is what we will refer to in this thesis as the detection rate of a sample. In this thesis we will assume that if a species is present it will always be observed, hence we assume perfect detectability. It is important to note that the term ‘detection rate’ should not be confused with detectability or detection probability which is the probability a

species is discovered after it has invaded a new area or the probability a species is detected when visiting a site in which it is actually present (Thompson and Seber, 1996; Moore et al., 2011). Various studies focus on these concepts of detection probability. For example, Moore et al. (2011) estimated the detection probability for large plant species. However, in this thesis, the focus is on the detection rate, thus on methods to select those units in which the species is actually present. An alternative term for the detection rate could have been the incidence, as for example used in Wright (1991). Arguably, incidence would have been a better terminology given the close, and likely confusing, resemblance between the terms detection rate and detection probability. However, in many biological papers, and more specifically in epidemiological papers, the definition of incidence is the number of new cases of an event, for example the number of newly infected persons or newly invaded sites within a certain time interval. Hence, in this thesis it was chosen to define the detection rate D of a sample of size n as

$$D = \frac{1}{n} \sum_{i=1}^n I_D(i), \quad (1.1)$$

where n is the sample size and $I_D(i)$ is an indicator function that equals one if the species is present in the i_{th} sampled unit and zero otherwise. For example, 100 equally sized sites or units are visited. The invasive species is abundant in 20 of those units. In this case, the detection rate is $D = \frac{20}{100} = 0.2$. Note that although this definition for the detection rate is intuitive, there are several other formats or definitions in which studies have represented the detection rate. For example, Crall et al. (2013) and Guisan et al. (2006) used contingency tables of presence/absence of a species in sampled units and Le Lay et al. (2010) used figures to illustrate the positive and negative predictive power of a sample. Ultimately, all these studies aim to improve the detection rate of a sample.

For pest control in general, and in surveillance and monitoring programmes for invasive species in particular, high detection rates are critical so that new incursions can be quickly eradicated or contained (Crall et al., 2013; Rout et al., 2014). Because selecting units in which the species is present or may be present during the course of monitoring is important, the detection rate

can be used to describe the efficiency of an eradication strategy (Kean et al., 2008; Le Lay et al., 2010; Crall et al., 2013). The higher the detection rate the better the eradication strategy. For example, Bogich et al. (2008) showed that the higher the detection rate, the lower the financial cost is of an eradication program and the higher the success rate will be. Many other studies discuss the importance of increasing the detection rate for eradication strategies or environmental management in general (Mehta et al., 2007; Kean et al., 2008; Le Lay et al., 2010; Epanchin-Niell et al., 2012; Whittle et al., 2013). For example, Harvey et al. (2009) analysed the relationship between the sampling intensity and the detection rate for a recently introduced water flea, *Cercopagis pengoi*, in Lake Ontario. They concluded that focussing the sampling effort on those areas with the highest species suitability is preferred, rather than less intensive sampling at a wider array of sites.

Despite the knowledge that detection rates are important for eradication, Kean et al. (2008) stated that one of the greatest constraints to implementing an eradication strategy is the lack of appropriate methods to improve the detection rate. In addition to eradication purposes, another reason for having high detection rates is for analytical purposes, such as species distribution modelling or species dispersal modelling (Alexander et al., 1997; Edwards et al., 2005). Many of these techniques give poor results when the number of presence units are small (Stockwell and Peterson, 2002; Lobo et al., 2010).

(3) Adaptive Monitoring

Strategies for invasive species management are usually long lasting programs that require repeated surveying. Repeated surveying or long term surveying in ecology is generally known as monitoring. A monitoring design can sometimes be designed prior to the initial survey and can remain unchanged for the entire duration of the study. These monitoring designs will often have an informed design based on the available prior information. This prior information could be any type of information about the invasive species, for example details about previous invasions, auxiliary habitat information or expert knowledge. By incorporating this prior information into the monitoring design better results can be obtained compared with randomly selected samples. Later in this thesis, we will illustrate that even when no prior information is available better

results can be obtained, compared with simple random sampling, by selecting samples with an improved spatial coverage.

These non-adaptive monitoring designs mainly rely on prior information to improve the detection rates. If possible, it could be useful, to verify the appropriate usage of the used prior information over time, or even to improve the design of the monitoring design so that it no longer depends on prior information only. Given the consecutive surveys and long term aspect of monitoring, it is possible to adjust the monitoring design as a function of the observed data over time. Adjusting the monitoring design over time is known as adaptive monitoring (Lindenmayer and Likens, 2009; Guisan et al., 2006; Stohlgren and Schnase, 2006; Crall et al., 2013). Lindenmayer and Likens (2009) defined adaptive monitoring as:

“A monitoring program in which the development of conceptual models, question setting, experimental design, data collection, data analysis, and data interpretation are linked as iterative steps. An adaptive monitoring program is one that can evolve in response to new questions, new information, situations or conditions, or the development of new protocols but this must not distort or breach the integrity of the data record.”

The concept of adaptive monitoring is intuitive and appealing because it mimics how field-biologists would like to collect data. Areas with a high abundance of the species, or with an expected future high abundance of the species, should be sampled more intensely compared with low density or low risk areas. For example, in conducting a survey of a rare species, a field-biologist may feel inclined to sample more intensively in an area where one individual has been observed to see if others are present (Brown et al., 2008). Lindenmayer et al. (2011) illustrated adaptive monitoring using two case studies where the pre-existing monitoring programs were redesigned over time: One of these case studies is on adaptive monitoring of Bitou bush, an invasive weed species in Booderee National Park, south-eastern Australia. Adaptive monitoring, in this case, leads to the decision to respray those areas with an expected high Bitou bush density. This is done to kill new Bitou bush germinants and to prevent new Bitou bush growth. There are numerous other papers that use the term adaptive monitoring or describe an

example in which the monitoring design was adapted over time. For example, Armstrong and Louw (2013) discussed an adaptive monitoring design for the eggs of the Karkloof blue butterfly and Tombre et al. (2013) discussed adaptive monitoring for two growing goose populations in Norway that use agricultural land as their breeding grounds.

Note, firstly, that adaptive monitoring, as defined in this thesis, is different from adaptive sampling. Adaptive sampling means that the data collection design depends sequentially on the observed values in the previously selected units (Salehi and Brown, 2010; Thompson, 2012). The most well-known sampling design is adaptive cluster sampling (Thompson, 1990). These adaptive sampling designs aim to improve the precision of population estimates. To achieve this, adaptive sampling designs select a batch of units in two (or more) phases. The first phase provides information to decide which units should be selected in the second phase. Even though adaptive sampling designs select a sample in several phases, generally the entire sample is selected within one survey (for example within a one year long survey). With monitoring, instead of collecting all the data during one survey, the sampling effort will be spread over multiple sequential surveys. This is because adaptive monitoring designs are generally used to monitor changes in the population over time.

Secondly, it is important to notice the difference between adaptive monitoring and adaptive management. Monitoring is the process of sampling and data collection. Management is the process in which practical measures are taken to control the invasive species population. Similar to adapting the monitoring design it can also be beneficial to adapt the management strategy (Buckley, 2008; Foxcroft, 2004). Hence, one can for example have adaptive management without adaptive monitoring.

Despite many case specific applications, literature on general frameworks for adaptive monitoring is difficult to find. Many methodological studies about adaptive monitoring mainly focus on specific parts of the (adaptive) monitoring process. For example, Bogich et al. (2008) and Epanchin-Niell et al. (2012) analysed the optimal sampling intensity as a function of cost constraints and applied it to a case study of the Gypsy moth in California, USA. Other studies have focussed on how the sampling effort should be allocated over the consecutive years (Lobo

et al., 2010; Epanchin-Niell et al., 2012). Interestingly, only few studies discuss how to actually allocate the samples over the study area for adaptive long term monitoring.

An attempt to define a general framework for adaptive monitoring was given in Stohlgren and Schnase (2006). They outlined an iterative monitoring scheme using a flowchart. Stohlgren and Schnase (2006) stated that the collected ecological auxiliary information from previous samples should be used to select future samples. In the same year, Guisan et al. (2006) gave a practical illustration of a similar iterative and adaptive monitoring scheme to the one that was introduced by Stohlgren and Schnase (2006). In their study, they described how to use the collected ecological information from previous surveys to set strata boundaries for sampling of rare plant species in Switzerland. They concluded that by using the collected information, the detection rate of a survey can be increased. A similar study was performed in Crall et al. (2013) in which environmental information was used to set and adjust the strata boundaries over time. In their study a similar iterative adaptive stratification sampling method was tested for two invasive plant species in Wisconsin, USA. Crall et al. (2013) showed how the detection rate can be improved for both invasive species compared with stratified sampling. Le Lay et al. (2010) used the same adaptive stratified monitoring method to help detect invasive species and non-invasive species in the Swiss Alps. In this case, the ecological information was modelled using two different distribution modelling techniques to predict the distribution of the species. Based on these predictions the area was stratified into a low, a medium and a high likelihood stratum. A stratified random sample was selected with half of the sample size allocated to the medium likelihood stratum and half of the sample size was allocated to the high likelihood stratum. They showed that the probability of finding a rare plant species could be increased by almost 50 percent by using ecological information to select the sample compared with random sampling, and for a common plant species the new number of detections increased by almost 100 percent.

1.2 Standing on the Shoulders of Giants: How can this Thesis make a Difference?

The last three cited studies on adaptive monitoring show that using ecological auxiliary information to update the monitoring method can improve the detection rate. In this thesis we will try to extend the methods used in those studies on adaptive monitoring to potentially increase the detection rate. To do this, we focus on three aspects which can potentially make the proposed method for adaptive monitoring popular among field scientist:

(I) Most of the previously discussed adaptive monitoring studies used stratified sampling to allocate samples over the study area. We extend the methods proposed in these studies by considering several alternative sampling techniques that could be used. A range of sampling designs will be tested and it will be evaluated which of these sampling design is most suitable for adaptive monitoring to improve the detection rate. The selected sampling design should first of all lead to higher detection rates. Additionally, the sampling method should be easy to implement and easy to adjust for consecutive surveys. That a sampling design should be easy to adjust is important, for example, when the sample size changes over time or when certain sites turn out to be inaccessible.

(II) We will consider different types of auxiliary information that can be used to design and consecutively update the monitoring designs. Many previous studies have used ecological auxiliary information to design and update the monitoring design (Guisan et al., 2006; Stohlgren and Schnase, 2006; Crall et al., 2013). Ecological information is any type of information that can be used to predict the species' habitat suitability. The species' habitat suitability is an index of the likelihood that a species will settle in a specific type of habitat. Thus, an area with a high species habitat suitability will be preferred by the species to settle compared with other less suitable sites. Various models have been proposed in the literature to estimate the species habitat suitability based on a set of ecological predictor variables, such as logistic regression (Nelder and Wedderburn, 1972), Maximum Entropy (MaxEnt) (Phillips et al., 2004) and support vector machines (Vapnik, 1999). These methods are more commonly known as Species Distribution

Models (SDMs).

A common assumption with SDMs is that the modelled species is in near equilibrium with its environment (Elith and Graham, 2009; Royle et al., 2012). A species is in equilibrium with its environment if the species is distributed proportional to the estimated habitat suitability. In other words: The species is expected to occupy all sites that have a habitat suitability that is adequate for that species to settle in. Many invasive species are not in equilibrium with the environment of the study area (Robertson et al., 2004; Guisan and Thuiller, 2005), as they often do not occur at all the locations with the highest estimated species habitat suitability. For example, if a species with a high affinity for a certain type of crop invades a new host area, then a SDM will assign a high species habitat suitabilities to sites containing that type of crop. However, especially with recent invasions, not all the sites containing that type of crop will be invaded immediately. In that case, the species are not yet in equilibrium with the habitat. Hence, it is important to realise that ecological auxiliary information, or auxiliary information based on SDMs should be interpreted primarily as a measure of species habitat suitability rather than a measure of current species prevalence.

A different type of information that might provide additional information about the current species distribution is spatial information. Spatial information is information that uses the observed geographical information about the invasive species. For example, a list of sites, in the format of latitudinal and longitudinal coordinates of each observed individual of the species (and sometimes also absence). Based on Tobler's First Law of Geography which states that "Everything is related to everything else, but near things are more related than distant things" and the idea behind spatial autocorrelation, which is a measure to which a set of spatial sampling points tend to be correlated or clustered together in space (positive spatial autocorrelation) or dispersed in space (negative spatial autocorrelation), one can try to make use of spatial information to estimate where an invasive species is likely to be present. Parnell et al. (2010), for example, illustrates by means of a simulation study that knowing the location of an invasive species is the key factor in designing eradication strategies. Hence, incorporating the observed spatial auxiliary information of the sites that are known to be already invaded is important to predict the species density. Therefore, spatial information is another type of information

that can be important when updating the monitoring design. One way to incorporate spatial information into the proposed method for adaptive monitoring is illustrated in chapter 5 and 6. In this context it is also important to discuss the differences between so called design based spatial sampling and model based spatial sampling (Wang et al., 2012, 2013; de Gruijter and ter Braak, 1990; Knotters et al., 1995; Brus and de Gruijter, 1997). Design based sampling methods are better known as probability sampling designs. Each sampling unit has a probability of being included in the sample and units are selected randomly with respect to those inclusion probabilities. Inferences about the total population are based on the sampled units and the set inclusion probabilities. In the design based methods the uncertainty of a sample depends on which sites are sampled and hence would completely disappear in the case each sampling unit in the sampling domain would be sampled. These methods often assume that the observed number of individuals is fixed for that unit, or at least that within the sampling domain the mean number of individuals per unit is fixed. However, in the case the number of individuals was a one-off realisation of an underlying stochastic process the true population size is hard to measure even if every unit is sampled. This is exactly what model based sampling methods assume, namely that the population at the time of a survey is only one realisation of the super population (which is the set of all the unique populations at different times). Model based sampling design hence try to estimate the parameters of this underlying stochastic process. An example of a techniques used for model based sampling is Kriging (Krige, 1951; de Gruijter and ter Braak, 1990). Kriging is the name of a set of kriging tools, developed in the field of geostatistics, to estimate properties at unsampled units based on the values of the sampled units. Geostatistics can be regarded as a collection of numerical techniques that deal with the characterization of spatial attributes, employing primarily random models in a manner similar to the way in which time series analysis characterizes temporal data (Olea, 2006). These geostatistical methods make use of the spatial autocorrelation to estimate the species density (or by extension any variable of interest) at unsampled locations. Based on a sample a map of the species distribution can for example be formed by Kriging, by using various interpolation techniques. Several of these distribution maps over time could be used to make inferences about the super population. In this thesis we will mainly focus on design

based sampling methods. However, various parts of this thesis, in particular Chapter 4, 5 and 6 could have greatly benefited from a more in-dept model based sampling approach. Additionally, in chapter 5 and 6, the usage of kriging methods could have been explored in more detail.

(III) Monitoring of invasive species often means collecting information about the species, by modelling the trend in the population size or distribution. This information can then be used to set up an invasive species management strategy, such as a eradication strategy. Additionally, the species monitoring can be used to evaluate the effect of an implemented eradication strategy and if necessary to adjust the eradication strategy. Even though detection and eradication are closely related to each other, many studies focus on improving the eradication strategy or on improving the detection of species (Bogich et al., 2008; Homans and Horie, 2011). One of the main problems with using these two separate approaches is that when focussing on improving the eradication rate, the search for newly invaded sites is reduced and vice versa (Bogich et al., 2008; Homans and Horie, 2011). However, few methods integrate both monitoring and eradication into the same application (Epanchin-Niell et al., 2012; Bogich et al., 2008). Or as Bogich et al. (2008) stated: ‘There is an urgent need to design and test joint monitoring and management strategies in order to achieve early detection and, in turn, provide more effective control’. In this thesis we aim at improving the detection rate of a monitoring design by increasing the inclusion probabilities in high density areas. We will test if this can then be used as a potential eradication strategy as well, namely by eradicating species in those the areas that have been given higher inclusion probabilities. We will explore the idea of adjusting inclusion probabilities over time not only for monitoring but also as as potential eradication strategy. We will investigate if an adaptive monitoring design, especially one that achieves high detection rates could potentially be used as an eradication strategy and how?

In this thesis, we develop the ideas of the iterative method, as proposed in Stohlgren and Schnase (2006) and extended by others, to introduce a new and improved general framework for adaptive monitoring. This method will incorporate and address all three of the potential points of improvement: Selection of the sampling design, using spatial and ecological infor-

mation to update the monitoring design and the potential incorporation of an eradication strategy.

1.3 Thesis Outline

In this thesis a new general framework for adaptive monitoring is introduced. This method will focus on increasing the detection rates of invasive species over consecutive surveys. The structure of this thesis is as follows:

- 1** The new framework for adaptive monitoring is introduced in Chapter 2. A brief background will be given at each step and the points of improvement compared with previous adaptive monitoring methods will be discussed in greater detail.
- 2** The selection of an appropriate sampling design to use for the proposed adaptive monitoring method is done in Chapter 3.
- 3** Chapter 4 and 5 will illustrate how to adjust the monitoring design over time using available auxiliary information. Chapter 4 focusses on ecological information and Chapter 5 focusses on the usage of spatial information.
- 4** Applications of the new method for adaptive monitoring for non-stationary species populations will be evaluated in Chapter 6 and it is assessed if the idea of adapting the inclusion probabilities over time can be used to set-up an eradication strategy.

Chapter 2

A New General Framework for Adaptive Monitoring

In this chapter a new general framework for adaptive monitoring is introduced. The proposed method is based on the method introduced in Stohlgren and Schnase (2006) and which was used in other studies such as Guisan et al. (2006) and Crall et al. (2013). The differences of the proposed methodology compared with these existing methods are addressed. Moreover, we discuss the key elements that are believed to improve the performance of this method compared with the existing methods. The focus of this thesis is to achieve an improvement in the detection rates of the adaptive monitoring method. Additionally, the practical implementation of the monitoring design is also evaluated. For example, is it easy to implement the method in the field or can it be used for management strategies of invasive species?

2.1 Adaptive Monitoring: Proposed Algorithm

The proposed methodology for adaptive monitoring comprises the following steps:

Step 1 In the (unlikely) case that no prior information is available, select an uninformed sample in the first survey. Most of the time information will be available and the initial sample will be an informed sample and hence the algorithm will start at Step 2 (however for

illustrative, simulation related and computational purposes we will often simplify this in this thesis and assume that no prior information is available).

Step 2 Estimate the species distribution using the auxiliary information obtained from the previous sample(s). The auxiliary information can be one of the following types:

- (a) Ecological.
- (b) Spatial.
- (c) Ecological and Spatial.

Step 3 Use predicted species distribution to:

- (a) Select an informed sample.
- (b) Eradicate species.

Step 4 If more surveys are required return to Step 2, otherwise finish.

The proposed algorithm for adaptive monitoring is also illustrated in *Figure 2.1*.

As aforementioned in Chapter 1, this algorithm for adaptive monitoring aims to improve the existing adaptive monitoring methods for the following three principles:

- 1** Select the most suitable sampling design that allows for adaptive monitoring.
- 2** Use spatial information, in addition to ecological information.
- 3** Use the adaptive monitoring technique to design eradication strategies.

Intuitively, it is easy to understand how the second and the third points can lead to an improvement in the adaptive monitoring method compared with the existing methods. These points will be discussed in detail in Chapters 5 and 6. The importance of point one, the selection of a sampling scheme, is perhaps less obvious and needs some additional explanation. Most existing studies used stratified sampling to allocate the samples. New to our proposal, is the use of spatially balanced probability sampling designs to select the samples during the adaptive monitoring scheme. The subsequent sections explain the concept of spatially balanced probability sampling designs and introduce their use for adaptive monitoring.

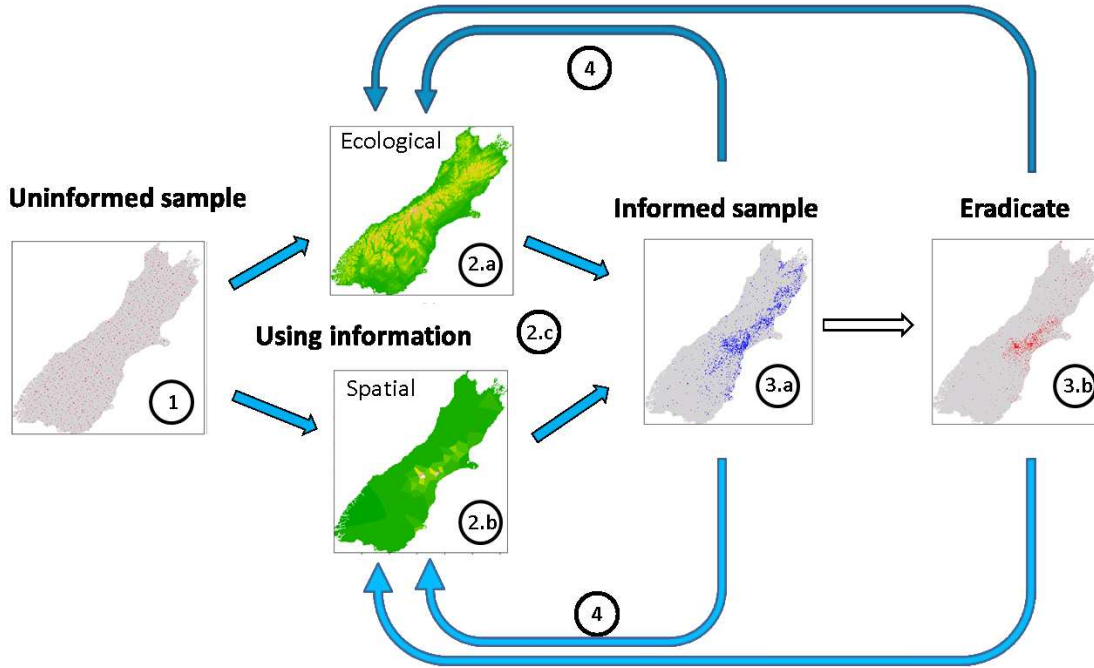


Figure 2.1: Illustration of the proposed method for adaptive monitoring. Each step of the proposed algorithm is added: 1. Select uninformed sample, 2. Use information to predict species distribution (a: Ecological, b: Spatial, c: Ecological+Spatial), 3. Select an informed sample (+ 3.b Eradicate), 4. Return to step 2. Note that this map of the South Island of New Zealand is chosen for illustrative purposes, and that spreading samples over the extent of the entire island is an oversimplification of any real applications.

2.2 Sampling

Once an invasive species is detected in a new host area, information about the state of the invasion can be collected. This information can for example be, information on the distribution of the invasive species or about the population size of a species within a study area. For most of the sampling designs and their applications in this thesis, the study area is partitioned into a grid of N quadrants of equal size, which we will call units. These N units constitute to the finite sampling frame, which is the set of units a sample can be selected from. The study variable to be recorded is the species abundance y_i within unit i . Let U be the set of all units in a population of size N

$$U = \{1, 2, \dots, N\}. \quad (2.1)$$

The parameter of interest is the total number of individuals of a species Y within the study area, which is

$$Y = \sum_{i \in U} y_i, \quad (2.2)$$

or the mean number of individuals per unit

$$\mu = \frac{Y}{N}. \quad (2.3)$$

Since it is difficult to visit every unit in a study area, a sample is selected from which inferences about the total area can be made. For example, suppose one is interested in the total number of possums in a certain national park. In that case, the national park is partitioned into N units. A limited number of n units are selected in such a way that inferences can be made about those N units. This could be done for example, by sampling each type of habitat. A sample is selected according to a certain sampling design or sampling design. Various sampling designs have been introduced for different surveying application (Cochran, 1977; Levy and Lemeshow, 2013). Selection of the most appropriate sampling design for ecological monitoring programs is important because of statistical (precision), practical or financial reasons (Gitzen et al., 2012). In this section, an introduction to sampling is given, to provide the reader with the necessary background. The reader is referred to Cochran (1977) and Levy and Lemeshow (2013) for a more comprehensive introduction on sampling. In the next chapter, several sampling designs are discussed and their suitability for adaptive monitoring is compared.

Sampling designs aim to obtain unbiased and precise estimate of the population total (Cochran, 1977; Thompson, 2006). The bias of an estimate is determined by the difference of the expected value of that estimate from the population parameter. For example, the difference of the estimated number of individuals \hat{Y} within the study area from the true number of individuals Y is $|Y - \hat{Y}|$ within that study area. Precision is the variance of the population estimate of the population size, where the smallest variance is preferred. Unbiasedness and precision of population estimation are both important when selecting a sample (Cochran, 1977). Biased parameter estimation, but with a high precision, leads to repetitive measurement of the same, but wrong, parameter estimate. Unbiased parameter estimation, but with low precision, leads

to an unbiased estimate but with a large variance.

2.2.1 Probability Sampling

Probability sampling designs are those sampling designs that assign each unit a probability of selection and thus have a form of randomness included when selecting the sample, so that statistical inferences can be made (Cochran, 1977; Tillé, 2006). This is compared with non-probabilistic or deterministic sampling designs which lack this randomness.

For a finite sample space of N units, unit i is assigned an inclusion probability π_i , which is the probability of including unit i in the sample, such that

$$\sum_{i \in U} \pi_i = n.$$

In this thesis we will attempt to use the concept of inclusion probabilities, generally used for population estimation, to hopefully increase the detection rate. Although the main focus is on increasing the detection rate, by adjusting the inclusion probabilities we will evaluate if our proposed method for adaptive monitoring can simultaneously be used for more precise population estimation as well. If all the inclusion probabilities can be accurately determined, then it is possible to produce unbiased estimates of the population total by weighting the sampled units according to their inclusion probabilities. A commonly used unbiased estimate for Y is the Horvitz-Thompson (HT) estimator (Horvitz and Thompson, 1952), given by

$$\hat{Y}_{\text{HT}} = \sum_{i \in U} \frac{y_i}{\pi_i} I_s(i), \quad (2.4)$$

where $I_s(i)$ is one if the i^{th} unit is selected in the sample and zero otherwise. The variance of \hat{Y}_{HT} for a fixed sample size can be written as

$$V(\hat{Y}_{\text{HT}}) = -\frac{1}{2} \sum_{i,j (i \neq j)} (\pi_{ij} - \pi_i \pi_j) \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j} \right)^2. \quad (2.5)$$

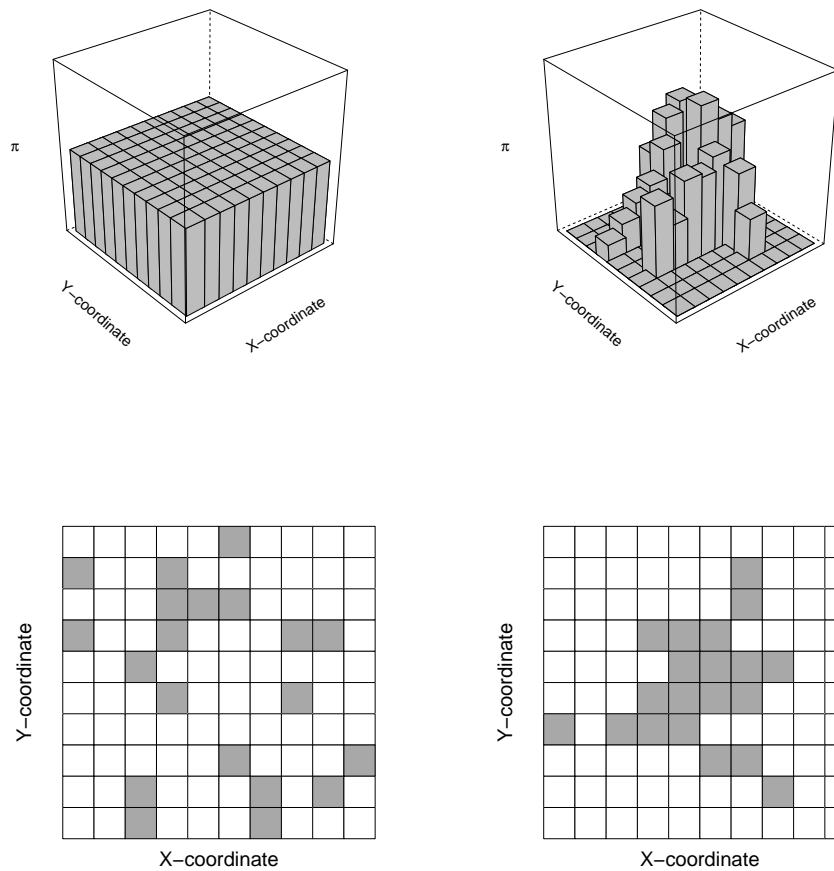
where π_{ij} is the second-order inclusion probability. A second order inclusion probability is the probability that both unit i and unit j are included in the sample. The variance, in 2.5, can be estimated from a sample using the unbiased Sen-Yates-Grundy estimator

$$\hat{V}_{\text{SYG}}(\hat{Y}) = -\frac{1}{2} \sum_{i,j (i \neq j)} \frac{(\pi_{ij} - \pi_i \pi_j)}{\pi_{ij}} \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j} \right)^2 I_s(ij),$$

where $I_s(ij)$ is one if both units i and j are included in the sample and zero otherwise.

If all units have equal inclusion probabilities, then $\pi_i = \frac{n}{N}$ for $i = 1, 2, \dots, N$. Hence all units have an equal chance of being selected in a sample. This is called equal probability sampling or equiprobable sampling. If the inclusion probabilities vary among the units, then some units have a higher chance of being selected in a sample as compared with other units. This is called unequal probability sampling. *Figure 2.2* illustrates the concept of both equal and unequal probability sampling. In this case, an equiprobable sample is selected using Simple Random Sampling (SRS) and an unequal probability sample is selected using a design called Balanced Acceptance sampling (BAS, see next section). A review of some classic sampling designs is given in Chapter 3.

Sampling designs are typically designed to obtain precise population estimates. In this thesis, the idea of probability sampling is used for adaptive monitoring. The reason for using probability sampling for adaptive monitoring is twofold. First, if one can assign those units with a high species densities a high inclusion probabilities, then a corresponding probability sample can be expected to achieve a high detection rate. A logical argument would be: ‘Why not simply select the n units with the highest inclusion probabilities’. This arguments leads to the second reason: Surveillance sampling. In this thesis, surveillance sampling is defined as the selection of a certain number of units from low density areas. This is done, for example, to ensure that monitoring programs do not miss out on unknown existing hotspots or newly invaded areas by the invasive species. Many studies discuss the trade-off between eradication and surveillance (Bogich et al., 2008; Rout et al., 2014). The need for surveillance sampling is one of the reasons why studies, in which the area is divided into a suitable and an unsuitable strata, allocate a proportion of the



(a) Equal probability sampling

(b) Unequal probability sampling

Figure 2.2: Example of an equiprobable sample (selected using SRS) and an unequal probability sample (selected using Balanced Acceptance Sampling (BAS), see next later). For both the equiprobable and the unequal probability case, $N = 100$, the sum of all π_i is equal to the sample size n , in this case $n = 20$. The top figures show the inclusion probabilities for each unit: the height of each bar indicates the inclusion probability for that unit. The bottom figures show the selected units based on the pre-set inclusion probabilities.

sample to the non-suitable strata, despite there being a low chance of finding the species. This is, for example, the case in Guisan et al. (2006) and Crall et al. (2013). Guisan et al. (2006), for example, used species distribution models to define a high suitability stratum and a low suitability stratum and allocated 50 percent of the sample size to the low suitability strata.

Deciding on the allocation of sampling effort between eradication and surveillance is often the point of discussion in the literature (Bogich et al., 2008; Rout et al., 2014). By selecting an unequal probability sample, we expect that most units are selected from expected high density areas, but generally some units from assumed low density areas are still added to the sample. This concept is further illustrated in the next chapters.

2.2.2 Spatial autocorrelation and Spatially Balanced Sampling

Natural resources are often observed to be spread over the sampling area following some kind of spatial autocorrelation (Stevens and Olsen, 2004; Theobald et al., 2007). Spatial autocorrelation is a concept that follows Tobler's First Law of Geography which states that 'Everything is related to everything else, but near things are more related than distant things.'. Legendre (1993) loosely defines spatial autocorrelation as 'the property of random variables taking values, at pairs of locations a certain distance apart, that are more similar (positive autocorrelation) or less similar (negative autocorrelation) than expected for randomly associated pairs of observations.'. In short, points closer together tend to have the same value for ecological variables compared with points that are more distant. Spatial autocorrelation is a spatial version of the more general concept of autocorrelation. Autocorrelation is a property of many ecological (and non-ecological) variables whose measurements are correlated in geographic space (spatial autocorrelation) or in time (temporal autocorrelation, which is the topic of interest in time-series analyses). In this thesis the focus is on spatial autocorrelation, although a valuable extension of this thesis would be to take into account the temporal autocorrelation of measured variables in sampled units as well. In this thesis, spatial autocorrelation means that nearby units tend to have similar species densities. An example of resources having a spatial autocorrelation are atmospheric pollutants, like traffic exhaust particles. The concentration of these particles will be higher in urban areas compared with more rural areas. For example, a bird species can locally have high densities close to certain preferable breeding areas, but the abundance will be lower in other areas.

The statistical problem that arises with spatial autocorrelation is that the value at a site can at least be partially predicted using the value from a neighbouring site. Hence, sampled units are no longer stochastically independent from each other, which is a common assumption in

statistical inference. Legendre (1993) says ‘This may come as a surprise to ecologists who have been trained in the belief that nature follows the assumptions of classical statistics, one of them being the independence of the observations. However, field ecologists know from experience that living beings in nature are distributed neither uniformly nor at random.’. In this thesis we assume that spatial autocorrelation is isotropic. This mean that the spatial autocorrelation only depend on the distance between two locations. In the case that things are more alike for shorter distances in some directions than in other directions the spatial autocorrelation is called anisotropic.

In the case of invasive species, one can assume that a unit which is located next to a unit with a known high density of the species will contain a high number of individuals as well. This is compared with a more distant unit. Hence, sampling nearby units provides similar information, and thus adds less information to the sample, compared with sampling more distant units. Furthermore, having large unsampled areas adds uncertainty to the sample. Therefore, it is desirable to distribute the selected sample over the study area such that the sample is evenly spread over the extend of the target resource, which, in this thesis, is the underlying species density. Such a sample is called a spatially balanced sample (Stevens and Olsen, 2004; Grafström and Tillé, 2013; Robertson et al., 2013). Spatially balanced sampling designs usually result in more efficient sampling designs by providing more information per sample unit, as each sample is distributed across the population. Therefore, they can improve the efficiency of population estimates by selecting a sample with few nearby units (Theobald et al., 2007).

In the case of isotropic spatial autocorrelation, it is possible to use Stevens and Olsen (2004) method of using Voronoi cells to define a spatially balanced sample. This is is illustrated in *Figure 2.3*. The Voronoi polygon ψ_i for sampled unit i is the collection of all the units in the study area that are closer to unit i than to any other sampled unit in the sample. Note that in the case of a small finite population, some units can be equidistant from the sampled unit. To avoid this problem one can assume at this stage, that the population is sufficiently large with very small units, for example, pixel sized units. Let v_i be the sum of the inclusion probabilities

of all the units within Voronoi polygon ψ_i , thus

$$v_i = \sum_{i \in \psi_i} \pi_i. \quad (2.6)$$

A sample is spatially balanced if v_i is approximately one for all Voronoi cells ψ_i . The test statistic introduced by Stevens and Olsen (2004) to measure the spatial balance of a sample is the variance in v , $\xi = V(v)$ with $v = \{v_1, v_2, \dots, v_n\}$. The smaller ξ , the higher the spatial balance of a sample. Note that in the case of an equiprobable design, all Voronoi cells are expected to have similar areas. Therefore, in the case of an equiprobable design, one can simply evaluate the variance in the areas of the Voronoi cells.

The difference between a spatially balanced sample and a sample with less spatial balance for an equiprobable design is illustrated in *Figure 2.3*. The points of spatially balanced sample are evenly spread, resulting in more equally sized Voronoi cells. Therefore, the variance among the areas of the Voronoi cells will be small. The sample with poor spatial balance has some points that are clustered and hence the variance of the Voronoi cells will be larger.

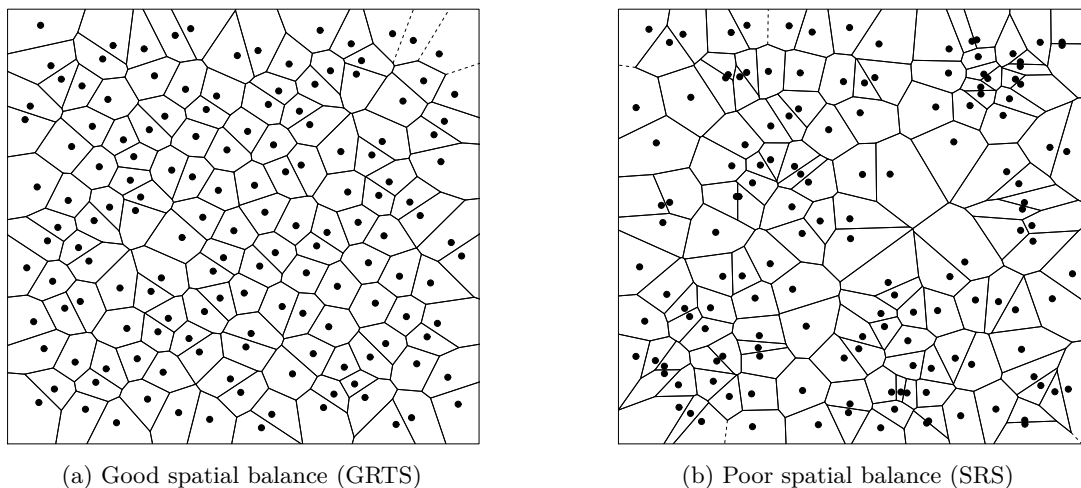


Figure 2.3: Illustration of a spatially balance sample (selected using generalised randomised tessellated stratified (GRTS), see next chapter) compared with a sample with less spatial balance (SRS). Note that the size of the Voronoi cells for the sample with high spatial balance is less variable compared with the sample generated by SRS.

Spatially balanced sampling designs are mainly used as a means to improve the precision of population estimates in the presence of spatial autocorrelation (Stevens and Olsen, 2004). Recently, Grafström and Lundström (2013) provided computational and theoretical evidence to support this claim that spatially balanced sampling designs can improve the precision of a sample. In their paper, they used a simulation study on some virtual datasets of natural resources to show that the variances of population estimates obtained by spatially balanced sampling designs are generally lower compared with the estimates of non-spatially balanced sampling designs such as SRS. More importantly for this thesis is that since spatially balanced sampling designs spread the sample evenly over the study area, this could also be used to increase the detection rate of a sample. For example, a sample selected by SRS might by chance not select units in a large section of the study area. In terms of surveillance sampling, this would mean that a potential invasion of this section would go unnoticed. A spatially balanced sample tends to spread out the sample more evenly over the study area. In the next chapter, several methods for spatially balanced sampling are introduced.

2.3 Adaptive Monitoring Using Probability Sampling

As discussed previously, one of the main advantages of selecting a probability sample is that the allocation of sampling units over the study will follow a similar trend as the pre-set inclusion probabilities. For example, as we will discuss in the next chapter, with stratified sampling one has to decide on an allocation scheme to distribute the sample over the strata. If the pre-set inclusion probabilities are high in areas of high species densities and lower elsewhere then more units will be selected in those high density areas and a small number of units will be selected in low density areas. However, how can one set these inclusion probabilities for each unit?

If the aim of the sampling design is to increase the detection rate, then expected high density units should be sampled more frequently and hence those units should be given higher inclusion probabilities. One way of doing this is by setting the inclusion probabilities proportional to the species density. If enough data are available then data driven computational techniques allow for

unit specific computation of the inclusion probabilities. Stratified sampling, for example, does not set unit specific inclusion probabilities, but gives all the units within the same strata an equal inclusion probability. However, for a single survey sufficient information is seldom available to accurately pre-set inclusion probabilities on the unit level. Therefore, it is difficult to accurately set the inclusion probabilities proportional with the true population density. In contrast, with monitoring, the repeated surveys provide increasingly more information over time, and this information can be used to update the inclusion probabilities. One can expect that if more information is available, then the computed inclusion probabilities can be expected to be more proportional with the true species density. In chapters 4, 5 and 6 we will illustrate how the species density information will be used to set unit specific inclusion probabilities. Because of simulation-related issues we will assume that the first survey is always uninformed and hence the sample will be uniformly distributed. As mentioned before, frequently it is the case that the initial sample is also an informed sample. After the first sample we will use statistical interpolation techniques (see chapter 4 and 5) to predict the species density at each unit in the sampling domain. Next, we can set the inclusion probabilities for each unit in the sampling domain using this predicted species suitability. By doing this, each unit will have an inclusion probability in the next survey to be selected to the sample. The proposed algorithm for adaptive monitoring uses two types of information to pre-set the inclusion probabilities: ecological information and spatial information. Using information to update the inclusion probabilities are discussed in Chapters 4, 5 and 6.

2.4 Discussion

In this chapter a general framework for adaptive monitoring has been introduced. This method is based on the iterative method for adaptive monitoring introduced by Stohlgren and Schnase (2006), which was used and modified by, for example, Guisan et al. (2006) and Crall et al. (2013). We developed our method by focussing on three main areas. Firstly, we will use a spatially balanced probability sampling design to select the samples. Second, both ecological as well as spatial auxiliary information will be used to update the sampling design. Third, this method for

adaptive monitoring can potentially also be used for eradication.

The concepts of probability sampling and of spatially balanced sampling have been introduced and an introduction on why adaptive monitoring would benefit from these methods has been given. Spatially balanced probability sampling can be expected to increase the detection rate, while selecting a surveillance sample to ensure that newly invaded areas will be detected as well. In the next chapter a selection of probability sampling methods and spatially balanced sampling methods are introduced and their performance are reviewed.

Chapter 3

The Selection of a Sampling Design for Adaptive Monitoring

Many previous studies on adaptive monitoring have used a stratified sampling design to select the units (Stohlgren and Schnase, 2006; Guisan et al., 2006; Crall et al., 2013). In this chapter, several other sampling designs are introduced and their suitability for (adaptive) monitoring is evaluated. We start by introducing a selection of common probability sampling designs. Next, several spatially balanced sampling design are introduced. Two spatially balanced designs are discussed in greater detail: Generalized Randomized Tessellation Stratified Sampling (GRTS)(Stevens and Olsen, 2004) and Balanced Acceptance Sampling (BAS) (Robertson et al., 2013). GRTS can be thought of as the current benchmark method for spatially balanced sampling. BAS is a recently introduced alternative to GRTS. The spatial balance of these designs is compared using a simulation study. Finally, we assess which sampling design can be useful for adaptive monitoring. This is done by testing which sampling method achieves the highest detection rate for a one-off (un)informed initial sample using simulation studies. Evaluation of the achieved detection rates for a smaller selection of sampling methods for consecutive surveys is done in the next chapters. Although the main focus is on detections rates, we will also evaluate if these sampling methods can increase the detection rate, without loss in precision or accuracy of common population estimates. Throughout this chapter, the practical (dis)advantages of the

proposed sampling designs are discussed.

3.1 Probability Sampling Designs

In this section, three probabilistic sampling designs are discussed, these are Simple Random Sampling (SRS), Stratified Random Sampling (StratRS) and Conditional Poisson Sampling (CP). All the considered sampling designs in this thesis are based on sampling without replacement. Note that there are many other designs that can select a probability sample. This is by no means a complete review. Note that when we introduce the estimators in the following sections for each discussed sampling design, we assume perfect detection. This is an assumption we will make throughout this thesis.

3.1.1 Simple Random Sampling

With SRS, each unit has an equal chance of being selected. To select a sample using SRS, units are randomly selected and added to the sample until the required sample size is reached.

Because each unit has an equal chance of being included in the sample, $\pi_i = \frac{n}{N}$ for each unit i . Therefore, we can rewrite the HT-estimator for SRS as

$$\hat{Y}_{\text{SRS}} = \frac{N}{n} \sum_{i \in U} y_i I_s(i). \quad (3.1)$$

The variance of this abundance estimator for SRS is given by

$$V(\hat{Y}_{\text{SRS}}) = N(N - n) \frac{\sigma^2}{n}, \quad (3.2)$$

with σ^2 being the finite population variance which is

$$\sigma^2 = \frac{1}{N - 1} \sum_{i \in U} (y_i - \mu)^2, \quad (3.3)$$

where μ is the population mean, the mean number of individuals per unit. When selecting a sample

of size n , an unbiased estimator of the variance for SRS is given by

$$\hat{V}(\hat{Y}_{\text{SRS}}) = N(N - n) \frac{s^2}{n}, \quad (3.4)$$

with s^2 being the sample variance given by

$$s^2 = \frac{1}{n - 1} \sum_{i \in U} (y_i - \bar{y})^2 I_s(i), \quad (3.5)$$

where \bar{y} is the sample mean number of detected individuals per unit. Proof that these estimators are unbiased can be found in Cochran (1977).

SRS is a straightforward design and easy to implement. SRS is often used as the benchmark methodology to compare the precision of other sampling designs with (Lohr, 2010; Levy and Lemeshow, 2013). This can be done by calculating the relative variance (RV), also known as the design effect or the efficiency rate. The RV is the ratio of the variances of the population estimators of two designs. Thus if SRS is the benchmark design:

$$\text{RV} = \frac{V(\hat{Y}_{\text{alternative.method}})}{V(\hat{Y}_{\text{SRS}})}. \quad (3.6)$$

If $\text{RV} > 1$ then the alternative design has a higher variance compared with SRS and thus that the alternative design has lower precision. If $\text{RV} < 1$ then the alternative sampling design has a smaller variance compared with SRS and thus has a higher precision. The RV can be generalized by using the variance of any other sampling design in the denominator.

3.1.2 Stratified Random Sampling

Equiprobable sampling designs, such as SRS, give each unit an equal probability of being selected. However many species populations are clustered or the species population has a strong spatial autocorrelation over the area. Stratified Random Sampling (StratRS) is a classic sampling design that is used to improve parameter estimation when dealing with those type of populations. To apply StratRS, the study area is partitioned into strata. A SRS sample is selected in each strata.

Stratification can be used for various reasons. If, for example, the stratification is done such that the values of the target variable in each stratum are as homogeneous as possible, thus by grouping those units that are as similar as possible, it is possible to obtain more precise estimates than the estimates of SRS (Cochran, 1977). For example, for a survey on invasive species, stratification may be based on the type of habitat, since the species density can be expected to be more homogeneous within the same type of habitat.

On the other hand, stratified sampling is often used to increase the spatial coverage and the spatial balance of a sample, compared with for example SRS. The use of StratRS to increase the spatial coverage of a sample will be discussed in more detail in the next section about spatially balanced sampling designs.

Unbiased population estimators are also available for StratRS. The estimated total number of individuals for StratRS when the study area partitioned into L strata is

$$\hat{Y}_{\text{StratRS}} = \sum_{h=1}^L \hat{Y}_h, \quad (3.7)$$

where \hat{Y}_h is the estimated number of individuals within stratum h thus

$$\hat{Y}_h = \frac{N_h}{n_h} \sum_{i=1}^{N_h} y_{ih} I_s(i), \quad (3.8)$$

where N_h is the number of sampling units in the h^{th} stratum and n_h is the selected sample size from that stratum. The variance of the total abundance estimator \hat{Y}_{StratRS} , can be written as

$$V(\hat{Y}_{\text{StratRS}}) = \sum_{h=1}^L N_h(N_h - n_h) \frac{\sigma_h^2}{n_h}, \quad (3.9)$$

with σ_h^2 the variance for the h^{th} strata given by

$$\sigma_h^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (y_{ih} - \mu_h)^2. \quad (3.10)$$

An unbiased estimator for this is given by

$$\hat{V}(\hat{Y}_{\text{StratRS}}) = \sum_{h=1}^L N_h(N_h - n_h) \frac{s_h^2}{n_h}, \quad (3.11)$$

where s_h^2 is the sample variance for the h^{th} stratum given by

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{N_h} (y_{ih} - \bar{y}_h)^2 I_s(i). \quad (3.12)$$

Proof that these estimators are unbiased can be found in Cochran (1977).

It is shown that StratRS will almost always improve the precision of a survey compared with SRS (Cochran, 1977). This is because the sum of the within strata variances is almost always smaller than the variance obtained by SRS. However, when strata are ly chosen, it is possible that the variance is not reduced (much) by stratification compared with SRS (Gitzen et al., 2012). This is discussed in more detail in Chapter 4. This can happen, for example, when stratification is implemented out of practical convenience instead of as a means to reduce the variation (Gitzen et al., 2012). An example of ‘practical stratification’ is stratification based on geographical borders, when for example each district has only a limited number of field scientists. In that case, the districts borders are the strata boundaries and the number of field scientists within each district could be used to determine the sample size within each district/strata. Practical stratification can also be useful when a guaranteed number of sampling units are required within each stratum. For example, in the case one would like to select ten units in each district. Practical stratification might result in lower financial costs. However, it does not ensure improvement in the precision of the survey design.

Setting the Strata

To increase the efficiency of StratRS when estimating the total number of individuals the sampling area should be partitioned into strata that are as homogeneous as possible (Cochran,

1977). This means that stratification should preferably be done based on the underlying species distribution. For example, the area could be stratified into a low species density strata, a strata of intermediate density and a high density strata. Unfortunately, most of the time the underlying species distribution is the variable of interest and is thus unknown. To circumvent this issue, auxiliary variables which are highly correlation with the species density are often used to base the stratification on (Lohr, 2010; Guisan et al., 2006; Crall et al., 2013). For example, the density of a population of invasive mosquito species could be correlated to the availability of water. Therefore, strata could be defined based on the proximity of water. Note that in the case stratification is used for other purposes than population estimation, for example mapping of a species population, the sampling intensity could be optimized in another way. For example one could focus the sampling intensity in those areas with the highest change in the density of the species.

Several algorithms have been designed to determine the strata boundaries and to determine the sampling allocation based on auxiliary information. Examples of these algorithms include: Cumulative Square Root Algorithm (Dalenius and Hodges, 1957), Ekman Algorithm (Ekman, 1959), Lavallée-Hidiroglou Algorithm (Lavallée and Hidiroglou, 1988) and the cube method (Tillé, 2011). These methods look for cut-off values in the auxiliary variable such that units can be combined into homogeneous group based on the level of the auxiliary variable. Many of these techniques have their background in disciplines other than ecology such as economics and agricultural sciences (Benedetti et al., 2010). In ecology, species distribution models and dispersal models have been used to set the strata boundaries (Guisan et al., 2006; Albert et al., 2010; Elith and Leathwick, 2009; Peterman et al., 2013). These models use environmental covariates to estimate the species habitat suitability. Most of these studies use the estimated habitat suitability to base the stratification on. These studies often partition the study area into two strata: A high suitability stratum and a low suitability stratum. In Chapter 4 the applications of species distribution models in designing sampling designs will be discussed in greater detail.

Deciding on the number of strata is another issue with stratification. Adding strata can improve the precision, but depending on the situation the addition of new strata must be traded off

against the additional time restraints and financial costs (Caughlan and Oakley, 2001). These time and financial restraints are issues that do not only occur with StratRS but also with almost any other type of sampling design for example for GRTS and BAS as well as for SRS. These practical issues will be outlined when introducing these sampling designs.

Allocation Sampling Effort

Once the study area is partitioned into strata the sampling units need to be allocated over the strata. The three most common sampling allocation schemes are: fixed allocation, proportional allocation and Neyman allocation.

Fixed allocation distributes a fixed proportion of the sample size to each stratum. The allocation is not based on the size of the strata but rather on other, often logistical or other convenient issues. For example, this can be based on the available number of volunteers in each stratum or based on the number of accessible sites in each stratum.

With proportional allocation, the sample size in each stratum is set proportional to the size of the stratum, thus

$$n_h = n \frac{N_h}{N}. \quad (3.13)$$

For example, if a stratum covers 40 percent of the study area, then 40 percent of the total sample effort will be allocated to that stratum. This type of stratification is convenient since it spreads the sample more uniformly over the study area compared with SRS (Gitzen et al., 2012).

Neyman allocation is an allocation scheme that allocates the sampling effort proportional to the population variance of each stratum (Neyman, 1934). Neyman allocation is given by

$$n_h = n \frac{N_h \sigma_h}{\sum_{k=1}^L N_k \sigma_k}. \quad (3.14)$$

More units will be selected in strata with higher variances. Unfortunately, σ_h^2 for each stratum is rarely known prior to sampling. A possible solution to obtain σ_h^2 is based on previous surveys or in the format of a small sample pilot study. In the case that these variances are known, or well approximated, Neyman allocation can lead to more precise population estimates. Optimal

allocation includes a cost function such that strata that are more time consuming or more costly to visit, will get allocated a relative smaller portion of the sampling effort.

3.1.3 Conditional Poisson Sampling

SRS is an equiprobable sampling design and each unit has an equal inclusion probability. With StratRS units within the same stratum have equal inclusion probabilities and a SRS is selected from each stratum. In the case that the study area is not stratified and units have different inclusion probabilities, an unequal probability sample can be selected. Many probabilistic sampling designs have been proposed in the literature that allow for unequal probability sampling. Brewer and Hanif (1983) list about 50 sampling schemes of this kind and recent reviews can be found in Tillé (2006) and Levy and Lemeshow (2013). Examples of unequal probability sampling designs are Probability Proportional to Size Sampling (pps) (Rosén, 1997) and Pivotal Sampling (Deville, 1998). Pivotal Sampling selects a sample in several steps. At each step a unit is selected and the inclusion probabilities of the remaining units are updated. In this thesis Conditional Poisson (CP) sampling (Hajek, 1981) is used as a reference method to compare it with other unequal probability sampling designs. This method is chosen because it was used as a reference method in Robertson et al. (2013) which introduced Balanced Acceptance Sampling (BAS, see later in this chapter).

Poisson sampling is a sampling design for which each unit is selected according to an independent Bernoulli trial. These Bernoulli trials use predefined probabilities p as the chance of being selected. In the case of probability sampling, these probabilities p for each unit can be obtained by $p_i = \frac{\pi_i}{\max(\pi_i)}$. Hence, with Poisson sampling a set of independent Bernoulli trials determine if the unit becomes part of the selected sample or not. CP sampling is a special case of Poisson sampling where the condition is added that a sample with a fixed sample size n should be selected. There are several versions of CP sampling. However, in this thesis, as well as in the BAS paper, CP-reject sampling as introduced in Hajek (1981) will be used. To achieve a fixed sample size, new Poisson samples are drawn until eventually a selected sample has the required sample size n . The fact that not all samples are accepted when selecting a CP sample affects the inclusion probabilities. A method to obtain adjusted inclusion probabilities for CP sampling

was given in Tillé (2006) and is

$$\pi_i^{(n)} = n \frac{\frac{p_i}{(1-p_i)}(1 - \pi_i^{(n-1)})}{\sum_{j=i}^N \frac{p_j}{(1-p_j)}(1 - \pi_j^{(n-1)})}. \quad (3.15)$$

Using 3.15, standard estimators such as the HT-estimator can be used for population estimation. CP sampling can be slow since in some situations many Poisson samples must be generated before a sample of size n is selected. Therefore, CP sampling is less suitable for large populations.

3.2 Spatially Balanced Sampling Designs

In this section, a selection of spatially balanced sampling designs is introduced. First, some standard methods that spread a sample evenly over an area are discussed, such as systemic sampling. Thereafter, more recent methods such as GRTS or BAS are introduced. Since these methods improve the spatial balance of a sample, are more flexible, and allow for unequal probability sampling, we will go into greater detail for these last methods.

3.2.1 Systematic Sampling

A classic method to select an evenly spread sample over a two dimensional area is systematic sampling (Bickford et al., 1963; Hazard et al., 1989). To begin, a grid is placed over the study area with l the number of intersection of the grid within the study area. Sampling starts by selecting a random intersection on the grid. Thereafter, every k^{th} intersection is selected such that $k = \frac{l}{n}$. Units including the selected intersections are selected. Alternatively, systematic sampling can be done by selecting a random unit in the study area. Thereafter, every k^{th} unit is selected such that $k = \frac{N}{n}$. Systematic sampling is illustrated in *Figure 3.1*. The method described here is the most well known form of systematic sampling. There are several variations of the base design, such as stratified random systematic sampling where a point is selected within each stratum and centric systematic sampling where points are selected at the centre of each stratum.

The efficiency of systematic sampling tends to depend on the spatial correlation of the species distribution. In the case that the species are uniformly distributed, one can expect a similar efficiency with systematic sampling as with SRS. In the abundance of spatial autocorrelation it is known that systematic sampling can have a higher efficiency compared with alternative sampling designs such as SRS or StratRS (Payandeh, 1970; Ripley, 1991; Stevens and Olsen, 2004; Wang et al., 2012, 2013). Because of this, systematic sampling is often implemented when, in the presence of spatial autocorrelation, a precise estimates of the number of individuals of a species is required or if mapping of the species distribution over the study area is wanted. Additionally, systematic sampling can be implemented for logistic reasons. It can for example be easier for field scientist to visit sites at regular intervals.

A potential problem could be that the species density follows a similar pattern as the pattern of the select units. For example, a ploughed field is sampled for seedlings of an invasive weed species. The sample units are placed in the same pattern as the furrows on the field. If you always count seedlings in the base of the furrow the variance will be underestimated. To solve this problem several variations of the classic system sampling algorithm have been proposed in the literature to perturb the purely systematic character of the selection of sampling units Olea (1984).

Although it is shown theoretically that systematic sampling can be more efficient than SRS a particular issue with systematic sampling is that there is no ‘design-based’ unbiased estimator. (Cochran, 1977; Wang et al., 2013). ‘Design based’ in this case means using probability based sampling theory. In short: Once the initial starting point is selected on the grid, the remaining points of the sample are part of a deterministic sequence. Because of this, the variance estimation can become biased. Furthermore, since the second order inclusion probabilities (which is the probabilities that two units are selected in the same sample) of neighbouring units are often (near) zero. It is therefore difficult to obtain an unbiased variance estimate. Several solutions have been proposed to solve this problem of having an unbiased estimator (Dunn and Harrison, 1993). One classic solution is to post-stratify the area. The area is stratified such that each stratum has (usually) two sampling units per stratum. The systematic sample is then treated as a StratRS sample. If similar sampling units are grouped within the same strata then the

within-strata variance will decrease and hence the overall variance of the estimator will decrease. In the case of anisotropic spatial autocorrelation this stratification is preferably done in the direction of the spatial autocorrelation, for example long narrow strata following the direction of the spatial autocorrelation McArdle and Blackwell (1989); Ripley (1991). Another method is to treat the systematic sample as a SRS and use the variance estimators for SRS. However this results generally in a conservative variance estimation.

Systemic sampling has a practically attractive design. Potentially, the fact that all the selected intersections on the grid need to be sampled to obtain a spatially balance sample can have its disadvantages. The design can lose its uniform spread when problems like non-response or premature endings arise. By non-response in this case we mean that when a sampling unit cannot be sampled in practise. For example, assume the unit is located on private property or the area is inaccessible for any other reason. Although the non-responses can have an effect on the spatial coverage effects on the precision of the population estimation are generally small. An example of a premature ending is when the sampling sequence starts in one corner but because of time or money related issues the survey is aborted before visiting all units. In that case, some areas end up being unsampled. This affects the spatial balance of the sample. Furthermore, it is difficult to add additional units to the initial sample while maintaining the spatial balance of the sample. Sampling additional units would be considered, for example, when extra funding becomes available to continue the survey. Of course, some of these issues also occur with other sampling designs. Finally, there is no obvious method to select an unequal probability sample with Systematic sampling.(Stevens and Olsen, 2004)

3.2.2 StratRS to Improve the Spatial Balance

Stratified sampling can also be used to improve the spatial balance of a sample. To do so, the study area is often tessellated into regular polygons such as squares, triangles or hexagons, or of any possible equally sized shapes. In fact, in the case of anisotropic spatial autocorrelation adjusting the shape of the strata according to the direction of the spatial autocorrelation

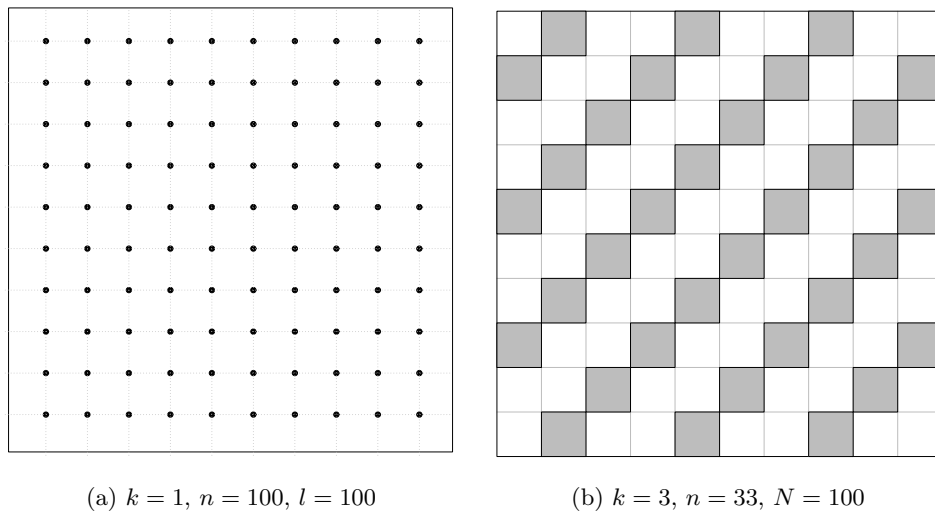


Figure 3.1: Illustration of Systematic Sampling.

can highly improve the spatial autocorrelation. For example, assume the species distribution increases from the east to the west (but there is no trend from south to north). In that case rectangular strata stretching from east to west can improve the efficiency of the sampling design. In this thesis we will assume only isotropic spatial autocorrelation hence we will not discuss this in detail. Next, one or two units are selected in each polygon. The tessellation of the area can also be based on natural boundaries or other arbitrary decisions. Note that, polygons that intersect the boundary of the study region can result in biased estimates, because some polygons will be smaller than other polygons (Lister and Scott, 2009). Also, it may be difficult to tessellate the area perfectly into the required number of equally sized polygons. An example of using StratRS as a means to increase the spatial balance of a sample is shown in *Figure 3.2*.

3.2.3 Generalized Randomized Tessellation Stratified Sampling (GRTS)

Stevens and Olsen (2004) presented a spatially balanced design called Generalized Randomized Tessellation Stratified sampling, or simply GRTS. The GRTS method is explained below. Compared with the previously described spatially balanced sampling designs GRTS has some desirable features. One advantage is that GRTS remains spatially balanced in the case of an

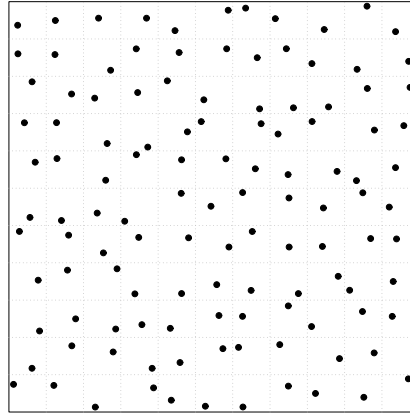


Figure 3.2: Illustration of Stratified Random Sampling with one unit per strata. In this case there are 121 strata.

advanced ending of a planned survey, as long as the units are visited in the specified order. Of course visiting sites in a specific order can be troublesome because of logistics, time and cost restraints. Additionally, GRTS adds the flexibility to add additional units to a sample while maintaining spatial balance. Again, this feature should be traded off against the same travel time problems. GRTS also allows for unequal probability sampling while remaining spatially balanced. These properties make GRTS increasingly popular among environmental scientists for designing sampling designs. GRTS has been applied to sample different types of natural resources such as aquatic resources (Hill et al., 2013; Lackey and Stein, 2013; Widmer et al., 2010), invasive plant species (Lemke et al., 2013) and chemical compounds (Dodder et al., 2012). GRTS has been cited by more than 380 papers. Examples of the use of GRTS are: Widmer et al. (2010) used GRTS to design a sampling designs for small-bodied fishes in a sand-bed river. Dodder et al. (2012) used GRTS to estimate the distribution of polybrominated diphenyl ethers in the Southern California Bight. Lemke et al. (2013) used GRTS in a study to estimate the effect of open surface mines on invasive plant species. Given this popularity of GRTS, we will use GRTS as the benchmark reference for spatially balanced sampling designs.

Methodology

The GRTS methodology is illustrated step by step in *Figure 3.3*. First, the study area is rescaled to the unit box. Next, the unit box is partitioned hierarchically into quadrants. This hierarchical partitioning is continued until the expected number of units in each quadrant is one. Each quadrant is assigned an unique hierarchical address. *Figure 3.3* shows the first three levels of hierarchical quadrant partitioning and the structure of the hierarchical addresses. For example, the shaded quadrant in *Figure 3.3* has the address 301.

After this partitioning, a transformation of all the addresses is applied. This is done by using a permutation algorithm on the separate digits of the address of each quadrant. This transformation introduces stochasticity to the sampling design. The reason for this is that otherwise the top left quadrant in *Figure 3.3* would always have the address 000 and the bottom right quadrant would always have address 333 and so on.

Next, the permuted addresses are reversed. For example, address 201 becomes 102. Finally, these reversed addresses are put in an increasing order on a line, also known as the real line. Each quadrant obtains a segment of equal length on the real line. A systematic sample is then selected using Brewer and Hanif (1983)' design. In short: To select a systematic sample from the real line the first segment is chosen as starting point.

Next, as with systematic sampling, segments are selected such that $k = \frac{N}{n}$, where n is the sample size and N the number of segments. Thus k is the number of segments on the real line between each selected segment. The permuted addresses of the selected segments are transformed back to the original addresses, using the same permutation algorithm but in the reverse direction. Finally, the selected quadrants can be mapped back using the original addresses.

The reason why GRTS selects well spread samples is because of the hierarchical stratification. Assume for example that there is only one level of partitioning, then there would only be 4 squares. If four units would be selected one unit would be selected in each partitioned rectangle. The remaining part of the algorithm is necessary to decide from within which rectangle the first unit would be selected and to allow GRTS to perform unequal probability sampling. In the

case the hierarchical partitioning has more than one level the logic behind the GRTS algorithm becomes more complex. The basic logic of GRTS is that the first four units are selected each within a different rectangle created by the first level of hierarchical partitioning, next the second set of four points will also be each selected from within a different rectangle of the first level of partitioning etc. The second level of partitioning ensures that the first four units that are selected within the first rectangle, created by the first level of partitioning, are each selected within a different rectangle created by the second level of partitioning and so on. An example of the GRTS sample is shown in *Figure 2.3a*.

Because of the way this algorithm works, the sample remains spatially balanced even in the case of advanced ending of the sample. This is of course only if the selected units are sampled in the order specified by the GRTS algorithm. In practise ensuring that points are sampled in such an order could come with a huge economical cost given the travel time and could turn out to be very ineffective.

Figure 3.3 illustrates the selection of an equiprobable sample using GRTS. Unequal probability sampling can be implemented as introduced in Brewer and Hanif (1983). The length of each quadrant on the real line is rescaled proportional to the inclusion probability of that unit. For example, unit A has an inclusion probability which is double the inclusion probability of unit B. Then the length on the real line of unit A will be twice as long as for unit B. *Figure 3.4* visualizes how unequal probability sampling is implemented in GRTS. Implementing unequal probability sampling with GRTS, units with higher inclusion probabilities will be more likely included in the sample, without the loss in spatial balance.

Population Estimation

GRTS has the ability to select an unequal probability sample. For population estimation standard design based estimators can be used, which allows for unbiased estimation, such as the HT-estimator. Second order inclusion probabilities are difficult to compute for GRTS. Furthermore, since it is a spatially balanced sampling design second order inclusion probabilities of neighbouring units are often (near) zero. Therefore, Stevens and Olsen (2003) derived the

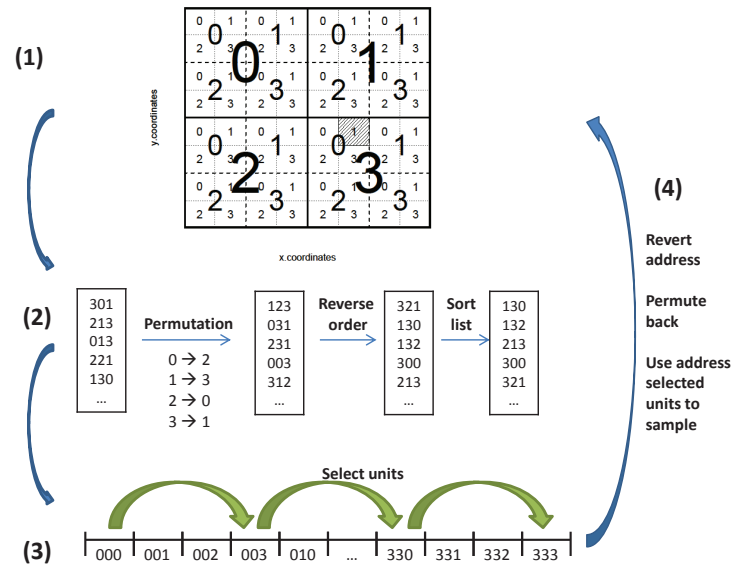


Figure 3.3: Step by step illustration of the GRTS methodology. (1) Rescale the study area to the unit box. The area is hierarchically partitioned into quadrants and assign hierarchical address. In this case, the first 3 levels of the hierarchical quadrant partitioning are shown. The address of the shaded unit is 301. (2) The addresses of all the quadrants are transformed using a digit specific permutation algorithm. Next, the transformed addresses are reverted. Finally, the reverted addresses are sorted and put on the real line. (3) Select units using Brewer and Hanif (1983)' method for systematic sampling on the real line. (4) Re-transform and re-revert the address of the selected units and map them back on to the study area.

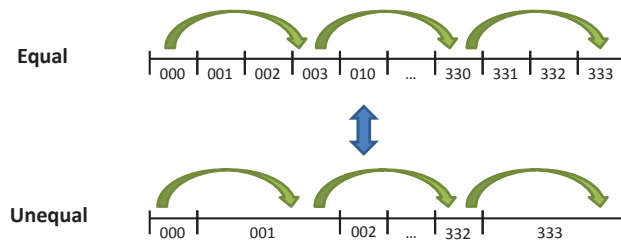


Figure 3.4: Illustration of Brewer and Hanif (1983)' method for systematic sampling on the real line for equal probability sampling and unequal probability sampling.

local mean variance estimator for the GRTS method. This approach of the variance estimator is very similar to the approach suggested for systematic sampling where two units are selected in

each stratum. The local mean variance estimator is given by

$$\hat{V}_{\text{NBH}}(\hat{Y}) = \sum_{i=1}^N \sum_{j \in D_i} w_{ij} \left(\frac{y_i}{\pi_i} - \bar{y}_{D_i} \right)^2 I_s(i,j), \quad (3.16)$$

where D_i is the neighbourhood to unit i that contains at least four observed units and \bar{y}_{D_i} is unit i 's neighbourhood mean. The weights, w_{ij} , decrease as the distance between unit i and unit j increases and satisfy $\sum_{j \in D_i} w_{ij} = 1$. For details on how to compute the weights, see Stevens and Olsen (2003). The local mean variance estimator is not an unbiased estimator but generally tends to overestimate the variance (this is compared with the observed simulated variance) as shown in Stevens and Olsen (2004), Grafström et al. (2012a) and Robertson et al. (2013). The local mean variance estimator performs best when the selected samples are well spread over the study area. This means that the local mean variance estimator works well for sampling designs that have small second order inclusion probabilities for units that are near each other, like spatially balanced sampling designs.

3.2.4 Balanced Acceptance Sampling (BAS)

Recently, Blair Robertson, Jennifer Brown, Trent MacDonald and the author of this thesis presented a new spatially balanced design called Balanced Acceptance Sampling (BAS) (Robertson et al., 2013). Jennifer Brown, Trent McDonald and Blair Robertson initiated the research on BAS when they attempted to extend GRTS to higher dimensions. GRTS can select spatially balanced samples in one and two dimensions. Examples of one dimensional problems are rivers systems or coastlines. Examples of two dimensional study areas are national parks or resources spread over a two dimensional area. In certain cases, it could be interesting to add a third (or even higher dimension) to the sample design. For example, for sampling of aquatic resources, depth could be added as a possible third dimension. In this section the BAS methodology will be explained in an easy and understandable manner, to make BAS more accessible to non-experts. Full details of BAS can be found in Robertson et al. (2013). My contribution to the BAS paper was mainly evaluating the performance of BAS and comparing its results with other spatially

balanced designs such as GRTS. Some of these results and additional simulation studies will be discussed later in this chapter.

Methodology

The Halton Sequence For SRS often pseudo random numbers are used to select a sample. Pseudo random numbers are the classical type of random numbers generated by (pseudo) random number generators using computers. The randomness of these pseudo numbers can lead to the selection of clustered units and samples with poor spatial balance. To select spatially balanced samples, BAS uses quasi random numbers to select the units. Quasi random numbers are typically uniformly distributed over an interval and are therefore generally not clustered. These quasi random numbers are generated by deterministic sequences.

The quasi random numbers used in BAS are known as the Halton sequence (Halton, 1960). The Halton sequence is a multidimensional extension of a van der Corput sequence. A van der Corput sequence is a sequence of points which are evenly distributed over the unit interval. Robertson et al. (2013) mentions that the k^{th} element in a van der Corput sequence with base p can be calculated using the radical inverse function $\phi_p(k)$ which is given by

$$\phi_p(k) = 0.\lambda_0, \lambda_1, \dots, \lambda_{K-1}, \lambda_K = \sum_{j=0}^K \frac{\lambda_j}{p^{1+j}}, \quad (3.17)$$

where $\lambda_j \in \{0, 1, \dots, p-1\}$ and K is a positive integer. For example, in the case of base $p = 2$, thus the binary system, the radical inverse for $k = 6$ is 0.011, or $\phi_2(6) = 0.011$. Converting $\phi_2(6)$ back to the decimal system with base 10, this gives $\frac{3}{8}$. The first ten members of the van der Corput sequence with base 2 are:

$$\{\phi_2(k)\}_{k=1}^{10} = \left\{ \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{5}{8}, \frac{3}{8}, \frac{7}{8}, \frac{1}{16}, \frac{9}{16}, \frac{5}{16} \right\}. \quad (3.18)$$

This example also illustrates an alternative and more intuitive method to reconstruct a van der Corput sequence. In the case of the sequence with base 2, start by dividing the interval $[0,1]$

in halves, then in fourths, eighths, etc..., To generate the sequence with base 3, we divide the interval $[0,1]$ in thirds, then ninths, twenty-sevenths, etc., which generates

$$\{\phi_3(k)\}_{k=1}^{10} = \left\{ \frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{4}{9}, \frac{2}{9}, \frac{5}{9}, \frac{8}{9}, \frac{1}{27}, \frac{24}{27}, \frac{10}{27} \right\}. \quad (3.19)$$

The Halton sequence is a combination of Van de Corput sequences, using prime numbers as the bases for the van der Corput sequences. Thus, the d -dimensional Halton sequence is

$$\mathbf{x}_k = \left(\phi_{p_1}(k), \phi_{p_2}(k), \dots, \phi_{p_d}(k) \right), \quad \text{for } k = 1, 2, 3, \dots, \quad (3.20)$$

where $p_1 = 2, p_2 = 3$ and p_j is the j^{th} prime.

As an example, assume we would like to select a sample of size seven in the unit square. Start by selecting the first two dimensions of the Halton sequence, the first to be based on prime number two and the second to be based on prime number three. Next, we pair them, to get a sequence of points in a unit square:

$$\left(\{\phi_2(k)\}_{k=1}^7, \{\phi_3(k)\}_{k=1}^7 \right) = \left\{ \left(\frac{1}{2}, \frac{1}{3} \right), \left(\frac{1}{4}, \frac{2}{3} \right), \left(\frac{3}{4}, \frac{1}{9} \right), \left(\frac{1}{8}, \frac{4}{9} \right), \left(\frac{5}{8}, \frac{7}{9} \right), \left(\frac{3}{8}, \frac{2}{9} \right), \left(\frac{7}{8}, \frac{5}{9} \right) \right\} \quad (3.21)$$

Mapping a Halton point on a two dimensional area is straightforward. The first dimension becomes after rescaling the x-coordinates of the sample. The second dimension becomes after rescaling the y-coordinates of the sample. This example only considers the first two dimensions of the Halton sequence. However, BAS can be extended to higher dimensions by using additional Halton sequences based on other prime numbers.

The Halton sequence is a deterministic sequence. This means that unless stochasticity is introduced, the same units will be selected for each sample. Robertson et al. (2013) therefore introduced the random start Halton sequence. Simply stated: The random start Halton sequence works as follows: skip the first u_d elements of the d^{th} Halton dimension. u_d can be different for each dimension.

Thus, the random start Halton sequence for d dimensions is

$$\mathbf{x}_k^{\text{RandomStart}} = \left(\phi_1(u_1 + k), \phi_2(u_2 + k), \dots, \phi_d(u_d + k) \right) \quad \text{for } k = 1, 2, 3, \dots \quad (3.22)$$

where $p_1 = 2, p_2 = 3$ and p_d is the d^{th} prime and $u_i \in \mathbb{N}$ for $i = 1, 2, \dots, d$. In this thesis all BAS samples are selected using the random start Halton sequence, with u_i an integer, randomly selected between zero and 500,000.

Selecting a sample using BAS This is best explained by means of an example. Suppose that a species population is distributed over a study area as shown in *Figure 3.5a*. The population density increases linearly with increasing x-coordinates and is uniformly distributed over the y-coordinates. The inclusion probabilities are set proportional to the species density, illustrated in *Figure 3.5b*. To select a sample, first generate a sufficiently long list of Halton points in $A + 1$ dimensions, where A is the number of dimension of the study area. In this thesis, we will only consider two dimensional study area, thus $A = 2$. The first two dimensions of each Halton point will be used to define the x- and y-coordinates. These coordinates refer to the unit containing these coordinates. The third dimension will be used for acceptance/rejection sampling. We will call this third dimension the acceptance/rejection dimension. BAS selects a sample by adding units to the sample using acceptance/rejection sampling (von Neumann, 1951) until the required sample size is reached.

To illustrate acceptance/rejection sampling we will call the grey area under the inclusion probability surface the acceptance area and the area above the surface the rejection area. To improve the computational efficiency of acceptance/rejection sampling the inclusion probabilities are rescaled such that the rescaled inclusion probabilities p_i over all units i has a maximum of one by

$$p_i = \frac{\pi_i}{\max(\pi_i)}. \quad (3.23)$$

The rescaled inclusion probabilities p_i are illustrated in *Figure 3.5c*. We can see that the acceptance area has become relatively bigger as compared with the rejection area. Acceptance/rejection sampling simply selects all these Halton points that fall within the acceptance

Table 3.1: Example of acceptance/rejection sampling for BAS

	1 st 3 Halton dimensions			Acceptance/Rejection	
	x	y	Accept/Reject	p_i	Decision
point 1	0.84	0.23	0.33	0.67	Accept
point 2	0.63	0.62	0.85	0.49	Reject
point 3	0.11	0.78	0.02	0.12	Accept

area.

For example, *Table 3.1* illustrates acceptance/rejection sampling for three Halton points, based on the pre-set inclusion probabilities illustrated in *Figure 3.5*. Note that these three points are chosen for illustrative purposes. The x- and y-coordinates of the first point, based on the first and second Halton dimension, are (0.84,0.23). The value of p_i at those coordinates is 0.67. Since the third Halton dimension of 0.33 is less than 0.67, the unit containing these coordinates is therefore be accepted and added to the sample. The second point, has the coordinates (0.63,0.62). For these coordinates $p_i = 0.49$. The third Halton dimension for this second point is 0.85. Since 0.85 is higher than 0.49, this second point is rejected and discarded from the sample. Finally, following the same reasoning, the third point is accepted and added to the sample. Units are added to the sample until the required sample size is reached. A sample of size 25 selected based on the pre-set inclusion probabilities is shown in *Figure 3.5d*.

Note that when the inclusion probabilities are not rescaled, acceptance/rejection can be performed on the original inclusion probabilities π_i . However, as the difference between *Figure 3.5b* and *Figure 3.5c* suggests, fewer Halton points will fall in the acceptance area. Hence, more Halton points will have to be evaluated to select the required sample size. Therefore, the computational efficiency will be lower.

Population Estimation

A method to compute exact first order inclusion probabilities for BAS is given in Robertson et al. (2013). In the same paper an alternative method was also proposed to approximate the first

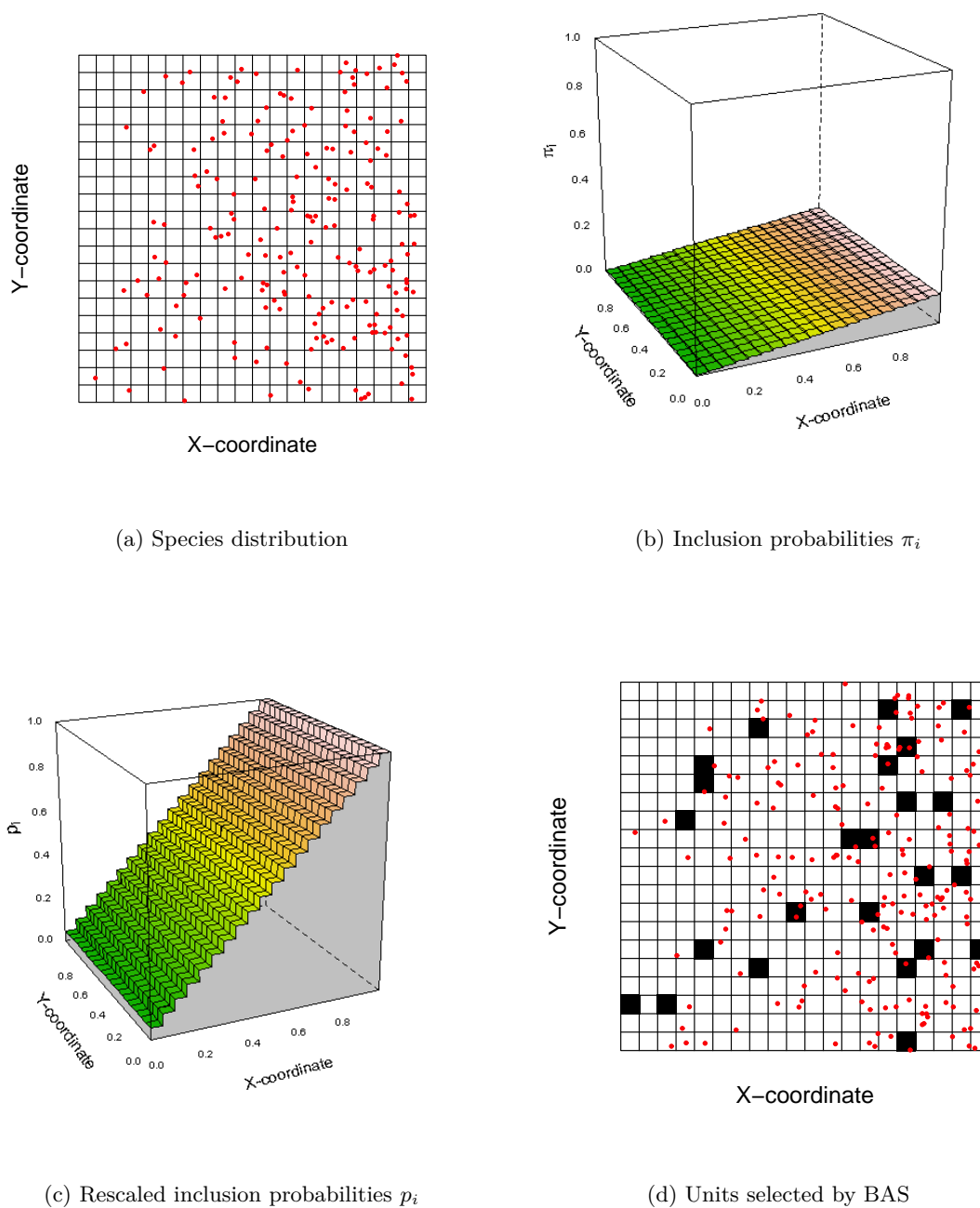


Figure 3.5: Illustration of sampling using BAS

order inclusion probabilities. For example, Robertson et al. (2013) showed by using Monte Carlo simulations that when using this approximation method for equal probability sampling $\pi_i \approx \frac{n}{N}$. Given these first order inclusion probabilities the HT-estimator and its variance estimator can be used for population estimation. This variance estimator can be estimated from a sample using the Sen-Yates-Grundy estimator. However, the Sen-Yates-Grundy estimator can be unstable for spatially balanced sampling designs, because it is common for spatially balanced designs to have second order inclusion probabilities that are close to zero for units that are close in distance. Robertson et al. (2013) found that, just like the other spatially balanced sampling designs such as GRTS or systematic sampling, in some simulation studies the second order inclusion probabilities were (near) zero. Therefore, they proposed the usage of the local mean variance estimator which Stevens and Olsen (2004) also proposed for the GRTS method.

3.2.5 Other Spatially Balanced Sampling Designs

Similar to BAS, other spatially balanced sampling designs were developed to select spatially balanced samples in more than two dimensions. For example, Grafström et al. (2012a) proposed a design called Spatially Correlated Poisson Sampling (SCPS). The SCPS selects sampling units sequentially and uses weights to preclude unobserved units close to previously observed units from being selected. More recently, the Local Pivotal Methods 1 and 2 (LPM1 and LPM2) (Grafström et al., 2012a) were introduced. These designs also select a spatially balanced sample from finite populations in multi-dimensional space. The Local Pivotal Methods are adjusted versions of the pivotal method introduced by Deville (1998) for spatially balanced sampling. In short: Units are selected sequentially. After each step, the inclusion probabilities are updated for a pair of units such that the sampling outcome is decided for at least one of the two units. The Local Pivotal Methods (LPM1 and LPM2) ensure, by decreasing the inclusion probabilities of those units near the latest selected units, that a more spatially balanced sample will be selected. Without going into the details: The LPM1 design has greater spatial balance, but LPM2 is simpler and faster (Grafström et al., 2012a). The Local Pivotal Methods can only sample discrete populations and can become computationally too expensive with increasing sample sizes.

3.2.6 Test for Spatial Balance

The test statistic introduced by Stevens and Olsen (2004) to measure the spatial balance of a sample, in the presence of isotropic spatial autocorrelation, is $\xi = V(\mathbf{v})'$ with $\mathbf{v} = \{v_1, v_2, \dots, v_n\}$. The smaller ξ , the higher the spatial balance of a sample. To compare the spatial balance of two sampling designs, we compare the ratio in ξ

$$\xi(\text{ratio}) = \frac{\xi(\text{design A})}{\xi(\text{design B})} \quad (3.24)$$

. If $\xi(\text{ratio}) = 1$ then both designs have equal spatial balance, if $\xi(\text{ratio}) < 1$ then design A has greater spatial balance and if $\xi(\text{ratio}) > 1$ then design B has greater spatial balance.

In the case of an equiprobable design, for a spatially balanced sample, the points will be evenly spread. Thus Voronoi cells are expected to have similar areas. Therefore, the variance among the areas of the Voronoi cells will be small. A sample with poor spatial balance, has some points that are clustered. Thus, the variance in the areas of the Voronoi cells will be higher. Therefore, in the case of equiprobable sampling, one can simply replace ξ by the variance of the areas of the Voronoi cells to obtain the test statistic for spatial balance.

Simulation Study

A simulation study was performed to compare the spatial balance of several sampling designs. For each sampling design 1000 samples were selected. For this simulation study points are selected in the unit box instead of selecting units from a discrete population. All sample sizes ranging from five to 250 were evaluated. To compare the spatial balance between the designs, the $\xi(\text{ratio})$ between the different methods were computed. The evaluated designs were SRS, GRTS and BAS. The results of this simulation study were published in Robertson et al. (2013) and the selection of the methods was based on the methods that were compared in the study that introduced GRTS (Stevens and Olsen, 2004). Systematic sampling was not included in the simulation study mainly because it can be difficult to select each tested sample size using systematic sampling in a easy way. However, it could have been worthwhile to include systematic sampling by, for example,

randomly selecting n polygons (on the grid) and select one point in each selected polygon, hence insuring a good spatial balance of each selected sample. In hindsight, this would have strengthened the conclusion of this and other simulation studies. LPM1 and LPM2 were also not included in this comparison. The reason for this is that these methods can only select a sample from a predefined discrete set of units. Therefore these methods cannot select a set of points in the unit square. This is a disadvantage of LPM1 and LPM2 (Robertson et al., 2013). As a possible solution to this problem, one could consider performing this simulation for a finite population. This could be done by tessellating the study area using a very fine grid (pixel size). Next, LPM1 or LPM2 could be used to select a number of units from this tessellated area. This is in fact similar to what GRTS does to select points in a box. However, for LPM1 and LPM2 this would be computationally too intensive. These designs only work efficiently for smaller population sizes.

The results are visualized in *Figure 3.6*. The $\xi(\text{BAS}/\text{SRS})$ curve and the $\xi(\text{GRTS}/\text{SRS})$ curve had consistently values of less than one for each sample size. This indicates that SRS has the lowest spatial balance for each evaluated sample size compared with GRTS and BAS. The $\xi(\text{BAS}/\text{GRTS})$ curve had values of less than one for each sample size. This shows that BAS has a greater spatial balance compared with GRTS for all evaluated sample sizes. The peaks in the $\xi(\text{BAS}/\text{GRTS})$ curve, for example for $n = 16$ and $n = 64$, can be explained by the algorithm of the GRTS methodology (Robertson et al., 2013). Recall that the GRTS methodology is based on hierarchical partitioning of the study area into quadrants. This partitioning splits the sampling area into $4, 16, \dots, 4^q$ squares of equal size. If $n = 4^q$, GRTS will have one point in each of these squares. From *Figure 3.6* we can see that the ratio of the variance between BAS and GRTS was the highest, so in favour of GRTS for sample sizes $4^2 = 16$ and $4^3 = 64$ and this ratio was increasing towards $4^4 = 256$. However, the maximum ratio between BAS and GRTS was 0.78, so even at those optimal sample sizes for GRTS, BAS still had greater spatial balance.

Adding Additional Units

In the previous simulation study for each repetition and at each sample size a new set of units was selected. However, a feature of both GRTS and BAS is the flexibility to add additional

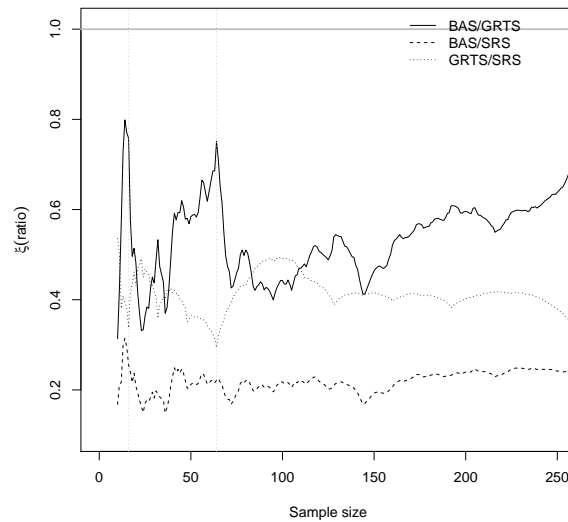


Figure 3.6: Results of the simulation study to compare the spatial balance for BAS, GRTS and SRS for using different sample sizes. For each repetition and for each sample size a new sample was generated. The results are represented as a ratio $\xi(\text{new/old})$. A ratio of one indicates equal variance and thus equal spatial balance, a ratio of less than one indicates greater spatial balance for the design in the numerator and vice versa.

units to a sample while maintaining the spatial balance. This is useful, for example, if additional funding becomes available to sample extra units, or if one decides to extend the survey by an additional year. In those cases, selection of additional sampling units is required. To review how adding units, to an already selected sample, affects the spatial balance of a BAS sample, a similar simulation study was performed as conducted in the previous section. In the previous simulation study a new sample was generated for each sample size. For example, if a sample was selected of size n then $n+1$ new points were selected for the next sample of size $n + 1$. For the simulation study in this section, sampling points were added one-by-one to the sample. The initial samples had a sample size of five and points were added one by one until a maximum sample size of 250.

The results of this simulation study are shown in *Figure 3.7*. There was no change in the $\xi(\text{BAS/SRS})$ curve compared with the previous simulation study. This indicates that BAS does

not lose its spatial balance when additional points are added to the sample. The $\xi(\text{GRTS}/\text{SRS})$ curve and $\xi(\text{BAS}/\text{GRTS})$ curve however, did follow a different trend compared with the previous simulation study. For smaller sample size, the $\xi(\text{BAS}/\text{GRTS})$ curve was slightly lower compared with the previous simulation study, however for higher sample sizes the curve increased and the values became more similar compared with the previous simulation study. This means that the spatial balance of GRTS is more affected by adding additional points to a sample when sample sizes are small. This can be explained as follows: For both GRTS and BAS gaps are filled in when points are added to the sample. However, it seems that for smaller sample sizes BAS does this more efficiently (relative higher spatial balance) compared with GRTS. However, once the sample size reaches a certain sample size one can expect that all gaps are more or less already filled in and that the effect of adding one point to the study will have less of an effect on the total spatial balance of the sample.

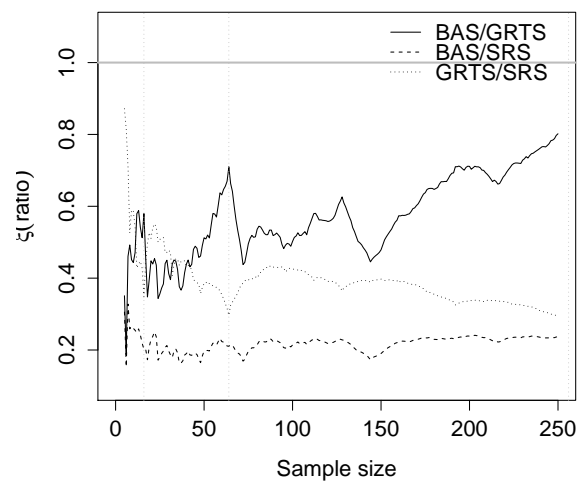


Figure 3.7: Results simulation study to compare the spatial balance for BAS, GRTS and SRS for different sample sizes. Units were sequentially added one-by-one to a sample. The results are represented as a ratio $\xi(\text{new}/\text{old})$. A ratio of one indicates equal variance and thus equal spatial balance, a ratio of less than one indicates greater spatial balance for the design in the numerator and vice versa.

Dealing With Non-Response/skipping points

An example of non-response is when a unit that got included in the initial sample cannot be sampled in practice, e.g. a selected unit might be located in inaccessible terrain or on private property. In this case certain sampling point are skipped and additional units have to be added to the sample to meet the required sampled size. In other words some points in the generated list of Halton points are skipped and units are added to the sample to meet the required sample size. In the previous simulation study it was evaluated how adding points to the sample would effect the spatial balance of the final sample. This simulation study is similar, however in addition to adding points to a sample, we skip a certain number of points and add points. This simulation study is important since GRTS and BAS provide a sequence of points which ideally should be visited in a fixed order to ensure spatially balanced samples.

To evaluate the effects of skipping points in the Halton sequence, due to non-response, on the spatial balance of the final sample a simulation study was performed. This simulation study is similar to the previously conducted simulation study. However, instead of selecting a sample in the unit box, a 20 percent non-response area was added inside the unit box. The distribution of the non-response area is illustrated in *Figure 3.8*. Voronoi tessellation was based on the final set of selected points, thus not including those points that fell within the non-response areas. The area calculation of the Voronoi cells excluded the non-response areas. This was done since these areas were assumed to be impossible to visit and should therefore not be included in the sampling population.

The results of this simulation study are shown in *Figure 3.9*. Both the $\xi(\text{BAS}/\text{SRS})$ curve and $\xi(\text{GRTS}/\text{SRS})$ curve had values less than one for all sample sizes. This indicates that BAS and GRTS have better spatial balance compared with SRS when 20 percent of the visited units would be for example turn out to be impossible to visits. BAS performed better compared with GRTS for almost all tested sample sizes. However, $\xi(\text{BAS}/\text{GRTS})$ is not as low as compared with the previous simulation study. This indicates that BAS will likely be more affected by skipping Halton points as compared with GRTS. This can possibly be explained by the fact that BAS has a better spatial balance when there is not a non-response area compared

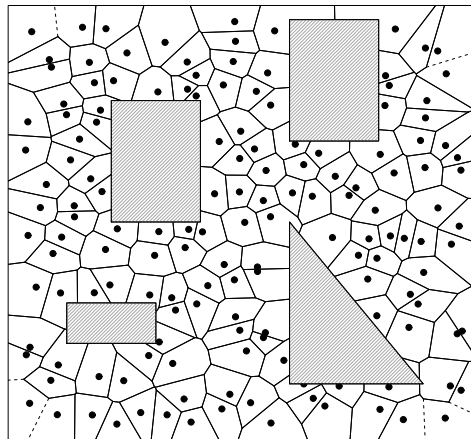


Figure 3.8: Illustration of the distribution of the 20 percent non-response area (grey areas). The Voronoi tessellation for a sample excluding the non-response area is added.

with GRTS. Since skipping points can have an impact on the spatial balance of a sample, this effect will be relatively bigger for those designs which have greater spatial balance in the case that no points are skipped. For example, there does not seem to be any effect on the spatial balance of SRS compared with the other designs. This is since the sampling process for SRS is completely at random and hence non-response will not effect the spatial balance of SRS.

3.3 The Selection of Sampling Design

In the previous sections, several sampling designs have been introduced. In this section, we will evaluate which sampling design is suitable for adaptive monitoring. Achieving high detection rates is the first and most important selection criteria. Achieving a high detection rate means knowing where the invasive species are. While we focus on achieving high detection rates in this thesis, it is still interesting the gain knowledge about the population size of the invasive species. Therefore, the second selection criteria is obtaining precise population estimates.

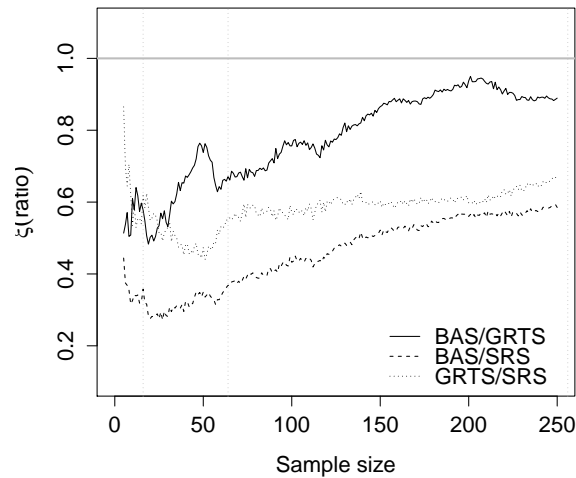


Figure 3.9: Results simulation study to compare the spatial balance for BAS, GRTS and SRS for different sample sizes with non-response area of 20 percent. Units were sequentially added one-by-one to a sample. The results are represented as a ratio $\xi(\text{new/old})$. A ratio of one indicates equal variance and thus equal spatial balance, whereas a ratio of less than one indicates greater spatial balance for the design in the numerator and vice versa.

3.3.1 Detection Rate Initial Sample

In this section we evaluate which sampling designs achieves the highest detection rate. In the case of the proposed methodology for adaptive monitoring, the initial sample, unless prior information is available, is an equiprobable sample. Therefore, first we will look at the achieved detection rate of different sampling designs when selecting an equiprobable sample. After the initial survey, information will be used to adjust the inclusion probabilities over time. Based on these adjusted inclusion probabilities, an unequal probability sample will be selected. Hence, a second evaluation will evaluate which sampling design achieves the highest detection rate for unequal probability sampling. Both hypotheses were reviewed by means of a simulation study. For these simulation studies, a dataset about a Rockfish population in Alaska was used. This dataset is adopted from Su and II (2003). The original population is highly clustered, as illustrated in *Figure 3.10a*. The study area consists out of 40 by 40 square units, $N = 1600$. The number of units occupied by Rockfish is 405.

Surveys using no prior information: Equiprobable Sampling

A Monte Carlo simulation study was performed for each sampling design, with the number of simulation $m = 1000$. The number of simulations of 1000 was chosen similar to the number used in Stevens and Olsen (2004) and Grafström et al. (2012a). This was due to computational limitations. GRTS and especially LPM1 and LPM2 can be very slow to run even on relatively fast cluster server computers (some simulations in the coming chapters took several days to finish). Each iteration a sample was selected and the detection rate D_j was calculated. These simulations were repeated for three different sample sizes, respectively: five, ten and 15 percent of the units were sampled. The sampling designs that are evaluated are: SRS, LPM1, LPM2, GRTS, BAS, StratRS with four equally sized strata (StratRS_4), StratRS with 16 equally sized strata (StratRS_16) and StratRS with 80 equally sized strata (StratRS_80). The strata partitioning for StratRS_4, StratRS_16 and StratRS_80 are illustrated in *Figure 3.11*. For all three stratified sampling designs, proportional allocation was conducted. Note that the strata sizes were selected such that an equal number of points could be selected in each strata. Because of this when selecting five percent of the units ($n=80$ units) StratRS_80 will have one unit per stratum. CP sampling was not included in the simulation study since in the case of equiprobable sampling CP is identical to SRS. For each sampling design and for each sample size the mean detection rate \bar{D} was calculated

$$\bar{D} = \frac{1}{m} \sum_{j=1}^m D_j, \quad (3.25)$$

as well as the simulated variance of the calculated detection rates $\hat{V}_{SIM}(D)$

$$\hat{V}_{SIM}(D) = \frac{1}{m} \sum_{j=1}^m (D_j - \bar{D})^2. \quad (3.26)$$

Spatially balanced designs are assumed to perform better than the non-spatially balanced designs when there is spatial autocorrelation in the target population. In the case of the Rockfish population, there is an obvious spatial clustering of the population and thus spatial autocorrelation is present. To evaluate for the effect of a spatial autocorrelation (clustering) on the achieved detection rates, the simulation study was repeated for a semi-clustered population and a random

population. These populations are altered versions of the original Rockfish population. To create a semi-clustered population, 40 percent of the units were randomly selected and their positions were randomly switched. To create the random population, all units were randomly switched. The resulting semi-clustered and random populations are illustrated in *Figure 3.10b* and *Figure 3.10c* respectively.

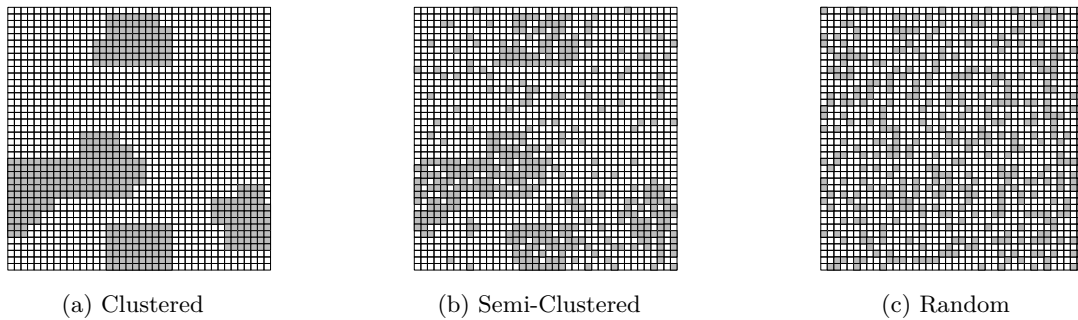


Figure 3.10: Illustration of the (modified) Rockfish populations. In *Figure (a)* the population is clustered, similar to the original Rockfish population adopted from Su and II (2003). *Figure (b)* and *figure (c)* are modified species distributions and were made less clustered.

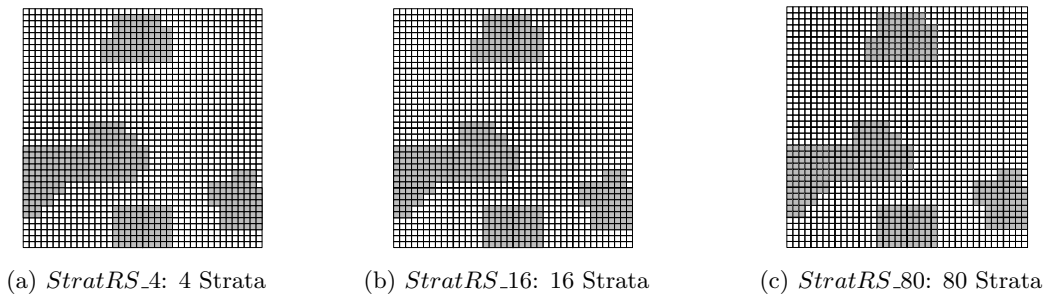


Figure 3.11: Illustration of the strata partitioning of the Rockfish population.

For each evaluated sampling designs the relative simulated variance RV_{SIM} of the detection rate compared with SRS was calculated, were

$$RV_{sim}(D) = \frac{\hat{V}_{SIM}^{Design}(D)}{\hat{V}_{SIM}^{SRS}(D)}.$$

The results are shown in *Table 3.2*. For all three Rockfish populations and all three sample sizes, the sampling designs achieved similar mean detection rates. The mean detection rate was close to 0.253. This is what one can expect since this is the total number of occupied units divided by the population size. This illustrates that if no prior information is available and an equiprobable sample is selected, then any sampling design is as good as any other in terms of achieving a high detection rate. However, when looking at the RV_{SIM} 's, some sampling designs did achieve less variable detection rates. For example, for the clustered Rockfish population, all designs had a RV_{SIM} of less than one. This means that they achieved less variable detection rates compared with SRS. The spatially balanced designs generally had the lowest RV_{SIM} . Of all the evaluated sampling designs, BAS achieved, for two out of three of the sample sizes, the lowest RV_{SIM} . The RV_{SIM} of BAS when sampling five, ten and 15 percent of the units was, respectively, 0.312, 0.201 and 0.250.

All three stratified sampling designs had RV_{SIM} 's of less than one. This means that they had less variable detection rates compared with SRS. This is because stratified sampling using proportional allocation, generally, selects samples with a higher spatial balance than SRS does. The importance of selecting a spatially balance sample is again illustrated by comparing the results of the stratified random samples. StratRS_4 distributes the samples over only 4 strata, StratRS_16 allocates the samples over 16 strata and StratRS_80 over 80 strata. StratRS_80 generally will ensure the highest spatial balance, with depending on the sample size one, two or three units per stratum. Because of this StratRS_8 has much lower RV_{SIM} compared with StratRS_4 and StratRS_4. Even though the strata for StratRS_16 are still large and not ideal as 'stratification to improve the spatial balance of the sample' it still performed clearly better than StratRS_4. The RV_{SIM} of StratRS_80 is close to the other spatially balanced sampling designs. StratRS_80 had even very similar results as GRTS for the five percent sample size. However, in general the other spatially balanced sampling designs performed slightly better. This can potentially be explained by the fact that with StratRS two (or more) selected units in neighbouring strata can still be clustered in the selected sample which is generally less likely with the other evaluated designs such as LPM2 and BAS.

For the semi-clustered population, similar results were obtained as for the clustered population. The main difference was that the RV_{SIM} 's were generally not as low as in the case of the clustered population. Although the spatially balanced designs still performed best, the difference with SRS and the stratified designs had become smaller. This indicates that the spatially balanced sampling designs perform better when there is a strong spatial autocorrelation in the species population. BAS performed generally best when five or ten percent of the units were sampled. LPM1 performed best when selecting 15 percent of the units.

The results for the random Rockfish population were different compared with the clustered and semi-clustered populations. For each sampling design, the RV_{SIM} 's was around one. This shows that if there is no spatial autocorrelation in the population, the sampling designs perform more or less similar. This means that if there is no spatial autocorrelation in the species population, then it is difficult to obtain higher detection rates or less variable detection rates for a sampling design, compared with SRS. This illustrates once more that spatially balanced sampling designs work best when nearby units have similar response values. If there is no such spatial autocorrelation in the response variable, like in this third Rockfish population, then spatially balanced sampling designs do not ensure better results compared with non-spatially balanced sampling designs.

Surveys using prior information: Unequiplorable Sampling

In this section the detection rates from the different sampling designs with unequal probability sampling are reviewed. A similar simulation study as for the equiprobable case is performed. The problem, however, is how to set the inclusion probabilities. Ideally, if a unit is not occupied, π_i should be zero. In that case the unit is no longer included in the sample population and thus cannot be selected in the sample. Since inclusion probabilities cannot be zero this problem was circumvented by setting the inclusion probabilities based on a small subset of units. This was done as follows: A semi-clustered Rock fish population, similar to the previous simulation study, was created. The study area was partitioned into 64 strata of 5 by 5 units, $N_h = 25$ for all $h = 1, 2, \dots, 64$. For each stratum, the number of occupied units c_h were counted. The inclusion

Table 3.2: Results of the simulation study on detection rates initial survey: Equiprobable sampling. The best performing method per population per percentage sampled is highlighted in bold.

Units Sampled		5%		10%		15%	
		D	$RV_{\text{SIM}}(D)$	D	$RV_{\text{SIM}}(D)$	D	$RV_{\text{SIM}}(D)$
Clustered	SRS	0.252	\	0.252	\	0.252	\
	StratRS_4	0.253	0.896	0.254	0.837	0.252	0.883
	StratRS_16	0.253	0.686	0.252	0.593	0.253	0.657
	StratRS_80	0.252	0.428	0.254	0.314	0.254	0.310
	LPM1	0.253	0.313	0.253	0.236	0.253	0.212
	LPM2	0.255	0.344	0.253	0.232	0.252	0.213
	GRTS	0.253	0.421	0.254	0.279	0.254	0.301
	BAS	0.251	0.312	0.252	0.228	0.252	0.250
Semi Clustered	SRS	0.253	\	0.252	\	0.254	\
	StratRS_4	0.253	0.950	0.254	0.897	0.252	0.952
	StratRS_16	0.254	0.894	0.254	0.847	0.253	0.935
	StratRS_80	0.254	0.739	0.254	0.757	0.252	0.779
	LPM1	0.254	0.741	0.252	0.674	0.253	0.707
	LPM2	0.253	0.696	0.253	0.653	0.253	0.665
	GRTS	0.254	0.800	0.255	0.750	0.254	0.777
	BAS	0.258	0.618	0.258	0.606	0.255	0.705
Random	SRS	0.253	\	0.253	\	0.254	\
	StratRS_4	0.252	0.989	0.253	1.007	0.254	0.996
	StratRS_16	0.254	1.001	0.256	0.997	0.254	1.018
	StratRS_80	0.254	0.991	0.255	1.022	0.254	0.994
	LPM1	0.252	0.997	0.253	1.155	0.253	1.081
	LPM2	0.254	1.051	0.253	1.143	0.254	1.042
	GRTS	0.254	0.989	0.254	1.176	0.254	1.042
	BAS	0.251	1.093	0.252	0.980	0.253	1.051

probability $\pi_{i,h}$ for unit i in the h^{th} stratum was set equal to

$$\pi_{i,h} = \frac{1}{25} \frac{c_h n}{\sum_{k=1}^{64} c_k}. \quad (3.27)$$

Thus each unit within a stratum is assigned the same inclusion probability, but each strata had different inclusion probabilities. For example, *Figure 3.12* shows a semi-clustered species distribution and the resulting inclusion probabilities. Strata with a large number of occupied

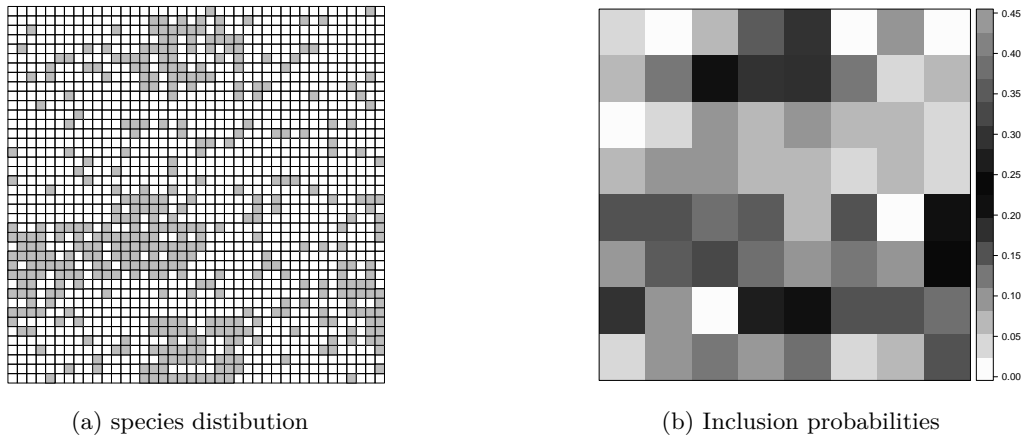


Figure 3.12: Illustration of the pre-set inclusion probabilities based on a simulated species distribution. Different shades of grey mean different inclusion probabilities

units have high inclusion probabilities and vice versa. In the case, when a stratum has no occupied units the inclusion probabilities for all units in that strata will be zero. Therefore, it was ensured that in each stratum at least one occupied unit was present. To achieve this, new species distributions were generated until this requirement was satisfied. An unequal probability sample is selected from the study area using the pre-set inclusion probabilities. The sampling designs that were evaluated are: SRS,CP, LPM1, LPM2, GRTS and BAS. Stratified sampling was not included since in the case of proportional allocation, the sample would be an equiprobable sample. Hence, the results for stratified sampling would be similar to the previous simulation study. Alike the previous simulation study, 1000 samples were selected for each sampling design. The simulation study was repeated for three different sample sizes, these are 5%, 10% and 15% of the units sampled.

For each evaluated sampling design \bar{D} and $RV_{SIM}(D)$ were calculated. The results are shown in *Table 3.3*. All evaluated sampling designs achieve higher detection rates than those achieved by SRS. This means that the selection of an unequal probability sample can improve the detection rate compared with SRS. The spatially balanced designs, LPM1, LPM2, GRTS and BAS achieved slightly higher mean detection rates compared with CP, which is not a spatially balanced sampling design. The resulting $RV_{sim}(D)$ s show that the precision in the detection rates was

Table 3.3: Results of the simulation study on detection rates initial survey: Unequiprobable sampling

Units Sampled	5%		10%		15%	
	\bar{D}	$RV_{\text{SIM}}(D)$	\bar{D}	$RV_{\text{SIM}}(D)$	\bar{D}	$RV_{\text{SIM}}(D)$
SRS	0.253	\	0.253	\	0.251	\
CP	0.329	1.211	0.331	1.020	0.339	1.032
LPM1	0.341	1.141	0.355	0.944	0.361	0.952
LPM2	0.340	1.061	0.356	0.884	0.370	0.996
GRTS	0.332	1.130	0.348	0.935	0.380	1.032
BAS	0.340	1.032	0.350	0.974	0.371	0.936

similar for all designs. However, CP had the lowest precision for each sample size.

3.3.2 Population Estimation

In this section the efficiency of the different sampling designs for population estimation is reviewed. This is done by means of Monte Carlo simulations using two virtual study populations. The results of these simulation studies can also be found in Robertson et al. (2013).

Example 1

The first virtual population is an example taken from Grafström et al. (2012b). The target value y_i is defined by the function

$$f(x, y) = 3(x + y) + \sin(6(x + y)) \quad (3.28)$$

with (x, y) in the unit box. This function $f(x, y)$, which has a spatial autocorrelation, is visualized in *Figure 3.13*. The study area is divided into 20 by 20 equally sized squared units. For unit i , the response variable y_i is the volume under *Equation (3.28)* within that unit. The population total is $Y = \sum_i^{400} y_i \approx 2.9994$. In this example 1000 equiprobable sample were selected for each sampling design. The evaluated sampling designs were: SRS, LPM1 LPM2, GRTS, BAS and StratRS. For StratRS the study area was partitioned in 20 strata of four by five units. Since CP and SRS are identical in the case of equal probability sampling CP was not included in this

simulation study. The simulation study was repeated for three different sample sizes: $n = 20, 40$ and 60 . The precision of a sampling design was estimated by using

$$\hat{V}_{\text{SIM}}(\hat{Y}) = \frac{1}{m} \sum_{i=1}^m (\hat{Y}_j - Y)^2,$$

where $m = 1000$ is the number of simulated samples and \hat{Y}_j is the population estimate from the j^{th} sample. The local mean variance estimator $\text{Mean}(\hat{V}_{\text{NBH}})$ was also calculated to evaluate the performance of this estimator for the spatially balanced sampling designs. For SRS and StratRS the exact variance estimates were computed.

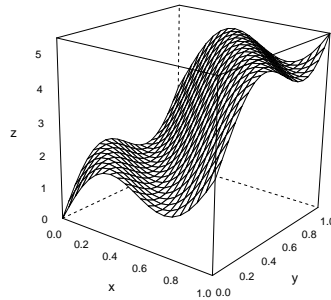


Figure 3.13: Continuous function $z = f(x, y)$ used to compute the response variable y_i (Example 1).

The results for this simulation study are given in the first four columns of *Table 3.4*. For each sampling design $\text{Mean}(\hat{Y})$ was approximately 2.9994 with no obvious bias. The simulated variances were similar for BAS, LPM1 and LPM2 and achieved overall the best performance. The simulated variances for GRTS were always higher than those achieved by BAS, LPM1 and LPM2. SRS had the highest simulated variance and performed the least efficient. StratRS had simulated variances similar to the other spatially balanced sampling designs, for $n = 20$ it has the highest precision and for $n = 40$ and $n = 60$ the variance is lower than the one observed for GRTS. What makes StratRS for $n = 40$ and $n = 60$ advantageous over the other methods is that, since at least two units per stratum are sampled, unbiased estimates for the variance are

possible. This is an important advantage for StratRS over the other spatially balanced sampling designs. This is especially since the mean local variance estimator was generally greater than the simulated variance. This indicates that the mean local variance estimator does not perform well for this problem. However, since this estimated variance is still smaller than the one observed for SRS, and since there is to our knowledge no alternative estimator that achieves a more precise estimate for the variance, this estimator can and should still be used. Of course, this is given the fact that one can expect a conservative estimate of the variance. Whereas the mean local variance estimator was generally a very conservative variance estimator for the spatially balanced sampling designs the exact variance estimates for SRS and StratRS were very close to the observed simulated variance especially for StratRS. This is advantageous for StratRS over the other spatially balanced sampling designs

Example 2

For the second example, an altered version of a artificial population introduced in Grafström et al. (2012b) was used. The population was defined by dividing a rectangular study area into 20 by 20 regular units. The units are numbered as follows: 1 to 20 for the first row, 21 to 40 for the second row, etc. The inclusion probability of unit i was defined by $\pi_i = nz_i / (\sum_j z_j)$, where z_i is the i^{th} value in the auxiliary matrix Z . This auxiliary matrix is visualized in *Figure 3.14*. Note that Z is an auxiliary matrix which has a spatial autocorrelation. The target value y_i had been given a linear relationship in the ratios y_i/z_i by setting

$$y_i = z_i \left(1.1 - 0.2 \frac{z_i - \min_j(z_j)}{\max_j(z_j) - \min_j(z_j)} \right) + \epsilon_i,$$

where ϵ_i is a random error term uniformly distributed on $[0, 0.1]$. This error term is added to distort the ratio between π_i and y_i and make it variable between the units. If this ratio is equal for all units, then the variance will be zero. This error term leads to a situation where some units will create negative bias and other units will create positive bias when added to a sample (Grafström et al., 2012a). The population total is $Y = \sum_i y_i \approx 525.37$. The evaluated sampling designs were: SRS, CP, LPM1, LPM2, GRTS, BAS. StratRS was not included since it was not

possible to select an exact (integer) sample size within each stratum. The simulation study was repeated for three different sample sizes: $n = 20, 40$ and 60 . The following test statistics were calculated: $\text{Mean}(\hat{Y})$, \hat{V}_{SIM} and $\text{Mean}(\hat{V}_{\text{NBH}})$. For SRS the exact variance estimates were computed.

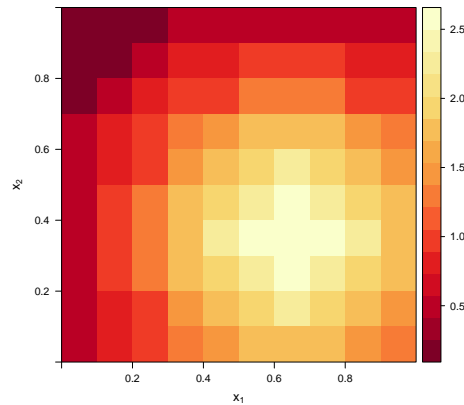


Figure 3.14: Visualization auxiliary matrix Z .

The results of this second example are given in the final three columns of Table 3.4. All evaluated methods achieved unbiased population estimates. The spatially balanced designs GRTS, BAS, LPM1 and LPM2 performed better for the variance estimation compared with the two non-spatially balanced methods CP and SRS. BAS had similar precision and accuracy to the other spatially balanced designs. However, when running the simulations BAS was observed to execute faster (see later this chapter). In this second example, the mean local variance estimator worked well for all spatially balanced designs. This illustrates that the mean local variance estimator can be a good estimator when there is a spatial autocorrelation in the response y_i .

Importance of a Spatial Autocorrelation in the Response for Population Estimation

In both previous examples, there was a form of spatial autocorrelation in the response value. As shown in the first example, spatially balanced sampling designs give more precise population estimates when there is a spatial autocorrelation in the target value in the case of equal probability sampling. To evaluate for the importance of having such a spatial autocorrelation in the response value, the simulation study for the first example was repeated. This time distortion was added

Table 3.4: Results simulation study for population estimation.

Design	Example 1			Example 2		
	Mean(\hat{Y})	\hat{V}_{SIM}	Mean(\hat{V}_{NBH})	Mean(\hat{Y})	\hat{V}_{SIM}	Mean(\hat{V}_{NBH})
<i>n</i> = 20						
BAS	2.9972	0.0122	0.0209	525.61	34.91	40.52
LPM1	2.9961	0.0122	0.0266	525.60	26.15	40.30
LPM2	3.0004	0.0124	0.0264	525.26	27.63	39.26
GRTS	2.9844	0.0961	0.0220	525.74	33.70	38.35
CP	-	-	-	525.50	87.45	34.19
SRS	2.9951	0.0943	0.0991 ($V_{SRS}(\hat{Y})$)	527.97	2608	2714.4 ($V_{SRS}(\hat{Y})$)
StratRS	2.9925	0.0101	0.0102 ($V_{StratRS}(\hat{Y})$)	-	-	-
<i>n</i> = 40						
BAS	2.9995	0.0033	0.0070	525.89	11.57	12.32
LPM1	2.9977	0.0037	0.0073	525.24	10.66	13.63
LPM2	3.0024	0.0037	0.0070	525.41	9.83	13.74
GRTS	3.0024	0.0051	0.0070	525.27	13.56	13.02
CP	-	-	-	525.66	38.59	12.02
SRS	3.0001	0.0487	0.0473 ($V_{SRS}(\hat{Y})$)	526.43	1127	1109.5 ($V_{SRS}(\hat{Y})$)
StratRS	2.9978	0.0048	0.0049 ($V_{StratRS}(\hat{Y})$)	-	-	-
<i>n</i> = 60						
BAS	3.0008	0.0016	0.0031	525.85	7.484	8.284
LPM1	3.0000	0.0018	0.0034	525.36	6.332	7.350
LPM2	2.9978	0.0018	0.0033	525.40	6.460	7.405
GRTS	2.9977	0.0032	0.0032	525.29	6.823	7.080
CP	-	-	-	525.55	25.79	6.677
SRS	2.9937	0.0288	0.0306 ($V_{SRS}(\hat{Y})$)	524.87	791.0	801.2 ($V_{SRS}(\hat{Y})$)
StratRS	2.9999	0.0030	0.0030 ($V_{StratRS}(\hat{Y})$)	-	-	-

to the target value of each unit to break the ‘smooth’ spatial autocorrelation in the target value. The new response values y'_i after distortion of the spatial autocorrelation were created by setting the distorted response variable y'_i equal to $y'_i = y_i + \delta_i$, with $\delta_i \in [-d(\text{range}(y_i)), d(\text{range}(y_i))]$ and d being the intensity of the distortion. Adding the distortion was done in such a way that $\sum_{i=1}^N y'_i$ remained equal to $\sum_{i=1}^N y_i$. This was done by selecting 200 uniformly random values of δ_i , thus $\delta_1 = (\delta_1, \delta_2, \dots, \delta_{200})$ and then setting $\delta_2 = (-\delta_1, -\delta_2, \dots, -\delta_{200})$. Finally for each unit i select a value δ_i from $\delta = (\delta_1, \delta_2)$. We repeated the simulation study for Example 1 in the previous section for four different levels of d : 0 (= identical as in the previous section), 0.1, 0.2 and 0.3. For each level of d , 1000 samples were selected, each of size 40. *Figure 3.15* gives an illustration of the response density function used in Example 1 after adding distortion.

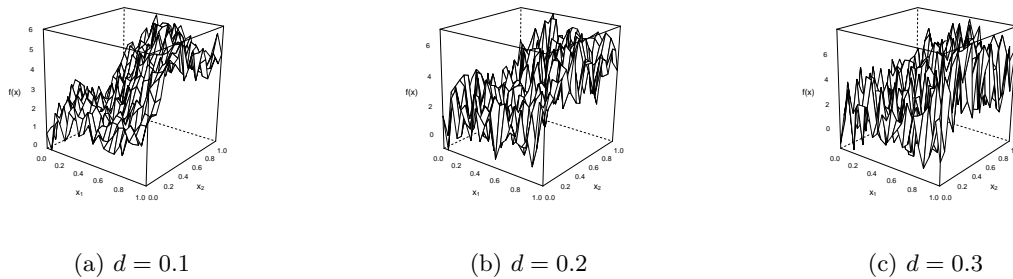


Figure 3.15: Example of added distortion to density function target population Example 1.

The results are shown in the first two columns of *Table 3.5*. The variance of all the spatially balanced designs increased with increasing distortion d . For SRS, however, the variance remained more or less constant. One would expect that the variance would be exactly the same for SRS. However, due to the relative small numbers of simulations there can still be some differences in the actual observed simulated variance. This confirms the general expectations that for equiprobable sampling, spatially balanced designs work better than SRS only if there is spatial autocorrelation in the response variable. These results also confirm the general perception that, since the results of spatially balanced sampling designs are always better or as good as the results of SRS for any distortion rate, it is advisable to use spatially balanced designs regardless of the spatial autocorrelation of the response variable. Moreover adding distortion did not seem

to affect the accuracy. Overall the population estimates were found to be unbiased.

Importance of constant proportionality between y_i and p_i for Population Estimation

In the second example, y_i was perfectly proportional with p_i and thus have a perfect linear relationship, except for some added error ϵ , for all values of y_i . If this proportionality is perfectly constant, then the variance of the HT-estimator will be zero. Unfortunately, unless the response value y_i is known for each unit, it is often impossible to set the inclusion probabilities so that there is a perfect linear relationship between the inclusion probabilities and the target variable. This issue is well documented in the literature (Cochran, 1977; Tillé, 2006). When using the HT-estimator, to obtain precise estimates the ratio in $\frac{y_i}{\pi_i}$ should be (near) constant for each unit. Hence, there should be a linear relation between the number of invasive species y_i and the set inclusion probability π_i . This means that units with a low species density (y_i) should ideally be assigned a low inclusion probability and units with a high species density should be assigned a high inclusion probability. Deviations from this linear relationship can lead to imprecise results. The linear relationship between y_i and π_i and possible effects caused by deviation from this linear relationship are illustrated in *Figure 3.16*. The black line illustrates the perfect linear relationship between $\frac{y_i}{\pi_i}$. In the case when every unit in the population is a point somewhere on this line, then any selected sample will result in accurate population estimates. If not, then points that fall above the line will add a positive bias to the population estimate and points that fall beneath the line will create negative bias. If the units are scattered symmetrically around this line, then the population estimate can be unbiased but is likely to have a larger variance. In most cases the species density, and thus y_i , is not known. Auxiliary information, such as the species habitat suitability, can be used to approximate the species density. This approximation will never be 100 percent perfect. Therefore, it is difficult to set the inclusion probability perfectly linear with the species density based on auxiliary information. Hence, population estimates based on inclusion probabilities set by using auxiliary information can be inaccurate and imprecise. For example, if a unit has a low \hat{S}_i then that unit is believed to be unoccupied and will be assigned a small inclusion probability. Assume that unit is selected in

the sample and, totally unexpected, the species would be observed in that unit after all. Then that would create a large bias and the population total would be overestimated. Similarly, if a unit is assigned a high inclusion probability and no species has been observed, this will create an underestimating estimate.

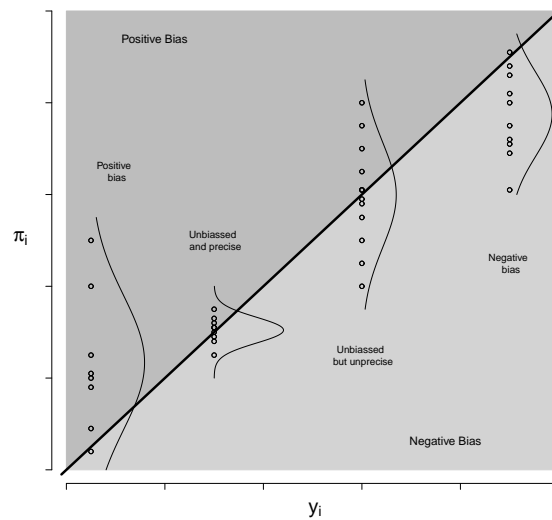


Figure 3.16: Illustration of the importance of a linear relation between $\frac{y_i}{\pi_i}$ when using probability sampling and the HT-estimator. The black line shows the perfect relation between $\frac{y_i}{\pi_i}$. Each dot represents a sampling unit.

To test the effect of deviations from this ideal proportionality, a modified version of the simulation study that was conducted for Example 1 was performed. The ideal inclusion probabilities have perfect linear relation with the response, thus $\pi_i = y_i \frac{n}{\sum_{i=1}^N y_i}$. For this simulation study four different sets of inclusion probabilities were compared with increasing deviations from the ideal linear relation in $\frac{y_i}{\pi_i}$. These four sets were obtained by adding distortion to the ideal inclusion probabilities π_i . The new inclusion probabilities with added distortion π'_i were calculated by setting $\pi'_i = \pi_i + \delta_i$, where $\delta_i \in [-d(\text{range}(\pi_i)), d(\text{range}(\pi_i))]$ and d being the intensity of the distortion. Adding the distortion was done in such a way that $\sum_{i=1}^N \pi'_i$ remained equal to $\sum_{i=1}^N \pi_i = n$. To achieve this requirement we applied the same method as was done in the previous section. Four different levels of d were evaluated: 0 (perfect linear autocorrelation), 0.1,

0.2 and 0.3. For each level of d , 1000 samples were selected, each of size 40. *Figure 3.15* gives an illustration of the response density function used in Example 1 after adding distortion.

The results are shown in the last two columns of *Table 3.5*. The accuracy and precision of SRS did not get affected by distorting the linear relation in $\frac{y_i}{\pi_i}$. However, all the sampling designs that can select an unequal probability samples lost precision when the linear relation in $\frac{y_i}{\pi_i}$ became more distorted. The decrease in the precision seemed similar for the spatially balanced designs. The simulation variance for CP was similar to the spatially balanced designs. The accuracy for CP was a little more biased compared with the spatially balanced sampling designs. These results indicate that having a linear relation in $\frac{y_i}{\pi_i}$ is important when selecting an unequal probability sample. These results suggest that the precision of an unequal probability sample can decrease rapidly with increasing distortion and thus caution should be taken when the inclusion probabilities are not set (nearly) proportional with the response variable.

Decision

In this section, simulation studies were performed to evaluate several sampling designs for their potential use for adaptive monitoring. The main criteria for being a good sampling design was achieving high detection rates. Additionally, it was considered which sampling design is best for population estimation.

All the sampling designs achieved very similar mean detection rates. However, the spatially balanced sampling designs generally achieved the lowest simulated variances in the detection rates. This was especially the case when the underlying species population was clustered or semi-clustered, thus when spatial autocorrelation was present. The stratified sampling designs also achieved lower variances compared with SRS. Especially in the case when the strata are small enough, such that there are only one, two or three units selected per stratum, the difference in the observed simulated variance is nearly as good as for the other spatially balanced sampling designs. Out of all the evaluated spatially balanced sampling designs, BAS generally performed at least nearly as good as or better than the other spatially balanced sampling designs in terms

Table 3.5: Results for the “test for importance spatial autocorrelation in response” and “test for importance proportionality $\frac{y_i}{\pi_i}$ ”.

Design	d	Spatial autocorrelation y_i		Linear proportionality $\frac{y_i}{\pi_i}$	
		Mean(\hat{Y})	\hat{V}_{SIM}	Mean(\hat{Y})	\hat{V}_{SIM}
BAS	0	2.998	0.004	2.999	0.000
	0.1	2.996	0.006	3.000	0.002
	0.2	2.999	0.012	2.998	0.013
	0.3	2.989	0.019	2.993	0.020
LPM1	0	3.001	0.005	2.999	0.000
	0.1	3.001	0.006	2.997	0.002
	0.2	3.004	0.011	2.995	0.012
	0.3	2.993	0.014	2.996	0.019
LPM2	0	3.003	0.004	2.999	0.000
	0.1	2.999	0.005	3.000	0.002
	0.2	3.005	0.011	3.000	0.012
	0.3	3.003	0.014	3.003	0.021
GRTS	0	3.001	0.005	2.999	0.000
	0.1	3.001	0.006	2.998	0.002
	0.2	2.997	0.012	3.002	0.011
	0.3	2.995	0.016	2.996	0.018
SRS	0	2.993	0.050	2.941	0.049
	0.1	3.001	0.048	3.000	0.048
	0.2	2.989	0.049	2.988	0.050
	0.3	2.997	0.049	2.993	0.049
CP	0	/	/	2.999	0.000
	0.1	/	/	2.981	0.004
	0.2	/	/	3.002	0.014
	0.3	/	/	3.010	0.021

of observed variance in achieved detection rates. It is important to note that these results are conditional on the fact that the tests for spatial balance were only performed for BAS, SRS and GRTS and the tests for the detection rates did not include all possible spatially sampling designs such as systematic sampling.

The results of the simulation studies on population estimation, in Table 3.4, showed that spatially balanced sampling designs generally performed better compared with SRS and at least as good as CP. When spatial autocorrelation is present (example 2) BAS achieves accurate results often with a higher precision compared with the other spatially balanced sampling designs. One important note is that StratRS had a higher precision than BAS for a sample size of 20 and had equal or even better unbiased estimates for all three test sample sizes. Furthermore StratRS with $n = 40$ or $n = 60$, which is 2 or 3 replicates per stratum, is one of the simplest spatially distributed methods which has an unbiased estimate of the variance. This is not the case for BAS or GRTS. Furthermore StratRS is an intuitive design and is simple to apply. The main reason why StratRS as such is not considered in the next chapters is because our main focus was on using methods that would allow for probability sampling when the inclusion probabilities are set on the unit level. In hindsight, it would have been a good idea to include StratRS with two (or more) units per strata to be able to compare the results with a wider variety of relevant sampling designs.

The last two simulation studies on population estimation show that it is important that there is a linear autocorrelation in the response y_i and that there is a linear relation in $\frac{y_i}{\pi_i}$ when selecting a (unequal) probability sample. If this is not the case, and the inclusion probabilities are not set proportional (enough, for example $d < 0.2$) with the response then the results of a probability sample can become imprecise and the advantage over selecting an equiprobable sample disappears. The fact that population estimates can become imprecise if the inclusion probabilities are not set precise enough (proportional with Y_i 's) means that one has to be careful with selecting unequal probability samples. This is especially the case when not enough information is available to base the inclusion probabilities on, like before the first survey. Therefore, it is important to notice that one has to be careful with selecting an unequal probability in the initial sample.

In general we can conclude that BAS performed (nearly) as good as, or better than the other designs. Note that in this chapter, the performance of different sampling designs were evaluated when selecting a sample for a single survey. In the next chapters, the performance of these sampling designs for repeated surveying or long term monitoring will be evaluated.

3.4 Practical aspects of BAS

BAS does not only achieve high detection rates and precise population estimates, it also has some practical features that make this sampling design desirable for long term adaptive monitoring. In this section we discuss some of the practical features of BAS for adaptive monitoring.

3.4.1 Long Term Sampling

The selection of a spatially balanced sample can improve the population estimates compared with a non-spatially balanced sample. In the case of repeated sampling, such as long term monitoring, for each year, the selected sample should preferably be spatially balanced. In addition to the selection of a spatially balanced sample in each year, it sometimes is important that the cumulative set of all units selected over all previously years are spatially balanced. For example, assume a spatially balanced sample is selected in year one and a second spatially balanced sample is selected in year two. Sometimes it is preferable that the selected units in year one combined with the selected units in year two are a spatially balanced set of units as well. By doing this the combined set of sampling units over the two years would be a spatially balanced set of units which would allow for better mapping of the species distribution after year two for example since the total sample will have a higher spatial balance. If this was not the case, sampling points for year two could be very close to the units already visited in year one, which would provide less information. Note that it is not always necessary that all sampling units are cumulatively spatially balanced. For example, if one is interested in capturing a temporal change, for example in the context of model-based sampling, it is often useful to sample a certain number of units each year (Gitzen et al., 2012). Having a certain number of units that are

sampled could also be useful to map the changes in the species distribution over time. Deciding on which type of sampling design to use will depend from case to case and will not be discussed in further detail.

Since BAS can add points to a sample without losing its spatial balance, as was illustrated previously in the chapter, this can easily be achieved with BAS. Assume that each year in a 3 year long survey, 50 units will be selected. To ensure that both the year specific samples and the overall sample are spatially balanced, simply generate a list of the first 150 selected units by a BAS sample. The first 50 units should be sampled in year one, the second cohort of 50 units should be sampled in year two and finally the last 50 units will be sampled in year three. It is also possible that, for example in the case funding would be available for another three years, to add units to the sample while keeping the spatial balance of the sample. This is, as long as the starting points of the random start Halton sequence are known. Note that a similar approach is possible for GRTS, however, as illustrated in *Section 3.2.6*, GRTS loses spatial balance compared with BAS when adding points to a sample. *Figure 3.17* shows a spatially balanced sample selected over a three year long survey using BAS in 2D and 3D. We can see that the points are spatially balanced within each year and also over all three years. Note that this method for long term sampling works for both equal and unequal probability sampling. For example, in the case of non-adaptive monitoring in which one would like to allocate a larger share of the sampling effort to certain parts of the study area, an unequal probability sample can be selected using BAS. By visiting the selected units in the order of the Halton sequence, the visited units will be spatially balanced within each survey. Furthermore, target inclusion probabilities are achieved each year as well as over all the years. In the case of adaptive monitoring in which the (unequal) inclusion probabilities will be updated at the end of each year, the sample within each year will be spatially balanced. However, the overall sample (all units combined) is likely to be no longer spatially balanced, this is especially the case when there are large changes in the set inclusion probabilities.

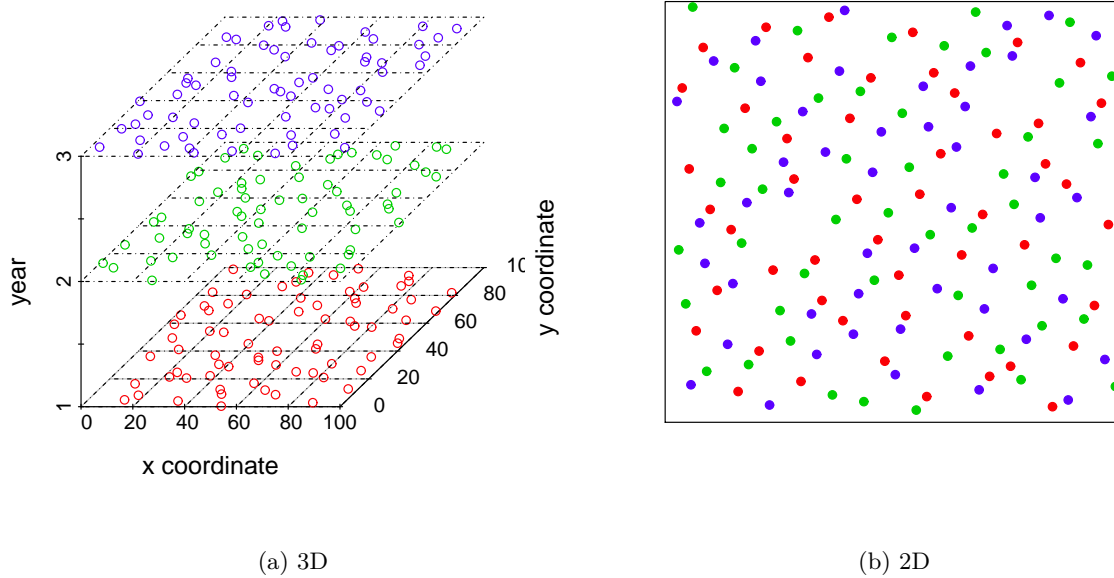


Figure 3.17: Illustration of a long term sample selected using BAS in 2D and 3D. During the three year long survey 50 points were selected in each year.

3.4.2 Premature Ending, Non-Response and Oversampling

In a previous section of this chapter, it was illustrated that adding points to a sample or skipping points does not affect the spatial balance of a BAS sample much. This provides BAS with additional desirable features which can be used for long term sampling.

The first one is dealing with premature endings of a planned sequence of sampling units. In the case when not all units are sampled in a certain year, one can simply continue sampling using the original list of selected Halton points. For example, in *Figure 3.17*, if only 30 points could be sampled in year one, one can start with point 31 in the next year and sample the next 50 points and so on. This allows for greater flexibility for researchers in case of unforeseen circumstances. Another advantage of BAS over GRTS is the fact that oversampling is not necessary for BAS. Oversampling is a sampling technique in which more sampling points are selected than the required sample size was. Oversampling can be useful when it is not clear what the final sample size would be. An example of the usefulness of this feature would be where funding for an additional survey needs to be approved annually. An example of oversampling is when additional sample units are points are generated when there is a risk of non-response. Oversampling is also often applied to ensure that the required sample size will be reached when non-responses are likely to occur. GRTS, for example, requires oversampling to maintain its spatial balance when dealing with non-response and additional units need to be selected. GRTS cannot add points to a sample (unless the permutation scheme and the details of the hierarchical partitioning are stored, which is difficult to program). With BAS, only storage of the seeds u_1, u_2, \dots, u_d used to perform the random start Halton sequence is required to select additional sampling points.

3.4.3 Computational Efficiency of BAS

In appendix B a R script is given for implementing BAS and StratBAS. These scripts require shape files (.shp) as input files to specify the study area. Generating an equiprobable BAS sample over a rectangular area is fast. The speed is comparable or even faster than GRTS. For example, an equiprobable sample of size 10000 selected within the unit square required a running time of 1.1 seconds for BAS and 2.4 seconds for GRTS, this includes reading in the shape files. Most of the times, however, the study area is not rectangular shaped. In that case a bounding

box, this is the smallest possible rectangle that includes each point of the sampling area, is placed around the sampling area. Halton points will be generated within the bounding box. Each Halton point has to be evaluated to see if it is within the study area or not. Only points that fall within the sampling area are added to the sample. The higher the fraction of the non-study area within the bounding box, the more Halton points will be rejected and thus the slower BAS will be. Another problem with BAS sampling is that if the study area is not of one polygon but exist out of several polygons scattered within the bounding box, it can become computational slow to evaluate if a point falls within the sampling area or not. An illustration of these two problems is given in *Figure 3.18*. The sampling are is all the broad leaved trees areas on the Chatham Islands, New Zealand. The bounding box is indicated by the blue rectangle. The total sampling area is the collection of several small polygons scattered within the bounding box. The majority of the bounding box is 'empty'. Evaluating if a point falls within a polygon is computationally slow.

The computational speed of BAS sampling can be improved as follows: Place a bounding box around each separate polygon. Next, generate a set of Halton points within the overall bounding box (blue rectangle). Next, evaluate for each Halton point if that point falls within one of the polygon specific bounding boxes (red rectangles). If not, reject the point from the sample. Otherwise evaluate if that point falls within the polygon within that particular bounding box. To select a sample of size 100 over the sampling area as shown in *Figure 3.18*, GRTS takes almost 24 seconds, BAS without polygon specific bounding boxes takes around 41 seconds and BAS with the polygon specific bounding boxes takes 13 seconds. This method is faster since evaluation of a point within a rectangular box is generally faster compared with testing if a point lies within an irregular polygon.

For study areas rasterized into squared units, the computational efficiency of BAS can be improved by mapping the two dimensional cells onto the one dimensional real line. This method is similar to GRTS's hierarchical partitioning and mapping on the real line. In short each unit will receive a unique address based on the coordinates of the unit. Since the dimensions of each unit are known, it is possible to link any two dimensional Halton point to the right address.

Therefore, a set of two dimensional Halton points can be transform to a set of addresses. The idea is that before the selection of a BAS sample, the first two dimensions of a generated list of Halton points are combined into a one dimensional list of addresses. Based on this list of addresses it is faster to link a Halton point to a certain unit since the number of dimensions are reduced. This method could be applied to BAS sampling in higher dimensions. When performing this method selection of an equiprobable sample of size 500 from a study area of 200 by 200 units takes less the one second.

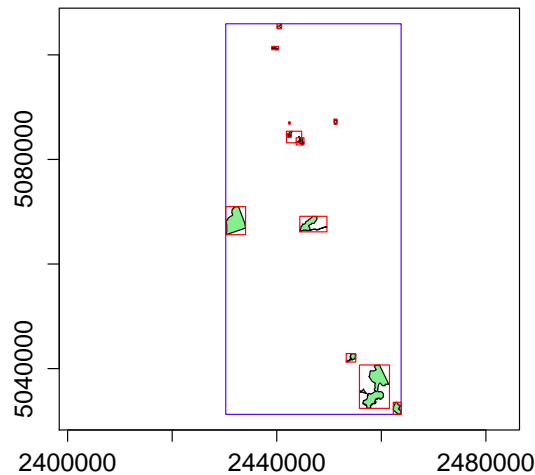


Figure 3.18: Illustration of the computational efficiency problems of BAS when sampling from a scattered landscape. The black polygons (.shp files) illustrate broad leaved trees area on the Chatham Islands, New Zealand. The blue rectangle illustrates the overall bounding box, this is the smallest rectangular area in which Halton points can be generated. The red rectangles around each separate polygon are the polygon specific bounding boxes used to improve the computational efficiency for BAS.

3.5 Discussion

In this chapter a selection of several sampling designs and especially spatially balanced sampling designs have been introduced. Simulation studies were performed to compare these sampling designs.

The first criteria, in this thesis for a suitable sampling design is to achieve high detection rates. In general spatially balanced sampling design achieved similar mean detection rates compared with the other sampling designs such as SRS, CP and StratRS. However, when looking at the variance of the achieved detection rates the spatially balanced sampling designs performed better. For many simulation studies BAS achieved observed variances in detection rate at least nearly as low or lower than the other spatially balanced sampling designs. However, in the case where there is no spatial autocorrelation present in the species distribution there is little advantage in using a spatially balanced sampling design.

The second criteria was to achieve precise population estimates. In general, the spatially balanced sampling designs had the best results in the case of unequal probability sampling. In the equiprobable case there was little difference between StratRS, BAS and GRTS. In fact, in the equiprobable case StratRS has the advantage that, as long as at least two units are selected per sample, an unbiased variance estimate is available.

Overall, BAS scored good for both criteria, except for the fact that no unbiased variance estimator is available. BAS is a new spatially balanced sampling design based on a relatively simple algorithm. Moreover, several practical features of BAS were explored such as BAS for long term monitoring. BAS also executes faster than other spatially balanced sampling designs. All this led to the decision to select BAS as the most consistent performing sampling design, at least for the initial survey. We will use BAS as our preferred spatially balanced sampling design that allows for unit specific inclusion probabilities in the coming chapters. It is important to note that, as illustrated in table 3.5, that when it is difficult to set unit specific inclusion probabilities, which is the case when insufficient prior information is available, population estimates can become imprecise. Hence, using methods that make use of unit specific inclusion probabilities are not always feasible to use for a single survey. In the next chapters, we illustrate how one can

set and adjust these inclusion probabilities for consecutive surveys over time.

Chapter 4

Adaptive Monitoring using Ecological Information to update the inclusion probabilities

4.1 Introduction

To increase the detection rate over consecutive surveys, areas with high species densities should be sampled more intensively than low species density areas. To achieve this, a method for adaptive monitoring has been introduced in Chapter 2. A key concept of the proposed method is that the inclusion probabilities should be adjusted over time such that the areas with expected high species densities are given a higher inclusion probabilities than those areas with an expected low species density. Accordingly, when selecting a probability sample, areas with a high abundance of the species of interest are expected to be sampled more frequently than lower density areas. Hence, with this approach for adaptive monitoring, we expect the detection rate to increase over time. This principle of setting the inclusion higher in observed and expected high density area is likely to be an oversimplification. Ideally this would be part of a much more complex process defined by the dynamic processes of the species population. In real life the density and distribution of the species depends on spatial and temporal deterministic changes, for example

seasonal changes in the population, as well as stochastic changes, for example random predator prey interactions. In later sections and the next chapters we will discuss how to adjust the monitoring strategy for changing, non-stationary populations. However, the focus in this chapter will be on showing how to adjust unit specific inclusion probabilities and how this can effect the detection rate.

If the underlying species distribution is known, then it is possible to set the inclusion probabilities proportional to the species distribution. We can say that the perfect inclusion probabilities, π_i' , have a perfect correlation with the actual species density. Unfortunately, the actual species distribution is almost never known. Hence, when adjusting the inclusion probabilities, ideally, the π_i 's are set as close as possible to the π_i' 's, such that the error ϵ_i in

$$\pi_i = \pi_i' + \epsilon_i \quad (4.1)$$

is minimized. If the ϵ_i 's are small, then it means that the chosen π_i 's will approximate the π_i' 's better. In that case, we expect that the detection rates will be higher than when the ϵ_i 's are larger. Note that as discussed in the previous chapter (section 3.3.2) this approach will only work well for population estimation if the ϵ_i 's are small. With the ϵ_i 's becoming larger using an unequal probability sampling method can become no better or even worse than when using the equal probability version of that method.

In Chapter 2, several sampling designs have been introduced that can select a probability sample based on the pre-set inclusion probabilities. In the current and subsequent chapters we discuss how we can actually set and adjust those inclusion probabilities. To minimise ϵ_i , we need to set the inclusion probabilities proportional to the species density distribution. Since the species distribution is usually unknown, we will make use of auxiliary information that is correlated with the species distribution. As mentioned in Chapter 2, there are two types of auxiliary information one can use to update the inclusion probabilities: ecological information and spatial information. In this chapter, we illustrate how to use the ecological information to update the inclusion probabilities for adaptive monitoring.

The idea is to use auxiliary information in the form of ecological data to predict the species distribution. All the available ecological data are modelled into a single auxiliary variable to predict the species distribution, the species' habitat suitability, say S . The S is an index of the likelihood that a species will settle in a specific type of habitat. Thus, an area with a high species habitat suitability will be preferred by the species to settle in over sites with lower S . We set the inclusion probabilities proportional to the estimated species' habitat suitability (\hat{S}). The species habitat suitability can be estimated from the available ecological data by using species distribution models (SDMs). These SDMs model the relationship between ecological covariates and the observed species distribution (Elith and Leathwick, 2009). Thus, based on the collected ecological information, such as elevation, precipitation, type of vegetation or road density, SDMs can estimate the species habitat suitability \hat{S}_i for each unit i . For example, during a bird survey, the type of vegetation is recorded for each visited unit, as well as the presence or absence of the bird species of interest in that unit. Two types of vegetation were observed: grassland and forest. Most birds were found in sampling units within the grassland vegetation and only a few birds were found in the forest units. In that case, a SDM will likely predict that grassland units have a higher S than the units with a forest habitat. Many different species distribution modelling techniques exist. Some examples of SDMs include: Generalized Linear Models such as logistic regression (Nelder and Wedderburn, 1972), Maximum Entropy (MaxEnt) (Phillips et al., 2004) and Support Vector Machines (Vapnik, 1999). The selection of the best species distribution modelling technique depends on the available information about the species, the research questions, the types of predictors and model expertise. A discussion of the differences between these species distribution models would go beyond the scope of this thesis and a detailed review can be found in Elith and Leathwick (2009).

The assumption that there exists a relationship between S and the species density is commonly made in ecology (Elith and Leathwick, 2009; Elith and Graham, 2009). Hence, the results of SDMs are often used to predict the species distribution over a study area (Comte and Grenouillet, 2013; Lyet et al., 2013). Some studies describe how SDMs can be used to improve the sampling design for future surveys (Edwards et al., 2005; Williams et al., 2009; Albert et al., 2010; Peterman et al., 2013). For example, Peterman et al. (2013) illustrated that for *Ambystoma*

jeffersonianum, an endangered salamander species in Illinois, U.S.A, a SDM can be used to predict if potential breeding ponds are occupied or not. They concluded that the available ecological information can be used to improve the detection rate when sampling for this species.

As aforementioned in chapter 1, several studies have used SDMs in an adaptive monitoring context. Stohlgren and Schnase (2006) suggested a methodology for adaptive monitoring in which the ecological information was used to update the monitoring design. Guisan et al. (2006) gave a practical illustration of the adaptive monitoring scheme suggested in Stohlgren and Schnase (2006). In Guisan et al. (2006), before each survey, a SDM is fitted using the collected data from all previous surveys to predict S for each unit in the study area. Next, Guisan et al. (2006) partitioned the study area into a suitable stratum and an unsuitable stratum. This partitioning was done based on the value of \hat{S}_i for each unit, by using a predefined threshold value. Units with an \hat{S}_i above the threshold value are grouped in the suitable stratum, and units with an \hat{S}_i below the threshold value are grouped in the unsuitable stratum. This threshold value was set such that the majority of the units would be partitioned into the unsuitable stratum and only those units with a high \hat{S}_i would be grouped into the suitable stratum. Next, a stratified random sample was selected with a fixed sampling allocation of 50 percent of the sampling units in each stratum. Since the suitable stratum is usually smaller than the unsuitable stratum, but the same number of units are selected in the suitable stratum as in the unsuitable stratum, the detection rate is expected to increase. According to Guisan et al. (2006), over time, more data will become available, and the estimation of S becomes more precise. Hence, the stratification becomes more accurate. In this case, more accurate means that most of the units that are occupied by the invasive species will be partitioned into the suitable stratum and vice versa. Other studies that have used SDM to update the monitoring strategy are Le Lay et al. (2010) or Crall et al. (2013). Similar to Guisan et al. (2006), these two studies partitioned the study area into a suitable stratum and an unsuitable stratum, and accordingly selected a sample using StratRS. We call this type of monitoring, adaptive monitoring using StratRS (AMStratRS). These three studies showed that the detection rate can be increased by using AMStratRS. With AMStratRS the strata partitioning is adjusted after each survey. Once the strata bound-

aries are set, units within the same stratum are assigned equal inclusion probabilities. These inclusion probabilities depend on the number of units to be selected in each strata. For example, say the suitable stratum has $N_h = 100$ and $n_h = 20$, then $\pi_i = \frac{20}{100}$ for each unit in that strata. Albert et al. (2010), stated that, since SDMs can compute \hat{S}_i for each unit in the study area, stratification based on \hat{S}_i is not necessary. Instead, Albert et al. (2010) discussed the use of other sampling designs that do not need any stratification. By doing this, there is no need for the often arbitrary decision for a threshold value for stratification or for a rule on the allocation of the sampling units. Albert et al. (2010) hypothesised that by using such non-stratified sampling designs, the detection rate might even increase more than when using StratRS. This idea, that other sampling designs could be used for adaptive monitoring instead of StratRS, as discussed in Albert et al. (2010), was one of the main improvements that we suggested when introducing a new framework for adaptive monitoring in Chapter 2.

In this Chapter, we begin by explaining how the ecological information and SDMs can be used to set and adjust the inclusion probabilities of each unit. Recall that the inclusion probabilities cannot be (near) zero. To avoid these problems with near zero inclusion probabilities three rescaling methods for the inclusion probabilities are introduced.

Next, sampling designs are reviewed for their performance for adaptive monitoring. First, it will be evaluated whether StratRS, as used in Guisan et al. (2006) to stratify the area into two homogeneous strata, a suitable stratum and an unsuitable stratum, can be replaced by other sampling designs which do not require any stratification, for example BAS or CP. Simulation studies in Chapter 3 showed that BAS overall performed best. Hence, of all the sampling designs considered here, especially the use of BAS will be evaluated. Second, in the Chapter 3, it was illustrated that spatially balanced sampling designs achieve higher or more consistent detection rates than those with SRS, StratRS and CP. Here, we evaluate how well the spatially balanced sampling designs, and especially BAS, perform in the context of long term adaptive monitoring.

4.2 Methodology

The methodology for adaptive monitoring using ecological information and SDMs is illustrated in *Figure 4.1*, and is: After the initial sample is selected, a SDM is used to estimate S_i for all units in the study area. For the estimation of S_i , the ecological information obtained during the initial survey is used. The inclusion probabilities for the next survey are set proportional to \hat{S}_i . Based on the set inclusion probabilities a sample is selected. After the second survey, the cumulative data from all previous surveys is used to update \hat{S}_i . The updated \hat{S}_i is then used to set the inclusion probabilities for the next survey, and so on.

Over time more and more information becomes available. Hence, the SDM's predictive power to estimate S_i will become more precise (Guisan et al., 2006). We expect that if the estimation of S_i is more precise, it will be more closely related to the actual species distribution. If so, the error ϵ_i in *Equation 4.1* is expected to decrease with an increasing number of surveys. This is of course if the population is stationary and is not changing over time. Later in this chapter and in the coming chapters these methods will be evaluated for more dynamic populations. Hence, the detection rate is expected to increase over time. In this section, each step of adaptive monitoring using ecological information is explained into greater detail.

4.2.1 Setting the inclusion probabilities using species distribution models

The monitoring strategy will be adapted over time by setting the inclusion probabilities proportional to \hat{S}_i . SDMs are used to estimate S . As aforementioned, there are many SDMs and the selection among the different SDM goes beyond of the scope of this thesis. In the following simulation studies, for the sake of simplicity, we use simple logistic regression as the SDM of choice. Logistic regression was chosen because many scientist are familiar with the basics of logistic regression, as opposed to many other SDMs.

First of all, we assume that \hat{S}_i is a strictly positive estimate. This is a fair assumption to make since many SDMs, for example logistic regression, will estimate S on a scale between zero and one, with one being the maximum suitability (Elith and Graham, 2009). We know that $\sum_{i \in U} \pi_i = n$ and π_i is set proportional with \hat{S}_i , i.e. $\pi_i = c\hat{S}_i$ where c is a constant. Hence, the

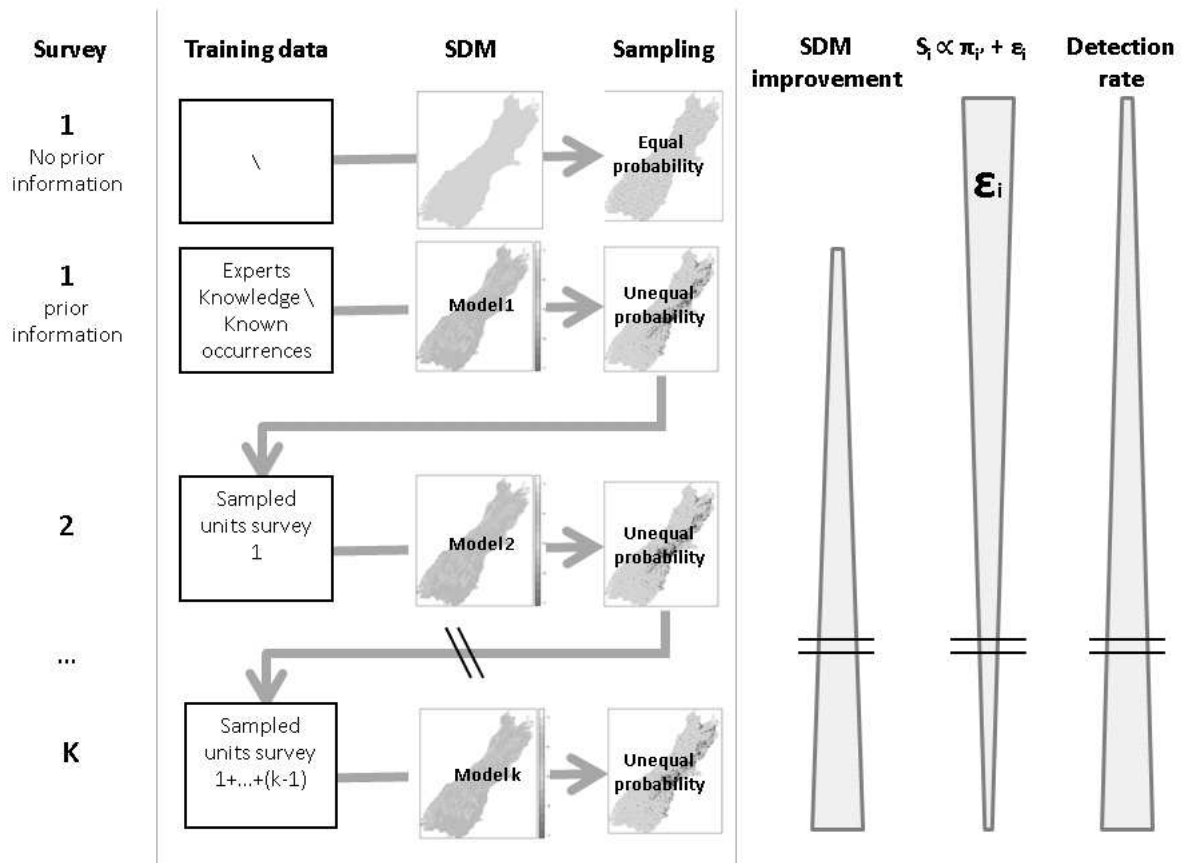


Figure 4.1: Illustration of the proposed methodology for adaptive monitoring. Figure is based on a figure in Guisan et al. (2006). Note that the initial sample is an equiprobable sample, unless prior information is available. In that case, a SDM can be fitted before the start of the initial survey and inclusion probabilities can be set proportional to \hat{S}_i . For species with a stable distribution over time the estimation of S is expected to become more precise, ϵ_i is expected to decrease and the detection rate is expected to increase.

inclusion probabilities for the k^{th} survey are given by

$$\pi_{i,k} = \hat{S}_{i,k-1} \frac{1}{\sum_{i=1}^N \hat{S}_{i,k-1}} n_k \quad \text{with } k \geq 2. \quad (4.2)$$

Note that k is always greater than one because the inclusion probabilities are only updated after the initial survey.

4.2.2 Rescaling Approaches

If \hat{S}_i is zero or near zero (by which we mean relatively small compared with the other \hat{S}_i 's), then given *Equation 4.2*, π_i will be (near) zero. This is problematic for two reasons. First, units with zero inclusion probabilities can never be sampled. As discussed in the introduction, it is important that when monitoring an invasive species that a part of the samples size is used for surveillance sampling. Hence, each unit should have some chance, even if small, to be included in the sample. This will be discussed in more detail later in this chapter. Second, inclusion probabilities of (near) zero make the HT-estimator non definable. This means that population estimation becomes impossible. To resolve these problems with (near) zero inclusion probabilities three rescaling techniques are proposed:

1. The minimum suitability approach: This method rescales \hat{S}_i such that for each unit the minimum estimated species habitat suitability is greater than zero. This is done by setting a lower limit on \hat{S}_i . Although the definition of this lower limit is arbitrary, we propose the following rule of thumb. The lower limit \hat{S}_{\min} , which will become the minimum suitability, is given by

$$\hat{S}_{\min} = l(\max \{\hat{S}_i\} - \min \{\hat{S}_i\}), \quad (4.3)$$

where l can take a value between zero and one. If a unit has a \hat{S}_i smaller than \hat{S}_{\min} , then replace it by \hat{S}_{\min} .

The rescaled values for \hat{S}_i , or \hat{S}_i^* , can then be used to compute π_i using *Equation 4.2*. *Figure 4.2* illustrates the minimum suitability approach.

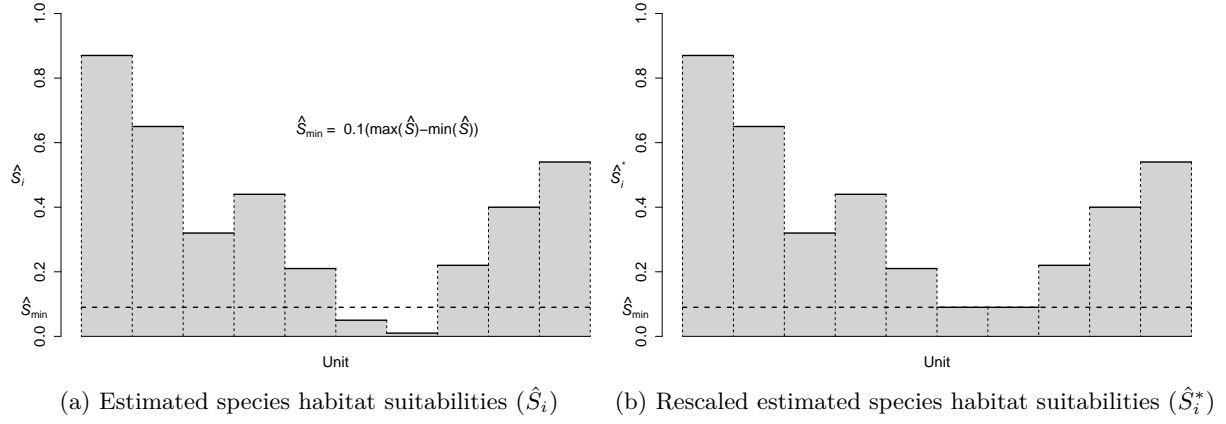


Figure 4.2: Illustration of the minimum suitability approach with $l = 0.1$.

2. The conservative π_i approach: The second method rescales the inclusion probabilities to a value greater than zero. This is done by rescaling the inclusion probabilities such that

$$\pi_i^* = a \frac{n}{N} + (1 - a)\pi_i, \quad (4.4)$$

where a is set to 0.5 for each survey. This method is called the conservative π_i approach since for each survey all units have an inclusion probability of at least $\frac{n}{2N}$. We call $\frac{n}{2N}$ the base inclusion probability. Note that this base inclusion probability is half of the value of the inclusion probability in the case of equal probability sampling.

3. The progressive π_i approach: This method is similar to the conservative π_i approach. The difference is that in the progressive π_i approach ‘ a ’ in *Equation 4.4* is replaced with

$$a = \frac{1}{k + 1}, \quad (4.5)$$

where k is the survey number. This method is called the progressive π_i approach since the base inclusion probability decreases over time. Hence, over time more weight will be given to set the inclusion probabilities based on \hat{S}_i . *Figure 4.3* illustrates the conservative π_i approach and the progressive π_i approach.

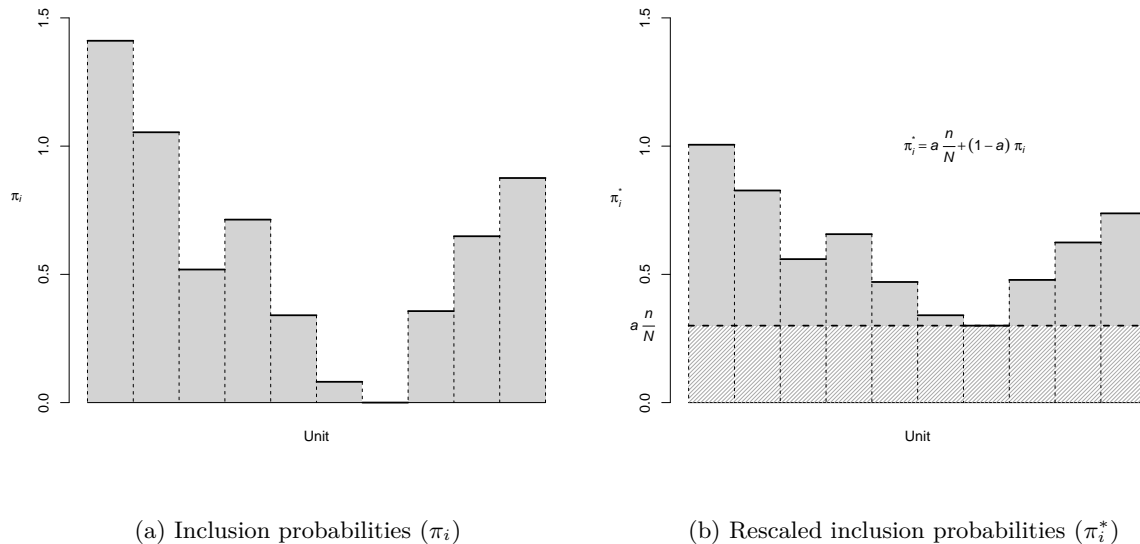


Figure 4.3: Illustration of the conservative π_i approach and the progressive π_i approach.

4.3 Simulation Study

A simulation study was performed to evaluate the performance of adaptive monitoring when using ecological information as proposed in this chapter. It was evaluated which rescaling technique provided better results. Moreover, the use of different sampling designs for adaptive monitoring was evaluated. AMStratRS was compared with adaptive monitoring that use non-stratified sampling designs. Additionally, it was reviewed whether the spatially balanced sampling designs perform better than the other sampling designs for adaptive monitoring or not.

4.3.1 Asian tiger mosquito Data

An artificial data set of the distribution of the Asian tiger mosquito in the South Island of New Zealand (*Aedes albopictus*) was chosen for this simulation study. This data set was made available by the Bio-protection group at Lincoln University, New Zealand and we are very grateful for the assistance of Professor Sue Warner and Senait Senay for providing access and support to this data. The Asian tiger mosquito is an invasive species which has recently been detected at

some of the main ports in New Zealand (Pitt et al., 2009). The Asian tiger mosquito is capable of carrying several serious human illnesses such as Dengue fever and Yellow fever. A dispersal model was used to generate future species distribution and densities. This dispersal model was simulated over a heterogeneous habitat suitability layer obtained by species distribution modelling. The locations of the initial dispersers in year one were chosen from cargo ports in the South Island that are similar in both the operation size and the habitat suitability with other international ports from which Asian tiger mosquito introductions were reported. Five independent introduction points were chosen accordingly as the initial dispersers which hypothetically could start the species dispersal. Species specific parameters such as birth and mortality rate were used to specify the rate of local spread of the species using population growth and spread rates found in the literature. Both short distance and long distance dispersal were included in the dispersal model. Details on the species distribution modelling and dispersal modelling used to generate the distribution maps can be found in Pitt et al. (2009) and Senay et al. (2013).

Figure 4.4 shows the artificial (worst case scenario) prediction for a possible invasion by the Asian tiger mosquito of the South Island of New Zealand between the years 2015 and 2018. In these maps, the South Island was tessellated into square units of three by three kilometres, in total 9716 units. The number of units with the species present for the years 2015, 2016, 2017 and 2018 were respectively 209, 415, 619 and 1118. These maps illustrate how the population of the Asian tiger mosquito is expected to expand over time. We will use these four distribution maps for our simulation study.

Auxiliary information on 20 covariates was used. These covariates were geological, climatological or ecological covariates such as altitude, rainfall or type of habitat. Details about these covariates can be found in Appendix A.

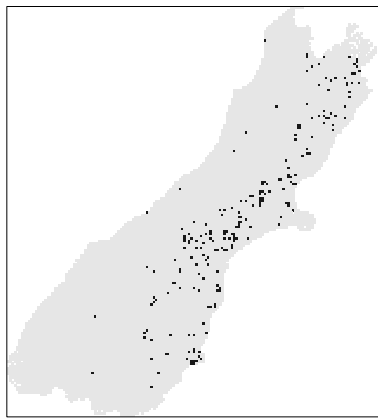
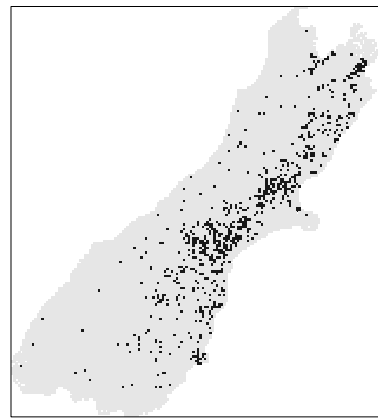
(a) Year 2015; $Y = 209$ (b) Year 2016; $Y = 415$ (c) Year 2017; $Y = 619$ (d) Year 2018; $Y = 1118$

Figure 4.4: Artificial distribution of a possible invasion of the South Island of New Zealand by the Asian Tiger mosquito between the years 2015 and 2018. These figures illustrate that the species population is not uniformly distributed of the South Island of New Zealand, with the majority of the invaded units on the East coast.

4.3.2 Testing the use of different Sampling Designs and Rescaling Techniques

In this simulation study the performance of a non-stratified sampling design for adaptive monitoring is compared with an AMStratRS method. Note that this simulation study mainly focuses on the rescaling approaches, in the next section we will repeat a similar simulation study to separate the advantages of spatial coverage from the advantage of having an adaptive sampling design. The following monitoring methods were evaluated:

1. Simple Random Sampling (SRS): An independent simple random sample was selected in each year. Note that this monitoring method is non-adaptive.
2. Adaptive Monitoring using StratRS (AMStratRS): This is adaptive monitoring using StratRS and SDMs, as was described in Guisan et al. (2006). No prior information was assumed before the first survey. Hence, in the first year an equiprobable sample was selected using SRS. Logistic regression was used as SDM. After each survey, S_i was estimated using the cumulative data from all previous years. Based on their \hat{S}_i s each unit was labelled suitable or unsuitable using a threshold value to dichotomize \hat{S}_i . The threshold value we used was obtained by maximum kappa estimation (Cohen, 1960). Maximum kappa is frequently used in species distribution modelling to illustrate a model's predictive capacity. We used maximum kappa since this was also used in Guisan et al. (2006). Other methods can be used to define a threshold value such as a fixed threshold (Manel et al., 1999), the minimal predicted area threshold (Engler et al., 2004) or other data driven thresholds. For a review on threshold selection in species distribution modelling see Nenzén and Araújo (2011). After this stratification two different allocation schemes were used:
 - (a) 50\50 : Half of the units were allocated to the suitable stratum, the other half were allocated to the unsuitable stratum, this is identical to the allocation method used in Guisan et al. (2006). Note that there are likely more efficient sampling allocation schemes, and it is also likely that better methods exist to stratify the area in two (or more) strata. Nevertheless it was decided to replicate the method described in Guisan et al. (2006). The reason why they used this method is as follows: First, after each survey they redefine the boundaries of the suitable stratum and the

non-suitable stratum. Because they used the maximum kappa method as a cut-off for the species suitability, the suitable stratum becomes smaller and smaller, but at the same time also more and more accurate at grouping high density units into one stratum. Hence, by selecting 50 percent of the units from a smaller area (assuming the suitable stratum becomes smaller) one can expect that the detection rate will increase. Second, Albert (2010) wrote a review on adaptive monitoring methods claiming that even though unequal probability sampling would likely perform better the reason why many studies stratify the area into a suitable and an unsuitable stratum is because unequal probability sampling is not as straightforward for many ecologist to implement as is random stratified sampling.

Note that if one of the strata is smaller than the number of samples that should be allocated to that strata, then first all the units in that stratum are selected. Next units are randomly selected from the other stratum until the required (total) sample size is reached.

- (b) PROP : This is proportional allocation. The units were allocated proportional to the strata sizes.

3. Adaptive monitoring using BAS (AMBAS): Adaptive monitoring using species distribution modelling, as shown in *Figure 4.1* and as explained in *Section 4.2*. No prior information was supposed to be available before the initial survey and therefore an equiprobable BAS sample was selected for the first year. Logistic regression was used as SDM. The species distribution model was constructed using the cumulative data from all previous years. The inclusion probabilities were set proportional to \hat{S}_i . BAS is used to select the samples in each consecutive year. To compared the effects of the different rescaling approaches, we repeated the simulation study for AMBAS using the following rescaling approaches:

- (a) The minimum suitability approach with $l = 0.05$ (LIM05).
- (b) The minimum suitability approach with $l = 0.10$ (LIM10).
- (c) The minimum suitability approach with $l = 0.15$ (LIM15).
- (d) The conservative π_i approach (CONS).

- (e) The progressive π_i approach (PROG).

Each monitoring method was performed for four consecutive years. Based on the distribution maps shown in *Figure 4.4a*, three different species distribution scenarios were evaluated:

1. Low density scenario: The mosquitoes distribution is stationary over time and has a low species density. The predicted distribution of the Asian tiger mosquito in the year 2015 was used for the four consecutive years. In this scenario the species is not in equilibrium with its habitat. A species is in equilibrium with its habitat if the species occupies all units that have a S_i which is above the tolerance level for that species (Elith and Leathwick, 2009). In this case, this means that the majority of the occupied units are having a high S , since those units are usually occupied first. However, not all units with a high S are already invaded.
2. High density scenario: The species distribution is stationary over time, with a high species density. The predicted distribution of the Asian tiger mosquito in the year 2018 was used for the four consecutive years. In this scenario the species is in an advanced state of equilibrium with its habitat for all years. Hence, most units with a high S are invaded by the species and some occupied units will have an intermediate to low S .
3. Increasing density scenario: The mosquitoes distribution changes over time. In this case the density increases, using the data from year 2015 to 2018. In this scenario the species population expands toward equilibrium over time. This non-stationary scenario is more realistic than the previous two stationary scenarios.

To compare these methods, Monte Carlo simulations were performed: The four surveys long monitoring scheme was repeated 1000 times for each method. Each year five percent of the units were sampled. For each iteration j and for each year k , the detection rate $D_{j,k}$ and the estimated population total $\hat{Y}_{j,k}$ were calculated. Based on these Monte Carlo simulations, with

$m = 1000$ is the number of simulations, the mean detection rate \bar{D}_k calculated per year k is

$$\bar{D}_k = \frac{1}{m} \sum_{j=1}^m D_{j,k}, \quad (4.6)$$

as well as the survey variance of the calculated detection rates for year k $\hat{V}_{\text{SIM}}(D_k)$

$$\hat{V}_{\text{SIM}}(D_k) = \frac{1}{m} \sum_{j=1}^m (D_{j,k} - \bar{D}_k)^2. \quad (4.7)$$

Additionally, the following performance statistics of the population estimation were calculated:

The observed bias of the estimated total number of individuals, \bar{B}_k , for the year k is

$$\bar{B}_k = \frac{1}{m} \sum_{j=1}^m (\hat{Y}_{j,k} - Y_k), \quad (4.8)$$

and the survey variance of the estimated population total $\hat{V}_{\text{SIM}}(\hat{Y}_k)$

$$\hat{V}_{\text{SIM}}(\hat{Y}_k) = \frac{1}{m} \sum_{j=1}^m (\hat{Y}_{j,k} - Y_k)^2. \quad (4.9)$$

Results: Simulation Study

The results of the simulation study are summarized in *Table 4.1*. The results of the achieved detection rates are discussed first. After that, the results of the population estimation are discussed.

Detection rates To begin, the results are discussed ignoring the differences between the different rescaling techniques. For all three scenarios, in year one, an equiprobable sample was selected for each evaluated sampling design. The observed mean detection rate for year one was similar for all sampling methods. This result confirms the results we found for equiprobable

sampling in the previous chapter. AMBAS had less variable detection rates than those with SRS and AMStratRS for year one. This result also confirms the results from the previous chapter. In the previous chapter it was illustrated that the spatially balanced sampling designs achieve less variable detection rates when selecting an equiprobable sample as compared with selecting a sample using SRS and StratRS (with a small number of strata).

From year two onwards, the adaptive monitoring techniques, AMStratRS and AMBAS, achieved consistently higher detection rates than SRS. This is because these adaptive designs allocate most of their sampling effort in those areas with an expected higher species densities, whereas SRS randomly selects units over the study area. In general, the detection rates for AMBAS were at least twice as high as for SRS. AMBAS achieved similar detection rates compared with AMStratRS. When using AMStratRS, the allocation technique 50 \ 50 achieved higher detection rates compared with AMStratRS using PROP. This is since with PROP the allocation over the strata is proportional to the size of the stratum. Hence, this does not mean that the (usually smaller) high suitability stratum will be sampled more intensively. However, the proportional allocation does lead to a more spatially balanced sample compared with SRS. This increase in spatial balance explains the slightly higher mean detection rates that are achieved by AMStratRS using PROP, as to the mean detection rates achieved by SRS. Interestingly the observed variances in the detection rates do not change much over time for the AMBAS methods after the second survey. One possible explanation is that this suggests that, in this simulation study, once a first batch of auxiliary information becomes available it is possible to set the inclusion probabilities relatively accurate. However, after that initial survey more information, gained from the following surveys, seems to add little to improve the detection rates of consecutive surveys.

For the low species density scenario, AMBAS achieved higher detection rates compared with SRS and generally there was only a small difference in the achieved detection rates compared with AMStratRS. In general those AMBAS methods that had a higher proportion of surveillance sampling, and thus allocated more units to the likely unsuitable areas performed worse compared with StratRS. The AMBAS methods that had a smaller proportion of the sample allocated

as surveillance sampling, such as AMBAS using CONS or LIM15 had marginally better mean detection rates than AMStratRS. However, AMBAS had generally smaller variances in the observed detection rates compared with the AMStratRS methods. For both AMBAS and AMStratRS, after the initial jump in detection rates from year one to year two, the increase in detection rates reduced. For example, the detection rate for LIM05 increased from 2.7 percent in the first survey to 13.6 percent in the second survey. The detection rate only increased with an additional 3.6 percent in the third survey and 3.9 percent in the fourth survey. This indicates that if a population is stationary or does not change much, the information gained from the initial survey is most important when updating the monitoring strategy. Additional information, gained during later surveys, is likely to have a smaller impact on the achieved detection rates. Note that the results for StratRS are only relative, since we only evaluated two types of stratification. If one would add more strata, change the strata boundaries or use a different sampling allocation scheme the results for AMStratRS may significantly improve. Additional simulation studies are needed to test this.

For the high density scenario, similar results were observed as for the low density scenario. Both AMBAS and AMStratRS had higher mean detection rates compared with SRS but the difference in mean detection rates between AMBAS and AMStratRS were small. Of course, since more units are occupied in the high density scenario as compared with the previous scenario, the values for the mean detection rates were higher. Another difference is that the increase in the mean detection rate over time is more constant over the consecutive surveys, compared with the low density scenario. A possible explanation is that in the case of the high density scenario, the mosquitoes are more in equilibrium with its habitat. This means that most of the highly suitable units are invaded. In addition, given the limited number of these highly suitable units, some units with a lower species habitat suitability are occupied by the species as well. For the low density scenario, after the initial survey, enough ecological information is available to predict which units will have a high S . However, it seems that when units with a lower S are invaded as well, more detailed and more accurate predictions are necessary to predict which of these units with a lower S are invaded and which are not invaded. Since more information becomes available over time, we see a more gradual increase in the detection rates for this high density

scenario.

For the third scenario, in which the density of the invasive species population increases, AMBAS seemed to achieve generally little higher mean detection rates after the second survey compared with AMStratRS and also has the most precise detection rates. This indicates that AMBAS works well for this particular non-stationary population example. An explanation why AMStratRS 50/50 performed not as well compared with AMBAS is because it is likely more difficult to set strata boundaries when the distribution of the species population changes over time. In Chapter 6 we will evaluate AMBAS for non-stationary populations in more detail.

These results show that by using an adaptive sampling design one can achieve higher detection rates. The observed differences in mean detection rates are small between the AMBAS methods and StratRS (50/50). On the other hand AMBAS does achieve less variable detection rates compared with StratRS which is an interesting advantage. A possible reason why AMBAS generally has less variable detection rates than AMStratRS is that no partitioning of the species habitat suitability for stratification is required when using AMBAS. In the case of (AM)StratRS, units within the same stratum obtain the same inclusion probabilities. Defining the strata boundaries and deciding on the sampling allocation over those strata can have a strong impact on the survey's detection rate and population estimates. Incorrect partitioning may result in a small, assumed 'suitable' strata, with few individuals present or a large, assumed 'unsuitable' strata, containing most of the individuals. With AMBAS unit specific inclusion probabilities are set, which means that each unit can be assigned a different inclusion probability. These unit specific inclusion probabilities make AMBAS more robust against the effect of misclassification, because misclassification can happen for a single unit or a small subset of units. Because of this, the detection rate is likely higher for AMBAS than that with AMStratRS. Again, it is important to notice that all these results are conditional on the chosen methods for AMStratRS. As previously discussed, by changing the stratification or the allocation scheme the results for StratRS could likely be improved.

The rescaling approaches are necessary to avoid (near) zero inclusion probabilities. For example,

in the first scenario, after the initial survey, on average 64 percent of the units had an \hat{S}_i less than 0.01 and thus would have near zero (or relatively small) inclusion probabilities. We evaluated five different rescaling approaches for AMBAS to avoid (near) zero inclusion probabilities. AMBAS, using any of these five rescaling approaches, gave generally similar mean detection rates compared with AMStratRS, sometimes marginally better and sometimes worse, but had smaller observed variances in the achieved detection rates. The results of AMBAS did not depend much on the chosen rescaling technique. Overall, LIM05 achieved often the highest or nearly the highest mean detection rate compared with the other rescaling approaches. The rescaling approaches CONS and LIM15 generally had marginally lower detection rates compared with the other rescaling techniques. This is because these two techniques use a higher proportion of the sample size as a surveillance sample and thus a smaller number of units will be sampled in expected high density areas. All the rescaling approaches had similar variances for the simulated detection rates.

Population estimation The results of the population estimation show that, except for the first year, AMBAS had more biased population estimates than SRS and AMStrat. The results for AMStratRS using PROP were similar to SRS. This is because AMStratRS using PROP selects an equiprobable sample, similar to SRS, and the sampling effort is not focussed on potentially high species density areas. AMStratRS using 50/50 generally had similar or even better results compared with the other methods. For AMBAS the average bias was mainly negative, which means that the true population total was underestimated. The differences in bias and precision among the rescaling approaches for AMBAS were small.

The reason for the poor performance of AMBAS is that due to the rescaling approaches, it forces the inclusion probabilities to be none proportional with the response y_i , as seen in the previous chapter. The rescaling approaches generally force all low suitability units to have a relatively higher inclusion probability.

AMBAS seemed to achieve marginally smaller observed variances compared with AMStratRS. Nevertheless, these results indicate, given the observed bias, that AMBAS does not perform well for population estimation at all. On the other hand, the fact that AMStratRS 50/50 performed best in this simulation study is advantageous for the stratified methods.

Table 4.1: Results of the first simulation study on the Asian tiger mosquito population. The value of the best performing method are in bold (per row)

	Year	SRS	AMStratRS		AMBAS						
			50\50	PROP	LIM05	LIM10	LIM15	CONS	PROG		
Low density	\bar{D}	1	0.026	0.026	0.026	0.027	0.026	0.026	0.027	0.026	
		2	0.026	0.123	0.027	0.134	0.125	0.133	0.134	0.133	
		3	0.027	0.142	0.028	0.162	0.141	0.135	0.122	0.139	
		4	0.026	0.181	0.027	0.201	0.183	0.174	0.173	0.188	
	$RV_{\text{SIM}}(D)$	1	1.000	0.989	0.978	0.728	0.725	0.738	0.732	0.721	
		2	1.000	0.988	0.952	0.745	0.783	0.845	0.776	0.744	
		3	1.000	1.009	0.899	0.722	0.794	0.731	0.808	0.717	
		4	1.000	1.052	0.904	0.694	0.725	0.743	0.785	0.736	
	\bar{B}	1	-1.541	-0.699	1.015	-1.024	0.502	-0.908	1.201	1.391	
		2	2.123	2.195	-0.642	-3.548	-4.124	-5.561	-3.981	-4.699	
		3	2.234	-1.805	-3.243	-2.997	-2.864	-3.512	-5.744	-4.123	
		4	-2.215	-1.215	-2.087	-3.197	-5.874	-6.937	-3.109	-3.553	
	$RV_{\text{SIM}}(Y)$	1	1.000	0.993	1.000	0.959	0.942	0.945	0.948	0.954	
		2	1.000	1.050	0.935	0.895	0.857	0.950	0.847	0.829	
		3	1.000	0.979	0.960	0.967	0.979	0.935	0.971	0.981	
		4	1.000	1.090	0.959	0.959	0.935	0.973	0.949	0.963	
	High density	\bar{D}	1	0.115	0.115	0.116	0.115	0.115	0.116	0.115	0.115
			2	0.115	0.188	0.118	0.185	0.186	0.183	0.187	0.186
			3	0.115	0.199	0.118	0.210	0.205	0.190	0.168	0.205
			4	0.115	0.229	0.120	0.273	0.274	0.259	0.224	0.271
$RV_{\text{SIM}}(D)$		1	1.000	0.999	0.991	0.888	0.887	0.878	0.888	0.889	
		2	1.000	1.008	0.952	0.615	0.783	0.745	0.771	0.754	
		3	1.000	1.239	0.929	0.762	0.744	0.731	0.808	0.757	
		4	1.000	1.152	0.930	0.714	0.745	0.783	0.785	0.806	
\bar{B}		1	-1.148	0.161	1.215	1.001	-0.829	0.882	-0.490	-0.810	
		2	1.146	-2.002	-0.642	-3.548	-4.124	-5.561	-4.481	-4.699	
		3	0.986	0.687	-2.243	-2.997	-2.864	-3.512	-3.744	-4.123	
		4	-0.551	0.299	-1.987	-3.197	-3.874	-2.937	-3.109	-3.553	
$RV_{\text{SIM}}(Y)$		1	1.000	1.001	1.000	0.959	0.942	0.945	0.968	0.974	
		2	1.000	1.089	0.935	1.052	1.050	0.999	1.020	0.991	
		3	1.000	0.918	0.960	0.971	0.935	0.935	1.079	0.971	
		4	1.000	0.901	0.959	0.947	0.997	0.973	0.935	1.003	
Increasing density		\bar{D}	1	0.026	0.026	0.026	0.027	0.027	0.027	0.026	0.026
			2	0.043	0.063	0.044	0.101	0.098	0.086	0.099	0.100
			3	0.064	0.170	0.065	0.229	0.231	0.201	0.188	0.210
			4	0.115	0.221	0.129	0.302	0.268	0.247	0.239	0.281
	$RV_{\text{SIM}}(D)$	1	1.000	0.987	0.981	0.751	0.749	0.750	0.751	0.777	
		2	1.000	1.108	0.922	0.775	0.783	0.745	0.776	0.769	
		3	1.000	1.239	0.949	0.882	0.744	0.731	0.878	0.857	
		4	1.000	1.452	0.940	0.842	0.855	0.983	0.985	0.836	
	\bar{B}	1	0.531	1.203	1.015	0.551	0.005	-0.508	-0.801	-0.588	
		2	1.102	1.449	0.594	-2.145	-2.896	0.002	-2.986	-3.138	
		3	0.201	1.007	-0.108	-2.997	-2.864	-3.512	-3.548	-4.002	
		4	-0.054	-0.055	-0.323	-3.481	-4.699	-3.548	-3.109	-4.124	
	$RV_{\text{SIM}}(Y)$	1	1.000	0.998	1.000	0.959	0.942	0.945	0.998	0.974	
		2	1.000	1.031	0.930	0.945	0.957	0.950	0.967	0.929	
		3	1.000	0.989	0.957	0.967	0.979	0.935	0.971	0.981	
		4	1.000	0.968	0.989	0.959	0.935	0.973	0.949	0.963	

Conclusion: Simulation Study

From this simulation study we can conclude that AMBAS and AMStratRS 50/50 achieved overall the highest mean detection rates. AMBAS generally achieved the greatest precision in observed detection rates. When comparing the rescaling approaches, AMBAS with the rescaling technique LIM05 gave generally the highest (or near highest) mean detection rates. Based on these results, we will use LIM05 as our preferred rescaling technique for the next simulation studies.

The results for the population estimation study showed that AMStratRS 50/50 generally resulted in good population estimates with smaller observed bias and better precision compared with SRS and AMBAS. AMBAS performed very poorly for population estimation. Hence, from here on, we will no longer evaluate AMBAS for population estimation.

4.3.3 Test for Spatially Balanced Sampling Designs vs Non-spatially balanced Designs.

In the previous simulation study we compared Guisan's AMStratRS with AMBAS. AMBAS uses BAS to select a sample in each survey. In the previous chapter, we illustrated the well known fact that a spatially balanced sampling design achieves higher detection rates compared with a non-spatially balanced sampling designs is, as expected, also applicable to BAS.

Based on these results, we performed a second simulation study to compare adaptive monitoring using GRTS, BAS and CP. This was done for two reasons. First, to evaluate whether the observed advantages of using AMBAS, a non-stratified sampling design for adaptive monitoring, were only because BAS is a spatially balanced sampling design. We will evaluate if the use of an unequal probability and non-spatially balanced sampling design such as CP could improve the detection rates as opposed to AMStratRS. Second, to review if spatially balanced sampling designs perform better than the non-spatially balanced sampling designs. Hence, the results of adaptive monitoring using BAS and GRTS were compared with the results of adaptive monitoring using CP.

The simulation study on the Asian tiger mosquito data was repeated using the following methods:

1. SRS.
2. AMStratRS: using PROP and 50\50.
3. Adaptive monitoring using BAS: This is identical to AMBAS (using LIM05) in the previous simulation study
4. Adaptive monitoring using GRTS: This is identical to AMBAS (using LIM05), but using GRTS to select samples instead of BAS.
5. Adaptive monitoring using CP: This is identical to AMBAS (using LIM05), but using CP to select samples instead of BAS.

These simulations were repeated for the same three scenarios and five percent of the units were selected in each survey. Only the test statistic for the detection rates were computed: \bar{D}_k and $\hat{V}_{\text{SIM}}(D_k)$.

The results of this second simulation study are summarized in *Table 4.2*. These results illustrate that adaptive monitoring using any of the three non-stratified sampling designs, thus using BAS GRTS or CP, all performed similar. Adaptive monitoring using any of these three non-stratified sampling designs generally achieved higher and less variable detection rates as compared with AMStratRS. When comparing the spatially balanced sampling designs, BAS and GRTS, with the non-spatially balanced sampling design, CP, the spatially balanced designs consistently had little higher mean detection rates. These results confirm the results from the previous chapter that spatially balanced sampling designs can achieve higher detection rates than the non-spatially balanced sampling designs. The results for BAS and GRTS were similar, with generally a small advantage for BAS over GRTS. The spatially balanced sampling designs had generally lower variances compared with SRS and AMStratRS. The observed variances for CP were generally the lowest out of the three tested spatially balanced sampling designs. The observed variances for GRTS and BAS were very similar in general, with a small advantage for BAS in some cases.

Table 4.2: Results of the second simulation study on the Asian tiger mosquito population. The value of the best performing method are in bold (per row)

		Year	SRS	AMStratRS		Non-stratified designs		
				50\50	PROP	BAS	GRTS	CP
Low density	\bar{D}	1	0.026	0.026	0.026	0.027	0.027	0.026
		2	0.026	0.123	0.027	0.136	0.137	0.117
		3	0.027	0.142	0.028	0.163	0.162	0.161
		4	0.026	0.181	0.027	0.201	0.199	0.188
	$RV_{\text{SIM}}(D)$	1	1.000	0.989	0.978	0.728	0.730	0.798
		2	1.000	0.988	0.952	0.745	0.752	0.859
		3	1.000	1.009	0.899	0.722	0.720	0.805
		4	1.000	1.052	0.904	0.694	0.702	0.790
High density	\bar{D}	1	0.115	0.115	0.116	0.116	0.116	0.115
		2	0.115	0.188	0.118	0.185	0.186	0.184
		3	0.115	0.199	0.118	0.211	0.209	0.194
		4	0.115	0.229	0.120	0.273	0.267	0.258
	$RV_{\text{SIM}}(D)$	1	1.000	0.999	0.991	0.888	0.894	0.959
		2	1.000	1.008	0.952	0.615	0.721	0.819
		3	1.000	1.239	0.929	0.762	0.760	0.806
		4	1.000	1.152	0.930	0.714	0.715	0.730
Increasing density	\bar{D}	1	0.026	0.026	0.026	0.027	0.027	0.026
		2	0.043	0.063	0.044	0.101	0.104	0.093
		3	0.064	0.170	0.065	0.229	0.210	0.209
		4	0.115	0.221	0.129	0.302	0.292	0.273
	$RV_{\text{SIM}}(D)$	1	1.000	0.989	0.978	0.728	0.730	0.798
		2	1.000	1.108	0.922	0.775	0.770	0.818
		3	1.000	1.239	0.949	0.882	0.886	0.900
		4	1.000	1.452	0.940	0.842	0.850	0.849

4.4 Discussion

Adaptive monitoring means that for each additional survey the monitoring design is updated based on the obtained information from the previous years. In this chapter we illustrated how to use ecological information to update the monitoring strategy. Based on the collected ecological information the species habitat suitability is estimated using a SDM. Next, the inclusion probabilities of all units in the study area are set proportional to the estimated species habitat suitability. A probability sample is then selected based on the set inclusion probabilities. Previous studies such as Guisan et al. (2006) followed a similar method, however they used the estimated habitat suitability to partition the study area in a suitable stratum and an unsuitable stratum. Based on this stratification of the study area a sample was selected using StratRS.

Simulation studies were performed to compare AMStratRS as used in Guisan et al. (2006) with alternative non-stratified sampling methods. Initially, we used BAS as our alternative sampling method and compared AMBAS with AMStratRS. In most of our simulation studies the detection rates for AMBAS were at least twice as high than those with SRS. When comparing AMBAS with AMStratRS the mean detection rates were generally very similar. AMBAS did achieve little smaller variances of the detection rates compared with AMStratRS 50/50.

We proposed three different rescaling approaches to avoid (near) zero inclusion probabilities. The results of the simulation study indicated that setting a lower limit on the estimated species suitability, and more particularly LIM05, was the preferred method to use in these cases. Setting a lower limit, means that the expected low density area can still be sampled, hence a surveillance sample is ensured. This surveillance sampling is necessary, especially when monitoring an invasive species to hopefully avoid new outbreaks. However, setting the lower limit to high can lead to a larger allocation of the sampling effort in the predicted low density areas and therefore a smaller portion is allocated to the high density areas. Hence, an increase of the lower limit might decrease the detection rate, which explains the results of our simulation study. Note that having a large or small surveillance sample depends on various often uncontrollable factors. For example, if the species expands over time a surveillance sample is essential to detect new and

emerging hotspots of the invasive species. Therefore it is difficult to make a general statement about the size of the best lower limit. One size does not fit all.

A second simulation study was performed to compare the differences between using spatially balanced sampling designs and using non-spatially balanced sampling designs for adaptive monitoring. The spatially balanced sampling designs achieved higher mean detection rates and a higher precision. Adaptive monitoring using BAS had marginally better results over adaptive monitoring using GRTS.

The results of the simulation study showed that AMBAS worked well for species populations that are in equilibrium (second scenario) with the habitat as well as for species populations that are not yet in equilibrium with the habitat (first scenario). The simulation study also illustrated that AMBAS worked well for both stationary (first and second scenario) and non-stationary populations (third scenario). In Chapter 6 we will evaluate AMBAS in more detail for its use for non-stationary populations.

The simulation study used in this chapter made use of binary presence-absence per unit data. We further evaluated AMBAS for count data using distribution data for *Solanum acaule*, a plant species that occurs in the high Andes of Peru and Bolivia. This data set is freely available in the package 'Dismo' of the R language (Hijmans et al., 2013). In total there are 594 *Solanum acaule* individuals spread over 3704 units. Many units had a species count of more than one, suggesting count data. However excessive zero counts were prevalent. Therefore, simple logistic regression on presence/absence per unit was chosen as species distribution modelling method. The results in *Table 4.3* indicate that in the case of count data AMStratRS 50/50 achieved the highest mean detection rates. Out of the non-stratified spatially balanced designs BAS performed best. Even though AMBAS had a lower observed variance than Guisan's AMStratRS these results are still surprising. One reason to explain this is that count information provides more powerful information to predict the species suitability than presence/absence data. Hence, the classification algorithm to stratify the units in suitable units or unsuitable units will become

more accurate and hence the allocation of sampling units to the high suitability stratum when using AMStratRS will become more accurate and precise.

Table 4.3: Results of the simulation study on the *Solanum acaule*. The value of the best performing method are in bold (per row)

	Year	SRS	AMStratRS		Non-stratified designs		
			50 \ 50	PROP	BAS	GRTS	CP
\bar{D}	1	0.035	0.035	0.035	0.035	0.035	0.035
	2	0.035	0.153	0.065	0.144	0.100	0.077
	3	0.035	0.163	0.075	0.139	0.095	0.071
	4	0.034	0.152	0.077	0.142	0.093	0.071
$RV_{SIM}(D)$	1	1.000	0.998	1.001	0.986	0.988	1.001
	2	1.000	0.924	0.989	0.829	0.840	0.948
	3	1.000	1.012	0.996	0.810	0.826	0.901
	4	1.000	1.223	0.992	0.902	0.831	0.887

Chapter 5

Adaptive Monitoring Using Spatial Information: Nearest Unit Tessellation

5.1 Introduction

In the previous chapter, ecological information was used to update the monitoring strategy. Species distribution models (SDMs) were used to model the ecological information to estimate the species habitat suitability. The estimated species habitat suitability was then used as an auxiliary variable to predict the species distribution. The inclusion probabilities for each unit were set proportional to the estimated species habitat suitability. A common assumption that SDMs make is that the modelled species is in near equilibrium with its environment (Elith and Graham, 2009). A species is in equilibrium with its environment if the species is distributed proportional to the estimated habitat suitability and occupies all sites that have a habitat suitability that is high enough for that species. However, often geographic processes are dominant over environmental ones (Elith and Graham, 2009; Lobo et al., 2010). This is the case particularly for invasive species, which by definition are not in equilibrium. This is since these species are new to an area and have not had the time yet to invade all places they can live in

(given there is some physical pathway they can follow to invade these area). In those cases, the rate of spread of a species is largely determined by geographical or spatial factors instead of the level of the species habitat suitability (Legendre, 1993; Elith and Leathwick, 2009). Examples of such geographical factors are contagious diseases, habitat connectivity, species dispersal or relevant (a)biotic interactions. These geographical factors can result in geographic clumping of the species population or spatially autocorrelated species distributions.

For example, in the previous chapter we used a virtual distribution of the Asian tiger mosquito. The initial invasion of the South Island of New Zealand by the Asian tiger mosquito was initiated from the east coast of the island. Over time the species spreads over most of the east coast but only few units are invaded on the west coast. However, the estimated species suitability is very similar for the west and the east coast. This is mainly caused by the Southern Alps, a mountain range that crosses the island from north to south (Pitt et al., 2009; Senay et al., 2013). It is difficult for the species to disperse beyond the mountain range. Because of this the west side of the island has almost not been invaded the Asian tiger mosquitoes. Hence, large areas with high suitabilities have not yet been invaded by the species. In the simulation study of the Asian tiger mosquito in the previous chapter, during each survey, samples were selected both on the west and east coast. By analysing the data it would have been easy to notice that there was a geographical or spatial effect on the species distribution. The SDM that was used to predict the habitat suitability ignored this spatial information.

Many SDMs only use the information of in which habitat a species is observed and disregard the observed spatial distribution of the species over the study area. Thereby, they often neglect information about possible deviations of the species distribution from its assumed equilibrium with the habitat (Lobo et al., 2010; Elith and Graham, 2009). This means that environmental information, or information based on species distribution models should be interpreted only as a measure of species habitat suitability, rather than a measure of the likelihood of species presence. This is especially the case when a species is unlikely to be in equilibrium with its environment. Hence, the observed spatial distribution of the species should therefore not be ignored when updating the monitoring strategy. Using spatial information is especially important when

surveying invasive or endangered species. Many of these species are not in equilibrium with the environment of the sampling area (Robertson et al., 2004; Guisan and Thuiller, 2005) and thus often do not occur at locations with the highest estimated species habitat suitability. Methods that use the observed spatial information of the species are called ‘geographical models’. The concept of geographical models is intuitive. Simply stated; units close to a sampled unit in which the species of interest was observed will be more likely to contain individuals compared to more distant units. Thus for sampling, the closer a unit is to a unit in which the species are observed the more likely it should be selected in the next survey. Thus these units should be given higher inclusion probabilities in the next survey to increase the detection rate.

Various geographical methods have been introduced to predict the species distribution (Bahn and McGill, 2007; Colwell and Rangel, 2009; Hijmans, 2012). An example are convex hull methods. A convex hull is the smallest polygon around a set of points with no angles exceeding 180 degrees. In ecology for example, a convex hull can be placed around all occupied units in a study area. The study area can then be stratified based on a unit being inside or outside the convex hull. This method was applied, for example, by Burgman and Fox (2003) to estimate the range of the distribution of *A. bonvouloiri*, a beetle species, in Europe. Another set of geographical methods is called Kriging (there are several different Kriging methods) (Krige, 1951). Kriging is based on the first law of geography, namely: ‘Everything is related to everything else, but near things are more related than distant things.’ (Tobler, 1970). Hence, Kriging techniques give nearby units similar values. Thus, if a unit is close to a known occupied unit, that unit is assigned a higher chance of finding a species in. Kerrya et al. (2013), for example, compared two types of Kriging methods for their accuracy to map the distribution of several large animals in Kruger National Park, South Africa. Other geographical techniques are spatial point processes modelling in which the expected density for each unit is predicted based on the number of occupied units in the neighbourhood of the unit (Seaman and Roger, 1996) or two dimensional kernel density estimation which is two dimensional variation of kernel density estimation.

The initial idea was to use these methods to adjust the inclusion probabilities for adaptive monitoring. We experimented with several of these geographical methods and compared their

practical usage for adaptive monitoring. It appeared that some of the previously discussed methods were not very suitable in practice for the adaptive monitoring method we introduced in Chapter 2. Other methods are difficult to use in a simulation study context. For example, convex hull methods have a binary outcome and thus simply stratify the sampling area into a suitable stratum and an unsuitable stratum. Our method for adaptive monitoring requires the assignment of inclusion probabilities on the unit level, preferably without stratification. Convex hull methods are therefore not suitable. Moreover, many studies (Burgman and Fox, 2003; Elith and Leathwick, 2009) describe how convex hull methods often give poor results, even after adding certain conditions to make the methodology more robust. Methods like Kriging and Kernel density estimation depend strongly on the chosen parameter settings (Kerrya et al., 2013; Oliver and Webster, 2014; Webster and Oliver, 1992). For example, in kernel density estimation the bandwidth selection has an important effect on the final estimated densities. Another example is the specification of the spatial correlation structure of the data, the variogram, for Kriging methods (Oliver and Webster, 2014; Webster and Oliver, 1992). Although this is not a problem as such, it does mean that each time these methods are used, optimization of the parameters is required (Oliver and Webster, 2014; Webster and Oliver, 1992). This can be a difficult, and often arbitrary, process. This case specific parametrisation makes these methods difficult to use for repeated surveying such as adaptive monitoring or when running simulation studies. Furthermore, many of these methods give poor or unrealistic results, or are difficult to fit if not enough data are available (Oliver and Webster, 2014; Webster and Oliver, 1992; Seaman and Roger, 1996). For example, the reliability of a variogram depends strongly on the size of the available data. In the case of presence/absence data, there should be at least 100 presence points available (Oliver and Webster, 2014). 100 presence units is often unrealistic to achieve when monitoring invasive species at an early stage of the invasion. Finally, methods like spatial point pattern analysis or convex hull methods use presence-only data. With adaptive monitoring, and especially in the case of invasive or rare species, the majority of the visited units are likely to be unoccupied. Hence, it can be important to include the information about unoccupied units to the geographical method.

Because of this, we tried to develop a new class of geographical methods called Nearest Unit Tessellation (NUT). The idea for NUT originated from the usage of Voronoi tessellations to measure the spatial balance of a sample. All NUT methods are based on some form of (discretised) Voronoi tessellation. In this chapter, we will illustrate and compare several NUT methods. A simulation study is conducted to show that each NUT method has specific properties and specific applications. Compared with some of the existing geographical methods, these NUT methods can be used to model the spatial information and to update the inclusion probabilities for our suggested method for adaptive monitoring. Note that the simulation study only compares the different NUT methods and does not provide any results of other potential geographical methods.

5.2 Methodology

The NUT methodology is based on discretised Voronoi cells. The Voronoi cell J for sampled unit j is the collection of all units that are nearest to unit j compared with any other unit in the sample. Equidistant units are randomly assigned to one of the potential Voronoi cells. A sampled unit occupied by the species (and assuming that the detectability is 100 percent) is called a presence unit, a sampled units unoccupied by the species is a absence unit. A Voronoi cell based on a presence unit is a presence Voronoi cell, a Voronoi cell based on an absence unit is called an absence Voronoi cell.

Here we introduce five different NUT methods. All five NUT methods follow the same three steps based on the selected sample(s):

Step 1 Tessellate the area into (discretised) Voronoi cells.

Step 2 Assign all units within the same Voronoi cell a weight ω_i .

Step 3 Set the inclusion probability π_i for each unit using the assigned weights.

The NUT methods can be different from each other at each of these three step. This will be explained in more detail later in this chapter. For example some NUT methods use presence/absence data for tessellation while others use presence data only for their tessellation. Another

difference is the way each NUT method sets the inclusion probabilities. Some methods only use two levels of inclusion probabilities based on presence\absence, other methods set inclusion probabilities with a continuous range for example based on the size of the Voronoi cell or based on the observed number of individuals in the sampled units.

To illustrate the NUT methods and to illustrate adaptive monitoring using these NUT methods, we again make use of the Rockfish data, adopted from Su and II (2003). However, this time the Rockfish data are modified differently. The original modelled population of Su and II (2003) had 20 by 20 units, $N = 400$. In this chapter, the Rockfish population is modified using bilinear interpolation to 80 by 80 units, $N = 6400$. The resulting distribution of the Rockfish is shown in *Figure 5.1*. The population of Rockfish is highly aggregated over the study area. For much of the study area, the fish were absent, but when present, their abundance peaked at 25 individuals per unit.

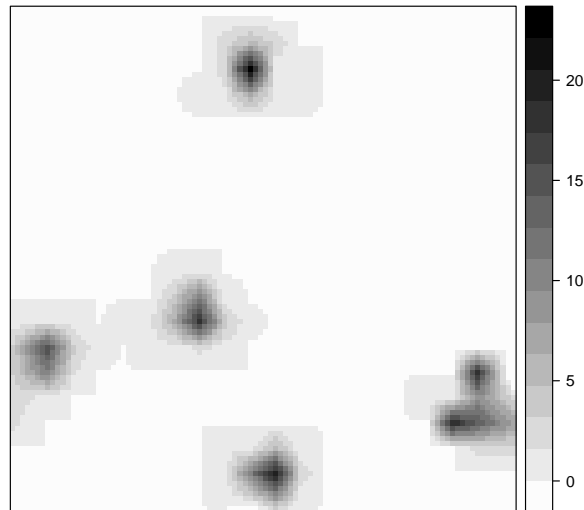


Figure 5.1: Distribution of the Rockfish population adopted from Su and II (2003), 80 by 80 units, $N = 6400$

5.2.1 Simple Nearest Unit Tessellation

The first NUT method is Simple Nearest Unit Tessellation (SNUT). SNUT uses both presence and absence units to set the inclusion probabilities. An illustration of SNUT, see *Figure 5.2a*, will make it easier to understand the basic principle of SNUT. Based on the sampled units in the previous survey(s) SNUT works as follows:

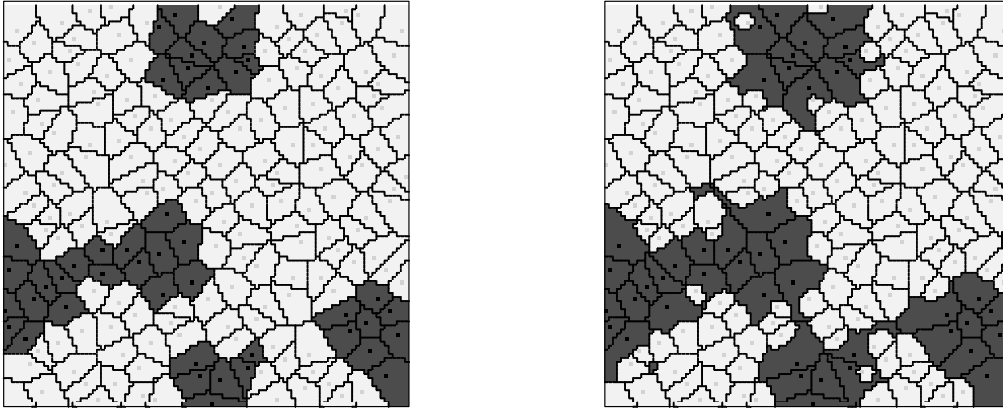
1. **Tessellate the study area into Voronoi cells:** Use both presence and absence units to construct presence and absence Voronoi cells.
2. **Assign weight ω_i to each unit:** Although the value of these weights ω_i are arbitrary, we suggest the following rule of thumb: If unit j is a presence unit, then $\omega_i = 2$ for all units in the presence Voronoi cell J . However if unit j is an absence unit, then $\omega_i = 1$ for all the units in the absence Voronoi cell J .
3. **Calculate π_i for unit:** For unit i the inclusion probability π_i can be calculated by

$$\pi_i = n \frac{\omega_i}{\sum_{i=1}^N \omega_i}. \quad (5.1)$$

Thus a unit within a presence Voronoi cell will get an inclusion probability that is double the inclusion probability of a unit in an absence Voronoi cells.

5.2.2 Unequal Nearest Unit Tessellation

A variation of SNUT sampling is Unequal Nearest Unit Tessellation (UNUT) and is illustrated in *Figure 5.2b*. The methodology of UNUT is similar to SNUT. However, for UNUT the Voronoi tessellation (step 1.) of the SNUT methodology is different and works as follows: The distance d between two units is the euclidean distance between the centres of the two units of interest. For each unit the distance d_p to the nearest presence cell p and the distance d_a to the nearest absence cell a is calculated. If for a unit $(1 - u)d_a < ud_p$, with u a value between zero and one, then that unit will be part of presence Voronoi cell P around presence unit p otherwise it will be part of absence Voronoi cell A around absence unit a . If $u = 0.5$ than UNUT is identical to



(a) SNUT

(b) UNUT

Figure 5.2: Illustration of Simple Nearest Unit Tessellation (SNUT) and Unequal Nearest Unit Tessellation (UNUT) with $u = \frac{2}{3}$. Two percent of the units are sampled. Black units are presence units, grey units are absence units. Darker Voronoi cells are presence Voronoi cells and will obtain a higher weight and thus a higher inclusion probability. Absence Voronoi cells are brighter and will obtain a lower weight and thus a lower inclusion probability.

SNUT. If $u > 0.5$, then more units will be allocated to presence Voronoi cells and thus presence Voronoi cells will absorb more units. Therefore more units will have $\omega_i = 2$. If $u < 0.5$, then more units will be allocated to Voronoi cells around absence units and thus more units will have $\omega_i = 1$. As a rule of thumb we suggest using $u = \frac{2}{3}$. The reason for this rule of thumb is as follows. After experimenting with UNUT, it was clear that UNUT was a better method than SNUT but definitely not as good compared with the NUT methods we will introduce in the next sections. However, because it is an intuitive extension of SNUT we decided to explain the method and include it in this thesis. The rule of thumb of $u = \frac{2}{3}$ was arbitrarily chosen after trying out various values for u . Nevertheless, $u = \frac{2}{3}$ is also intuitive. A value close to one will likely predict too many absence units a high inclusion probability and a value much lower than $2/3$ will decrease the advantage that UNUT offers over SNUT.

5.2.3 Presence Nearest Unit Tessellation

Presence Nearest Unit Tessellation (PNUT) only uses the spatial information from presence units, thus of those sampled units in which the species of interest was observed in the previous survey(s). A possible reason for only using presence units is that presence units provide reliable information about where a species is located. Absence units do not ensure total absence of the species. An absence unit may be a false negative due to observation error or can be an occasional unoccupied unit in an other than that high density area. Therefore absence units may be misleading observations. Another situation in which using presence only data can be useful is when dealing with citizen science data. Citizen science data can for example be the reported observations of a species by volunteers or data from achieved documentation. These reports mainly describe the occasions when somebody has actually observed the species. However, reports of a species being absence are rare, this is since not many people report not seeing a species.

PNUT is illustrated in *Figure 5.3*. PNUT works as follows:

1. **Tessellate the study area into Voronoi cells:** Only presence units are used to tessellate the study area into presence Voronoi cells.
2. **Assign weight ω_i to each unit:** The idea behind PNUT is that large Voronoi cells indicate low densities and small Voronoi cells indicate high densities. For each Voronoi cell the number of units that are part of that cell are counted. This count C can be seen as the relative size of the Voronoi cell. Set the weight equal to the reciprocal of the count, thus $\omega_i = \frac{1}{C}$.
3. **Calculate π_i for each unit:** For each unit the inclusion probability is set proportional to the weight

$$\pi_i = n \frac{\omega_i}{\sum_{i=1}^N \omega_i}. \quad (5.2)$$

Lower limit After performing PNUT, units in small Voronoi cells will have high inclusion probabilities since the inclusion probabilities are proportional to the reciprocal of the Voronoi

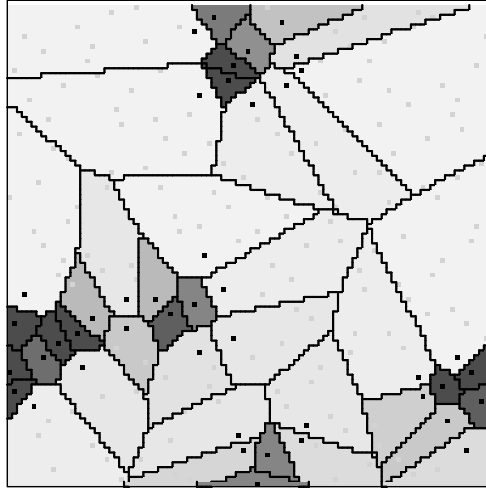


Figure 5.3: Illustration of Presence Nearest Unit Tessellation (PNUT). Two percent of the units are sampled. Black units are presence units, grey units are absence units. Units in smaller (darker) Voronoi cells will obtain a smaller relative size and thus a higher inclusion probability. Units in bigger (brighter) Voronoi cells will obtain a higher relative size and thus a smaller inclusion probability.

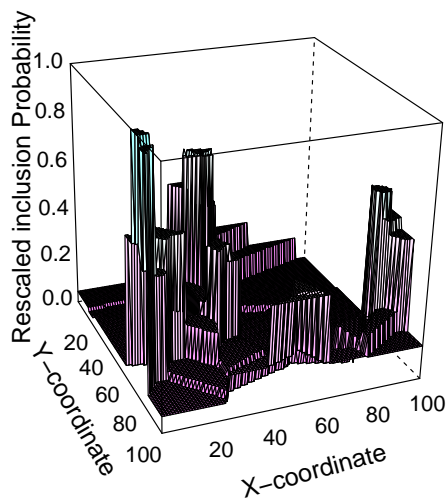
cell sizes. However, one possible drawback of PNUT is that if some Voronoi cells are very small, it will force the inclusion probabilities of the units in other Voronoi cells to be very small. Small Voronoi cells will look like sharp spikes on a 3D inclusion probability plot, see *Figure 5.4*. These extreme inclusion probabilities can be problematic because when selecting a sample it is likely that only units that are part of these small Voronoi cells will be selected.

If the sample size is higher than the number of units part of those small Voronoi cells, these extreme differences in inclusion probabilities do not have to be problematic. In that case, all the units within the small Voronoi cells will be selected first. After that, the remaining sampling units will be selected from the bigger Voronoi cells.

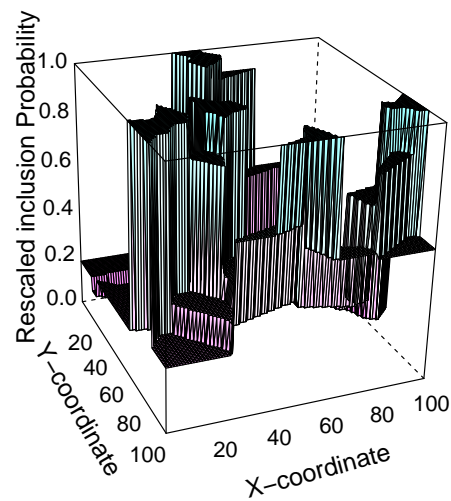
In the case when the sample size is small or if too many units have high inclusion probabilities we

suggest the following method. If the relative size C_i of a unit is smaller than a predefined lower limit C_{limit} , then replace that unit's C_i by the value of that lower limit C_{limit} . The remaining part of the PNUT sampling methodology remains the same.

A suggestion for a lower limit C_{limit} would be equal to the total number of units divided by the sample size of the performed survey, thus $\frac{N}{n}$. Note that $\frac{N}{n}$ does not have to be an integer. The logic behind this lower limit is that if both the species and the sampled units are uniformly distributed over the sampling area, then all Voronoi cells are expected to have the same size. Note that if each unit in the sampling area would be assigned this lower limit value, all inclusion probabilities would be equal to $\frac{n}{N}$ which is the same as for basic equal probability sampling.



(a) PNUT without using a lower limit



(b) PNUT using a lower limit

Figure 5.4: Example differences in inclusion probabilities when performing PNUT with and without using a lower limit for the Rockfish data. Note that the inclusion probabilities are rescaled such that the maximum is one.

5.2.4 Presence/Absence Nearest Unit Tessellation

Since PNU T only uses presence data, information provided by absence units is neglected. However, the majority of the sampling units often are absence units, this is especially the case with small and clustered populations. Presence/Absence Nearest Unit Tessellation (PANUT) is similar to PNU T but does include the information from the absence units into the geographical method. PANUT is illustrated in *Figure 5.5* and works as follows:

1. **Compute inclusion probabilities based on PNU T:** Perform PNU T using the set of presence units only to compute the inclusion probabilities, which we call $\pi_i^{\text{PNU T}}$.
2. **Compute inclusion probabilities based on Absence Nearest Unit Tessellation (ANUT):** Repeat the same steps as for PNU T but use the absence units only instead of the presence units, thus ANUT, and compute the inclusion probabilities π_i^{ANUT} using

$$\pi_i^{\text{ANUT}} = n \frac{\omega_i^{\text{ANUT}}}{\sum_{i=1}^N \omega_i^{\text{ANUT}}}. \quad (5.3)$$

3. **Set π_i^{PANUT} using $\pi_i^{\text{PNU T}}$ and π_i^{ANUT} :** To compute the final inclusion probabilities π_i^{PANUT} based on $\pi_i^{\text{PNU T}}$ and π_i^{ANUT} first calculate

$$p_i^{\text{PNU T}} = \frac{\pi_i^{\text{PNU T}}}{\max(\pi_i^{\text{PNU T}})} \quad (5.4)$$

such that $p_i^{\text{PNU T}}$ has a maximum of one, and

$$p_i^{\text{ANUT}} = \frac{\pi_i^{\text{ANUT}}}{\max(\pi_i^{\text{ANUT}})}. \quad (5.5)$$

such that p_i^{ANUT} has a maximum of one. Then use one of the following combination techniques:

- (a) The additive method:

$$p_i^{\text{PANUT}} = (p_i^{\text{PNU T}} - p_i^{\text{ANUT}}) + 1. \quad (5.6)$$

The +1 is added to ensure that p_i^{PANUT} is non-negative.

(b) The ratio method:

$$p_i^{\text{PANUT}} = \frac{p_i^{\text{PNUT}}}{p_i^{\text{ANUT}}}. \quad (5.7)$$

To get to final inclusion probabilities use

$$\pi_i^{\text{PANUT}} = n \frac{p_i^{\text{PANUT}}}{\sum_{i=1}^N p_i^{\text{PANUT}}}. \quad (5.8)$$

Note that the obtained inclusion probabilities π_i^{ANUT} are a measure of not observing the species in a unit, and are therefore not useful for probability sampling. However one could consider rescaling the reciprocal of π_i^{PNUT} to set inclusion probabilities for an unequal probability sampling. In addition, intuitively one could argue that for ANUT the relative size of the Voronoi cells C_i could be used as weights instead of the reciprocal thus

$$\pi_i^{\text{ANUT}} = n \frac{C_i^{\text{ANUT}}}{\sum_{i=1}^N C_i^{\text{ANUT}}} \quad (5.9)$$

and to obtain the final inclusion probabilities one could modify the additive method to

$$p_i^{\text{PANUT}} = p_i^{\text{PNUT}} + p_i^{\text{ANUT}}. \quad (5.10)$$

Equation 5.9 and *Equation 5.10* will set high inclusion probabilities in those areas for which less information is available about absence units. However, as illustrated in *Figure 5.5*, the main idea of PANUT is that areas with a high predicted species presence based on PNUT, will get penalized (thus reduced inclusion probabilities) if it is known that absence units are present within those areas, thus by information gained from ANUT.

5.2.5 Count Nearest Unit Tessellation

All previous NUT methods used binary presence/absence data. Count Nearest Unit Tessellation (CNUT) is a variation of NUT that uses the observed species count to set the inclusion probabilities. CNUT is illustrated in *Figure 5.6*. Based on all sampled units in the previous

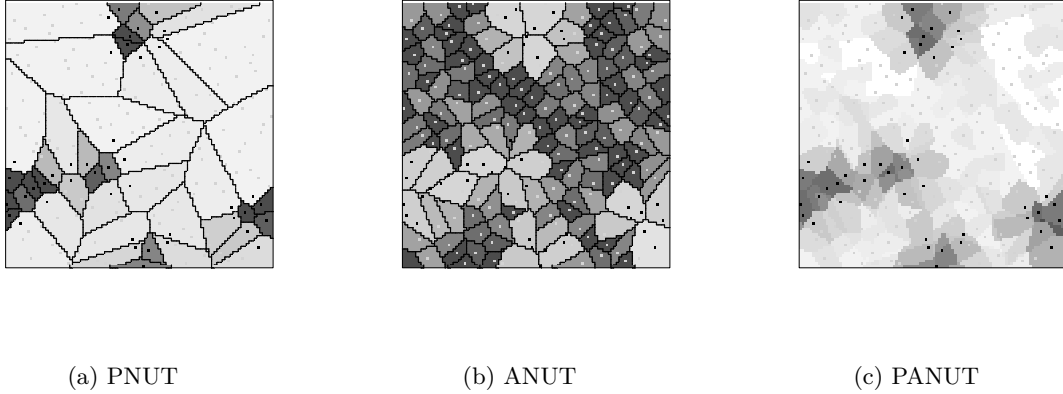


Figure 5.5: Illustration Presence/Absence Nearest Unit Tessellation (PANUT). Two percent of the units are sampled. In this case, the additive method is used for PANUT. Black units are presence units, grey units are absence units. Darker Voronoi cells for PNUT indicate a higher likelihood of observing a species and thus will increase the final inclusion probabilities. Darker Voronoi cells for ANUT indicate a higher likelihood of not observing a species and thus will decrease the final inclusion probabilities. For PANUT darker units have higher inclusion probabilities.

survey, CNUT works as follows:

1. **Tessellate the study area into Voronoi cells:** Both presence and absence units are used for the Voronoi tessellation.
2. **Assign a weight ω_i to each unit:** Set ω_i for all units in Voronoi cell J equal to the observed count of the species, y_j , in unit j .
3. **Calculate the inclusion probability for each unit:** For each unit set the inclusion probability proportional to the reciprocal of the assigned count by using

$$\pi_i^{\text{CNUT}'} = n \frac{\omega_i}{\sum_{i=1}^N \omega_i}. \quad (5.11)$$

If unit j is an absence cell then $\omega_i = 0$ for all units in Voronoi cells J and all these units will have an inclusion probability of zero. To avoid this the inclusion probabilities for each

cell are rescaled as follows:

$$\pi_i^{CNUT} = \frac{\pi_i^{CNUT'} + \frac{n}{N}}{2}. \quad (5.12)$$

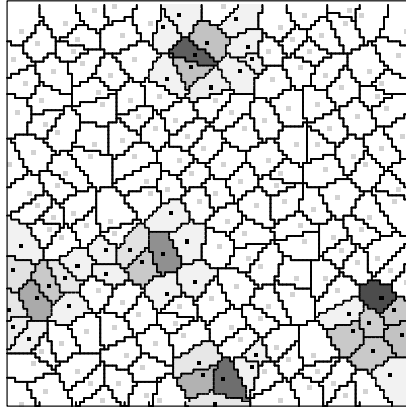


Figure 5.6: Illustration of Count Nearest Unit Tessellation (CNUT). Two percent of the units are sampled. Black units are presence units, grey units are absence units. Darker Voronoi cells have higher inclusion probabilities.

5.2.6 Adaptive Monitoring Using the NUT Methods

The NUT methods estimate the density of a species over a sampling area. Adaptive monitoring using geographical methods works as follows: Equal probability sampling is performed in the first survey. Thus $\pi_i = n/N$ for each unit. SRS is an example of an equal probability sampling method. However, in the case of these NUT methods this initial sample has to be a spatially balanced sample. This is since units selected with SRS can by chance be clustered near each other which can influence the NUT tessellation. In Chapter 3, BAS was chosen as the preferred spatially balanced sampling design. Therefore, we will use BAS to select the sample when using the NUT methodology. An equiprobable sample selected using BAS for the Rockfish data is illustrated in *Figure 5.7a*.

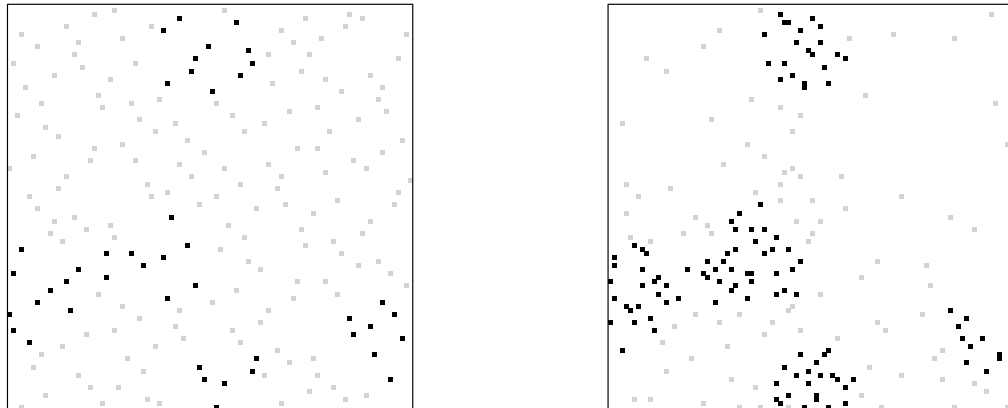
After an initial sample is taken, a NUT method is chosen and applied to set inclusion probabilities so that unequal probability sampling can be performed for the subsequent survey. The NUT methods are designed for adaptive monitoring and are therefore expected to provide good results when used in an adaptive long term context.

After setting the inclusion probabilities, unequal probability sampling can be performed to select the sample for the next survey. Since by using a spatially balanced sampling method the sampling units within each Voronoi cell will be more evenly spread. Therefore we recommend using BAS to select an unequal probability sample. An unequal probability sample selected using BAS based on the set inclusion probabilities using PANUT is shown in *Figure 5.7b*. Compared with the species distribution in *Figure 5.1*, this unequal probability sample is more densely distributed in and near areas with higher species densities.

5.3 Simulation Study

A simulation study was performed to evaluate the performance of the NUT methodologies. The simulation studies were performed using the Rockfish population and the Asian tiger mosquito population. The Rockfish population is a highly clustered population. The Asian tiger mosquito population, introduced in the previous chapter, is a less clustered or scattered population and has a non-homogeneous species distribution over the South Island. Most of the species are on the east side of the island.

Monte Carlo simulations were performed to evaluate and compare the NUT methods. The details about which NUT was used is discussed in the following subsection. For each NUT method, sampling was performed for ten consecutive years. This ten years long sampling scheme was repeated 1000 times for each method. Each year three percent of the units were sampled. For each iteration j , the detection rate $D_{j,k}$ and the population total $\hat{Y}_{j,k}$ were calculated for each year k . The mean detection rates \bar{D}_k per year k were calculated as well as the variance in



(a) Equal probability sample

(b) Unequal probability sample

Figure 5.7: Illustration of an equal probability and an unequal probability sample for the Rockfish data. For both samples two percent of the units are selected using BAS. Black units are presence units, grey units are absence units. In this case, the inclusion probabilities for unequal probability sampling are set by PANUT.

the calculated detection rates per year $\hat{V}_{\text{SIM}}(D_k)$.

5.3.1 Details of the Species Populations

Clustered Rockfish Population

The Rockfish dataset was introduced in *Section 5.2* and *Figure 5.1*. The Rockfish dataset is 80 by 80 units, $N = 6400$. Of those 6400 units, 1629 units are occupied by Rockfish, which is 25.4 percent of the units.

Scattered Asian Tiger Mosquito Population

The Asian tiger mosquito population dataset was introduced in the previous chapter. Compared with the Rockfish dataset, the Asian tiger mosquito is a scattered and less clustered population

with some presence units completely isolated from other presence units. To be able to compare the results for both species populations it is assumed that, in this chapter, the Asian tiger mosquito population is stationary, using the predicted distribution for the year 2018. In the year 2018, 15.8 percent of the units are occupied by the Asian tiger mosquito.

The distribution data of the Asian tiger mosquito is presence/absence data. However, count data are required for the simulation section of CNUT. Therefore, for the CNUT simulation study, the Asian tiger mosquito dataset was artificially transformed from presence/absence data to count data. This was done by first creating a species habitat suitability layer (using logistic regression, see Chapter 4). This suitability layer had values ranging from zero to one. For each presence unit the estimated species suitability was multiplied by 10 to get virtual counts. Counts for absence units were set equal to zero to ensure that the distribution and the number of occupied units remained unchanged. The resulting distribution and virtual counts are shown in *Figure 5.8*.

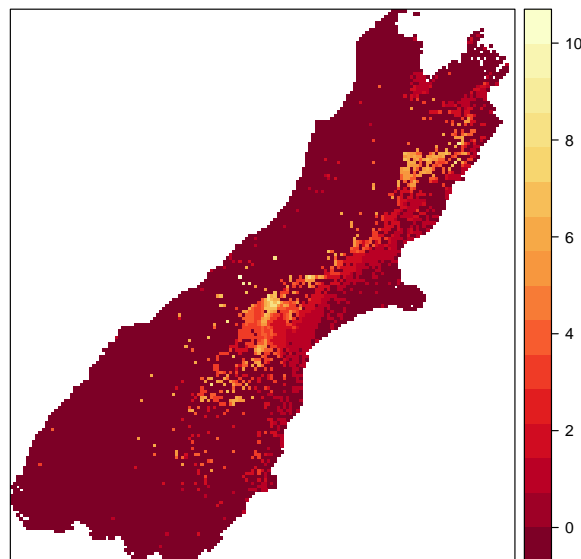


Figure 5.8: Predicted distribution and virtual count data of the Asian tiger mosquito in the South Island of New Zealand for the year 2018.

5.3.2 Results

SNUT and UNUT

In this section, we compare the results of the simulation study for adaptive monitoring when using the following geographical methods:

1. SNUT based on the cumulative data from all previous surveys (SNUT-CD).
2. SNUT based on data from the previous survey only (SNUT).
3. UNUT based on the cumulative data from all previous surveys (UNUT-CD).
4. UNUT based on data from the previous survey only (UNUT).

The results are shown in *Figure 5.9*. For both the Rockfish and the Asian tiger mosquito population the simulations using the cumulative data resulted in higher detection rates compared with those of the non-cumulative methods. The error bars are wide and tend to overlap for most methods at each year. Only the results for SNUT-CD and UNUT do not overlap from year 7 onwards for the Rockfish population. This indicates that the observed differences are likely not significantly different. For the Rockfish data SNUT performed better, whereas for the Asian tiger mosquito data UNUT performed better. In the case of a clustered population like the Rockfish data SNUT performs better than UNUT since one can expect that the bordering units of a presence unit will be presence units and the bordering units of a absence unit will be absence units. In the case of a more scattered, non-clustered population like the Asian tiger mosquito dataset, UNUT will perform better since one can assume that the proximate neighbourhood of a presence unit will contain more presence units. However these units do not have to border that presence unit. Using a similar logic, one can assume that the observed absence units could be bordering presence units or have presence units in their neighbourhood. Therefore, absence Voronoi cells in high density areas will be relatively small whereas absence Voronoi cells in low density areas will likely be relatively bigger. Therefore, choosing UNUT and thus having relatively bigger presence Voronoi cells and smaller absence Voronoi cells will result in higher detection rates for non-clustered populations.

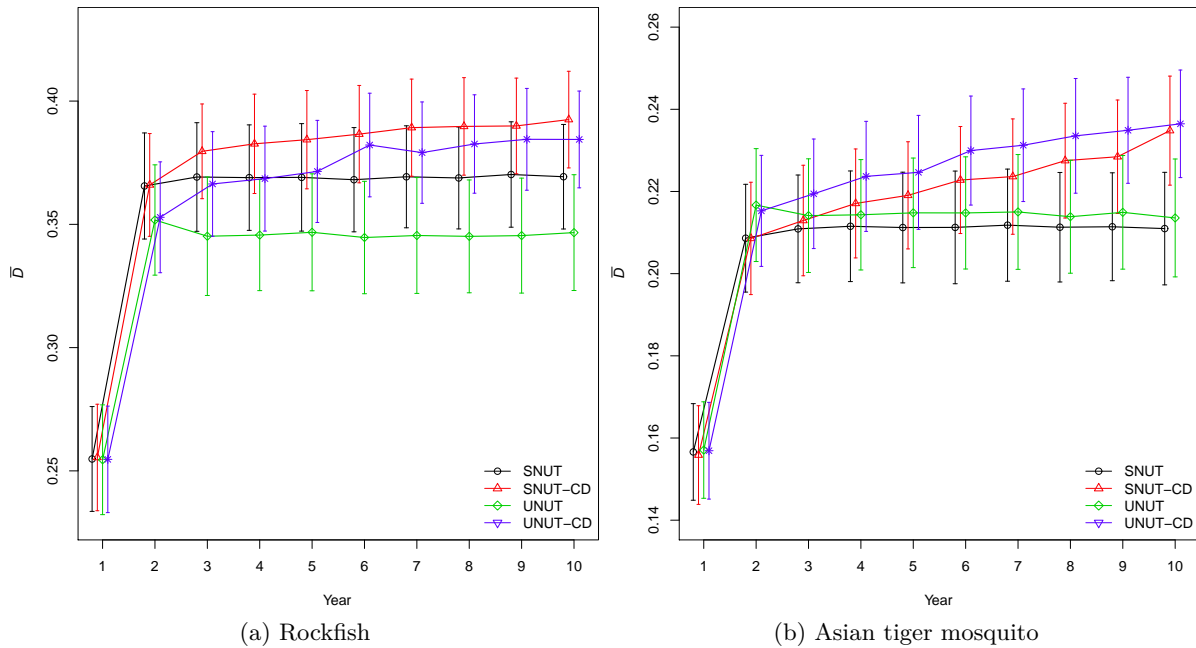


Figure 5.9: Results of the simulation study for SNUT and UNUT. The error bars are the 2.5% and the 97.5% percnetile obtained by the simulation studies.

PNUT

In this section, we evaluate adaptive monitoring using PNUT based on the cumulative data from all previous surveys with:

1. PNUT with no lower limit (PNUT-0).
2. PNUT with a lower limit of $\frac{N}{n}$ (PNUT-1).
3. PNUT with a lower limit of $0.5\frac{N}{n}$ (PNUT-0.5).
4. PNUT with a lower limit of $1.5\frac{N}{n}$ (PNUT-1.5).

The results are shown in *Figure 5.10*. For both the Rockfish and the Asian tiger mosquito populations the simulations using PNUT with no lower limit gave the best results followed by PNUT-0.5, then PNUT-1 and finally PNUT-1.5. Using PNUT without any lower limit was the only method for which the detection rates increased year after year and reached more than 80 percent for the Rockfish population and reached more than 60 percent for the Asian tiger

mosquito population. The other three methods, PNUT-0.5, PNUT-1 and PNUT-1.5 all reached the highest detection rate during the third year and then decreased. This is except for PNUT 1.5 in the Asian tiger mosquito population which surprisingly remains more or less constant from year three onwards. These results suggest that using a small lower limit or even better, using no lower limit, results in the highest detection rates. The problem is that PNUT with no lower limit, or a small lower limit is what we will call a greedy method. Greedy methods tend to allocate all their sampling effort in known high density areas, which we will call a greedy area. For PNUT, using no lower limit, once a high density area is located the inclusion probability will be much higher in that area compared with the other areas. This is as illustrated in *Figure 5.4*. Therefore this greedy area will absorb the majority of the sampling effort. Since more presence units will likely be found in that area, this absorbing effect will increase over time. The detection rate will increase because many units will be sampled in just one known high density area. However, this means that only few units are sampled elsewhere to screen the remaining areas for potential high density areas and thus no (or a very small) surveillance sample will be taken. Because of this, new invaded areas or existing hotspot are potentially not detected. Therefore, one should be careful when applying PNUT without using a lower limit.

The error bars are similar in size for each method and for each survey. The overlap in the error bars for the first two years indicate that there are only significant differences between the methods from the third year onwards.

The reason why PNUT-0.5, PNUT-1 and PNUT-1.5 (in the rockfish population only) decrease after the survey can be explained by the fact the PNUT does not take into account the absence units. Each time a presence unit is found the sampling effort will increase around that presence unit. However over time more presence units will be found in low density areas. These isolated units can be thought of as isolated presence units surrounded by absence units. Once many of those isolated presence units are discovered, which in this case is apparently after the third year, consequently more samples will be selected within actual low density areas. Because of this the detection rate might decrease again since the information from the absence units is not incorporated in the model when using PNUT.

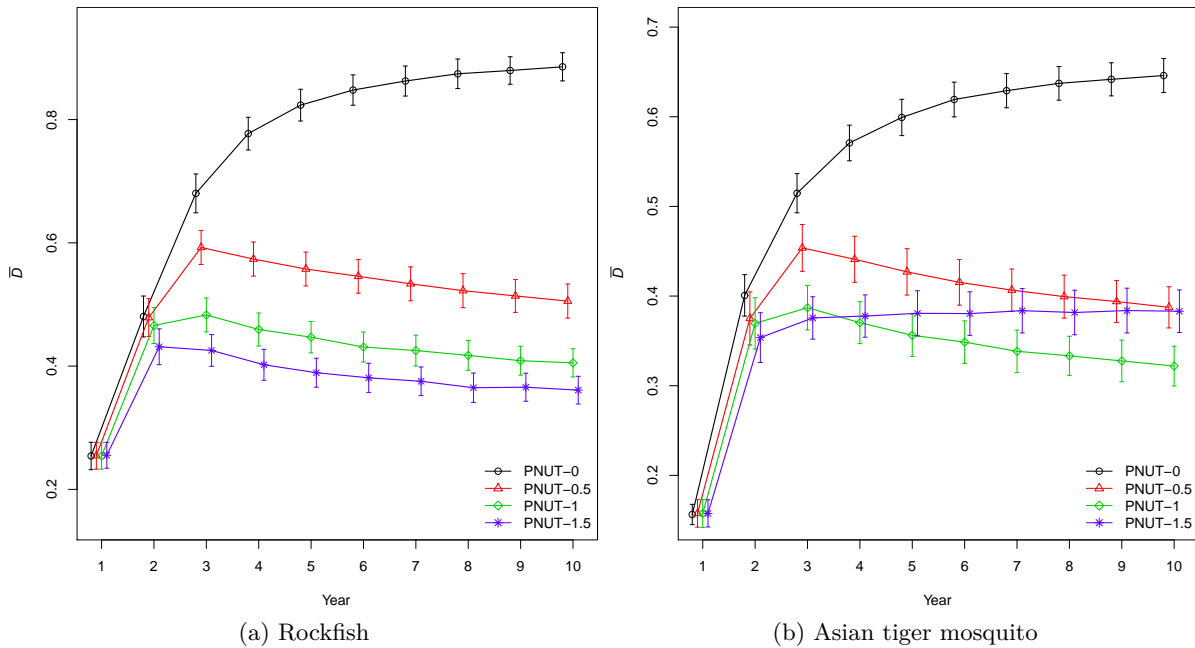


Figure 5.10: Results of the simulation study for PNUT. The error bars are the 2.5% and the 97.5% percentile obtained by the simulation studies.

PANUT

In this section we evaluate adaptive monitoring using PANUT:

1. Using the Additive method (PANUT-add) and using
 - (a) the cumulative data of both presence and absence units from all previous surveys (PANUT-add-CD).
 - (b) the cumulative information from all previous surveys for the presence units but only the information from the last survey for the absence units (PANUT-add-pres).
 - (c) the cumulative information from all previous surveys for the absence units but only the information from the last survey for the presence units (PANUT-add-abs).
 - (d) the information from the last survey only for both presence and absence units (PANUT-add).
2. Using the Ratio method (PANUT-ratio) using

- (a) the cumulative data of both presence and absence units from all previous surveys (PANUT-ratio-CD).
- (b) the cumulative information from all surveys dataset for the presence units but only the information from the last survey for the absence units (PANUT-ratio-pres).
- (c) the cumulative information from all surveys dataset for the absence units but only the information from the last survey for the presence units (PANUT-ratio-abs).
- (d) the information from the last survey only for both presence and absence units (PANUT-ratio).

The results are shown in *Figure 5.10*. First of all, the PANUT-add and PANUT-ratio methods were compared. PANUT-ratio had in general the highest detection rates. Interestingly, PANUT-ratio reached the highest detection rate in the second year, after using the information from the initial survey. This is similar to the non-zero PNUT methods which reached their highest detection rates after the third survey. The detection rates for PANUT-add-abs on the other hand still increased after the second survey. These result can be explained because the ratio method can become a greedy method. Recall that for the Ratio method $p_i^{\text{PANUT}} = \frac{p_i^{\text{PNUT}}}{p_i^{\text{ANUT}}}$. If p_i^{ANUT} is small, then p_i^{PANUT} can become very big. This will happen in high density areas. On the other hand if p_i^{PNUT} is small, then p_i^{ANUT} can solely become zero. Therefore with PNUT certain high density areas can absorb the majority of the sampling effort and will become greedy areas.

Using only the cumulative data of the absence units, thus PANUT-ratio-abs and PANUT-add-abs results in the highest detection rates. This unexpected result can be explained as follows: Since the aim is to increase the detection rate, over time more and more presence units will be found. The amount of presence units found in each survey will be high enough to construct a reliable PNUT layer. However, when combining the cumulative data of presence points, thus using PANUT-CD or PANUT-pres, inclusion probabilities will become very high, even for those units that are in the transition area between low and high density areas. This is problematic, since these areas often contain a relatively high number of absence units. Therefore, over time the detection rate will drop since many absence units in these intermediate density transition areas

will be sampled. This is also the reason why PANUT-CD and PANUT-pres resulted in almost identical detection rates since the number of available presence units will become too large and give misleading results. On the other hand, if the cumulative absence points are used, then these transition areas will get lower inclusion probabilities and the sampling allocation will be higher in the high density areas since enough presence units are available to map the true high density areas.

Finally the oscillating character of the completely non-cumulative methods can likely be explained by the fact that if in one survey a lot of presence units are found the next survey will likely focus too much on that area and thus a lot of absence units will be found. This in turn will cause the inclusion probabilities in those areas to drop, hence not enough samples will be selected in those areas etc. Over time this oscillating character seems to diminish since more information will become available.

CNUT

In this section, we evaluate adaptive monitoring by comparing

1. CNUT using the cumulative data from all previous surveys (CNUT-CD).
2. CNUT using the data from the previous survey only (CNUT).

The results are shown in *Figure 5.12*. For both the Rockfish and the Asian tiger mosquito population, the CNUT-CD resulted in higher detection rates compared with CNUT. For the clustered Rockfish population CNUT almost reached the highest detection rate in the second survey. After this initial jump the detection rates remained more or less constant. Note that for the Rockfish population there was an overlap in the error bars for each year indicating that there is no significant difference between the methods. For the Asian tiger mosquito there was also a big increase in the detection rate in the second survey. However, the detection rate kept increasing with each consecutive survey. This difference can be explained by the fact that exact mapping the density of a scattered population needs more information compared with the clustered populations.

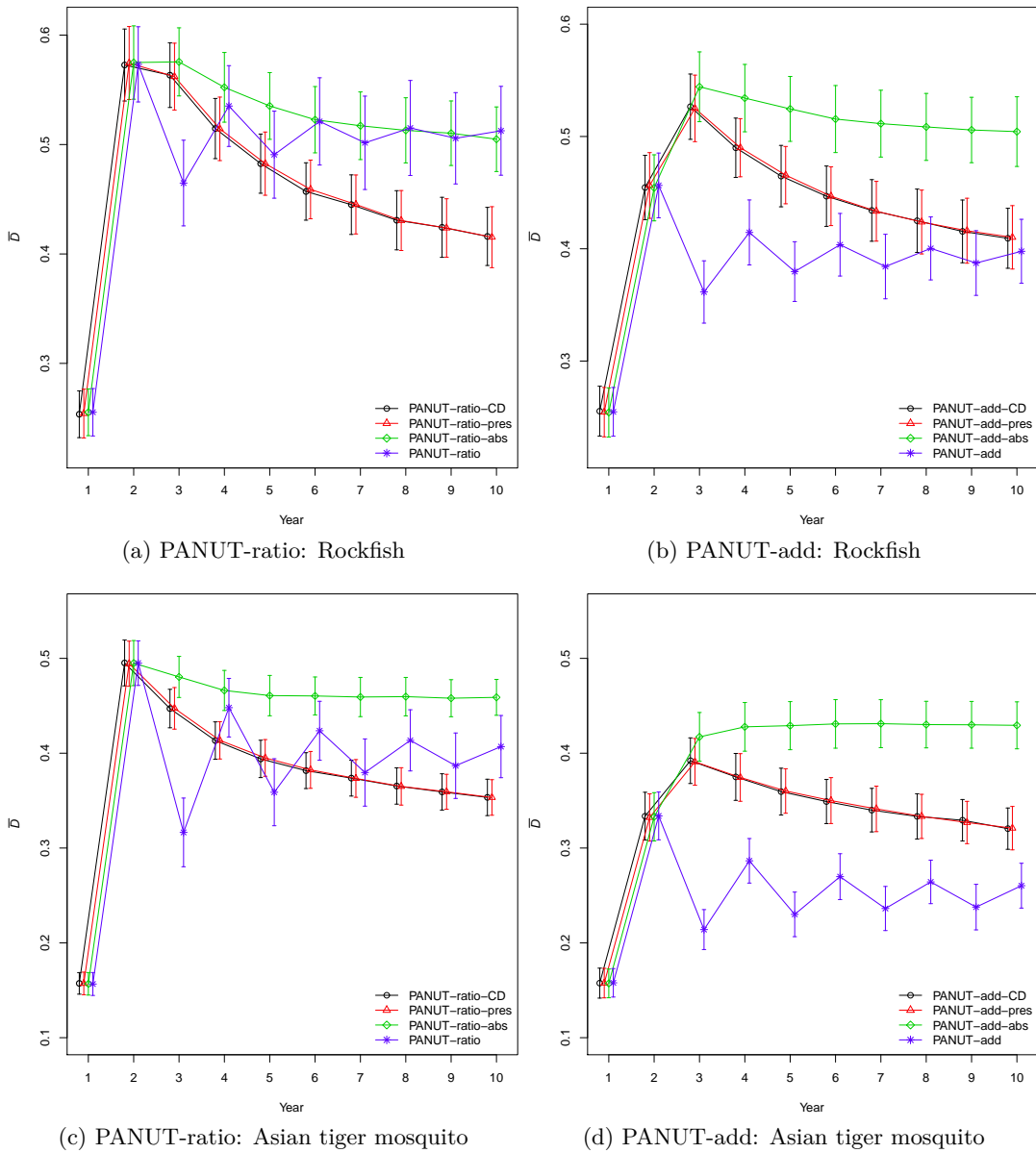


Figure 5.11: Results of the simulation study for PANUT. The error bars are the 2.5% and the 97.5% percnetile obtained by the simulation studies.

Summary of all NUT Methods

In this section, we select and compare some of the previous NUT methods based on the previous results and practical aspects. From each NUT type one method was chosen. Furthermore, in

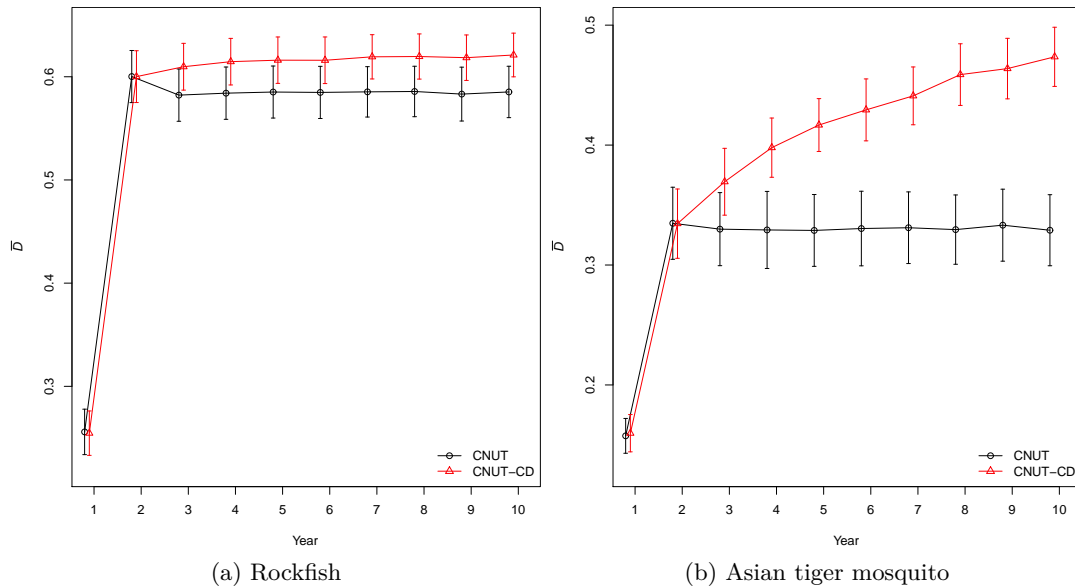


Figure 5.12: Results of the simulation study using CNUT. The error bars are the 2.5% and the 97.5% percentile obtained by the simulation studies.

general the method that would give the highest detection rates was chosen unless if that method would be too greedy as discussed in the previous sections, such as for example PANUT-ratio). For the Rockfish population the following NUT methods were selected:

1. SNUT-CD.
2. PNUT-0.5.
3. PANUT-add-abs.
4. CNUT-CD.

For the Asian tiger mosquito population the following NUT methods were selected:

1. UNUT-CD.
2. PNUT-0.5.
3. PANUT-add-abs.
4. CNUT-CD.

The results are shown in *Figure 5.13*. For the Rockfish population, CNUT had the best results. It had the highest initial jump in the detection rate after the initial survey. After this initial jump the detection rates remained more or less constant over time. CNUT had the highest detection rates for each consecutive year. Although both SNUT and PNUT reached their highest detection rates in year three, after the third year these detection rates gradually decreased. This is, as aforementioned, likely because of problems with units in the transition areas between the low and the high density areas.

For the Asian tiger mosquito dataset the best NUT methodology was less clear. PNUT is the first method that reached its maximum detection rate but the detection rates decreased after that. Although the detection rate of CNUT increased year after year, it only reached the detection rates of PNUT and PANUT after year five. However, from year five on, it had the highest detection rate. Therefore, it seems that CNUT for scattered data is efficient if a large amount of cumulative data are available to apply the CNUT method. PANUT had never detection rates as high as PNUT and CNUT, but did have high and stable detection rates than those with to the other methods. This illustrates that PANUT is a reliable method with high detection rates.

For both the Rockfish and the Asian tiger mosquito data the SNUT and UNUT methods resulted in the lowest detection rates. This can be explained because SNUT and UNUT can only assign two levels of inclusion probability: a relatively high inclusion probability or a relatively low inclusion probability. This is similar to Guisan's StratRS. Therefore, SNUT and UNUT cannot define any difference between for example low, intermediate or high density areas. Sampling these intermediate density areas with the same intensity as high density areas will result in lower detection rates.

5.4 Discussion

The NUT methods are a set of geographical techniques that use the observed spatial distribution of a species during a survey to update the inclusion probabilities for adaptive monitoring. NUT was developed since the existing geographical methods were not suitable for our proposed method

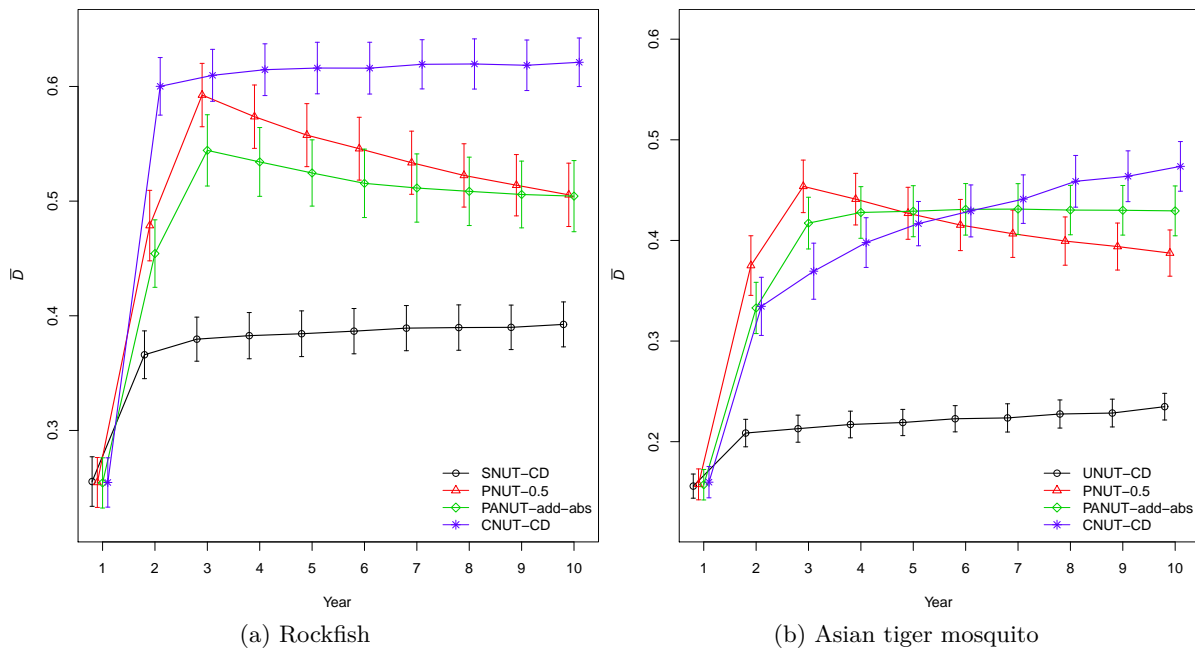


Figure 5.13: Results of the simulation study for comparing all NUT methods. The error bars are the 2.5% and the 97.5% percentile obtained by the simulation studies.

of adaptive monitoring. The NUT methods are intuitive and can easily be applied in practice. Using a simulation study, we illustrated that the NUT methods can lead to high detection rates. Of course these results are relative since no comparison is possible with the existing methods. We also illustrated that the detection rate of a monitoring strategy can be increased for both clustered populations and scattered populations. In the case of the previously performed simulation study, CNUT performed best for clustered data when count data are available. In the case of presence/absence data, PNUT or PANUT performed best. For the scattered Asian tiger mosquito population PNUT-0.5 is in our opinion the most reliable method. These conclusion of course depend on what proportion of the sample one wishes to assign as part of the surveillance sampling and on the species specific dynamics of the invasive population.

Although this chapter provides a good first introduction of the NUT methods, further development would be recommended. More extensive analyses should provide us with a better understanding of the use of each NUT method. Additionally the application of NUT to several other case studies would illustrate the robustness of the NUT methods better.

Even though future research is recommended to optimize the NUT methodology, in this chapter we have illustrated that NUT is a useful geographical method especially in the case of adaptive monitoring. Therefore, we will review the use of adaptive monitoring using NUT for non-stationary species populations in the next chapter.

Chapter 6

Adaptive Monitoring using Ecological and Spatial Information: Eradication of the Great White Butterfly

There are two types of auxiliary information that can be used to update the inclusion probabilities for the proposed method for adaptive monitoring. In Chapter 4 ecological information was used to pre-set the inclusion probabilities. This was done by computing the species' habitat suitability using SDMs. In Chapter 5 the observed spatial information of the invasive species distribution was used to improve the detection rate of the adaptive monitoring strategy. To set the inclusion probabilities using spatial information the NUT methods were introduced. In these chapters it has been illustrated that both ecological and spatial information contain valuable auxiliary data that can be used to increase the detection rate of the monitoring strategy. Hence, a logical next step is to use both ecological and spatial information simultaneously, to update the inclusion probabilities for adaptive monitoring.

In this chapter, we illustrate how to combine both types of information for adaptive monitoring. The methodology for using both types of information to set the inclusion probabilities for adaptive

monitoring is illustrated, and its performance is evaluated using a case study. For the case study we chose the current invasion of the Great White Butterfly (GWB), *Pieris brassicae*, of the Nelson District in the South Island of New Zealand (Kean and Phillips, 2013a,b; Phillips et al., 2013).

The GWB is a butterfly species that is non-native to New Zealand. The first observations of the presence of the GWB in Nelson were reported in 2010. The population has been notably increasing ever since. Similar invasions of the GWB in other parts of the world has lead to catastrophic effects on the local butterfly populations in the effected areas. In New Zealand, the native cress species are potential hosts for the GWB, and are therefore are considered threatened. Of the 79 cress species that are native to New Zealand, 57 are already at risk of extinction due to habitat loss and impacts from other pests (Phillips et al., 2013). A continuation of the current rate of dispersal and population growth of the GWB could also cause a large economic cost as the species also feeds on commercially grown Brassicae species. There are approximately 4,000 ha of Brassicae vegetable plantations in New Zealand and 6250,000 ha of brassica forage crops, which are estimated to be worth NZ\$ 80 million New Zealand dollars (Phillips et al., 2013). Because of these economical implications, a local eradication strategy has already been implemented in the areas where the species has been detected and in areas where it is assumed to be established (Kean and Phillips, 2013a). These local eradication areas are mainly in and near Nelson city. The decision to eradicate the GWB by the New Zealand Department of Conservation was made to avoid a further spread of the GWB to a more suitable habitat, identified in the Marlborough and Tasman districts which surround the Nelson district area. Invasion of these areas would increase the economic cost and would ultimately make future eradication unrealistic. Even though these more suitable areas are relatively far from where the species is currently established, the possibility of long distance dispersal through human aided transportation puts these areas at high risk. Human aided transport of species like the GWB is not uncommon since the pupae are known for pupating on artificial substrates in absence of their preferred host plants (Hagstrum and Subramanyam, 2010). Hence, effective and immediate eradication of the GWB is necessary.

Here, the GWB case study is used to evaluate the performance of adaptive monitoring using both types of information and to compare it with adaptive monitoring using only ecological information or using only spatial information. Initially, these comparisons will be performed by assuming a stationary population that does not change over time. This is similar to the simulation studies conducted in the previous chapter. However, the GWB is an invasive species, and one of the objectives of this thesis is to improve the detection rate of a monitoring strategy, so that it can be used for eradication management of invasive species. Thus, we explore the idea to use our proposed adaptive monitoring method to potentially link it to an eradication strategy. To evaluate the performance of our proposed method for adaptive monitoring as an eradication strategy, dynamic elements are added to the GWB population during the simulation study. This means that the GWB population that is used during the simulation study can grow, decrease, disperse and reproduce over time.

6.1 Methodology Adaptive Monitoring using Ecological and Spatial Information

Adaptive monitoring using ecological and spatial information is a combination of adaptive monitoring using ecological information (as introduced in Chapter 4) and adaptive monitoring using spatial information (as introduced in Chapter 5). Adaptive monitoring using ecological and spatial information simultaneously works as follows:

1. Select an equiprobable spatially balanced sample in the first survey using BAS.
2. Model the (cumulatively) obtained ecological and spatial information to set the combined inclusion probabilities. This is done by:
 - (a) Model the ecological information to compute the inclusion probabilities π_i^{eco} s:
 - i. Set the inclusion probabilities π_i^{eco} s proportional to the estimated species habitat suitability. This is identical to the method described in Chapter 4. However, no rescaling method is required at this stage.

- ii. Rescale π_i^{eco} such that it has a maximum of one and a minimum of zero. Thus compute

$$p_i^{\text{eco}} = \frac{\pi_i^{\text{eco}} - \min(\pi^{\text{eco}})}{\max(\pi^{\text{eco}}) - \min(\pi^{\text{eco}})}. \quad (6.1)$$

- iii. Set a lower limit l between zero and one such that $p_i^{\text{eco}} = l$ when $p_i^{\text{eco}} < l$. This lower limit assures that each unit has at least a small chance of being selected. Here we suggest $l = 0.05$. This value is chosen following the simulation studies performed in Chapter 4 which showed that the minimum suitability approach with $l = 0.05$ generally performed the best. Note that choosing a smaller value for the lower limit can increase the detection rate but comes at a cost by reducing the surveillance sample. Because of this trade-off, deciding on an optimal lower limit is not trivial and is often monitoring case specific.

- (b) Model the spatial information to compute the inclusion probabilities π_i^{spat} s:

- i. Select a NUT method to model the obtained spatial information, see Chapter 5. Calculate π_i^{spat} using the chosen NUT method.
- ii. Rescale π_i^{spat} such that it has a maximum of one and a minimum of zero. Thus compute

$$p_i^{\text{spat}} = \frac{\pi_i^{\text{spat}} - \min(\pi^{\text{spat}})}{\max(\pi^{\text{spat}}) - \min(\pi^{\text{spat}})}. \quad (6.2)$$

- iii. Set the same lower limit l as in step 2.a such that $p_i^{\text{spat}} = l$ when $p_i^{\text{spat}} < l$.

- (c) Combine p_i^{eco} and p_i^{spat} to set the final inclusion probabilities π_i^{both} for each unit by

$$\pi_i^{\text{both}} = \frac{n}{2} \frac{(p_i^{\text{eco}} + p_i^{\text{spat}})}{\sum_{i=1}^N \frac{(p_i^{\text{eco}} + p_i^{\text{spat}})}{2}}. \quad (6.3)$$

3. Select a probability sample using BAS according to the set π_i^{both} s for the next survey and return back to Step 2 .

An illustration of combining the inclusion probabilities based on ecological information and the inclusion probabilities based on spatial information is given in *Figure 6.1*. This figure shows how the set inclusion probabilities based on ecological information, π_i^{eco} (see *Figure 6.1a*), can be

different from the set inclusion probabilities using spatial information, (π_i^{spat} see *Figure 6.1c*).

Recall that the NUT methods perform best when the selected samples in each survey are spatially balanced samples. Therefore, when combining ecological and spatial data as described above, a spatially balanced sample should be selected in each survey. We suggest using a BAS for selecting these spatially balanced samples.

When a unit has a high species habitat suitability, it should follow that a high value for π_i^{eco} will be given, with a high observed species density based on the corresponding NUT method. Moreover, a high π_i^{spat} should also be observed, thus giving a high combined inclusion probability which is given by π_i^{both} . If a unit's π_i^{eco} and π_i^{spat} is low, then π_i^{both} for that unit will be low. If π_i^{eco} is high but π_i^{spat} is low, or vice versa, then the final inclusion probability π_i^{both} will have an intermediate value. Hence, using both types of information can be expected to lead to inclusion probabilities that are more closely correlated to the actual species density because it uses two types of information instead of one. We assume that these pre-set inclusion probabilities π_i^{both} will be more close to the ideal inclusion probabilities compared with π_i^{eco} or π_i^{spat} . As discussed in Chapters 4 and 5, setting the inclusion probabilities closer to the ideal inclusion probabilities should increase the detection rate of a sample. Hence we expect the detection rate of the monitoring strategy to increase when combining both types of information.

6.2 The Great White Butterfly Data

Like the Asian tiger mosquito data, which have been used in the previous chapters, the available GWB data set is a data set of a virtual invasion by the species. However, compared with the Asian tiger mosquito, the invasion of the GWB is not just a hypothetical invasion, in fact the invasion has actually already taken place. Therefore, the initial starting point of the invasion is known and data about the established invasion have been collected. For example, more than 9000 data entry points have been collected from the greater Nelson district area by AgResearch,

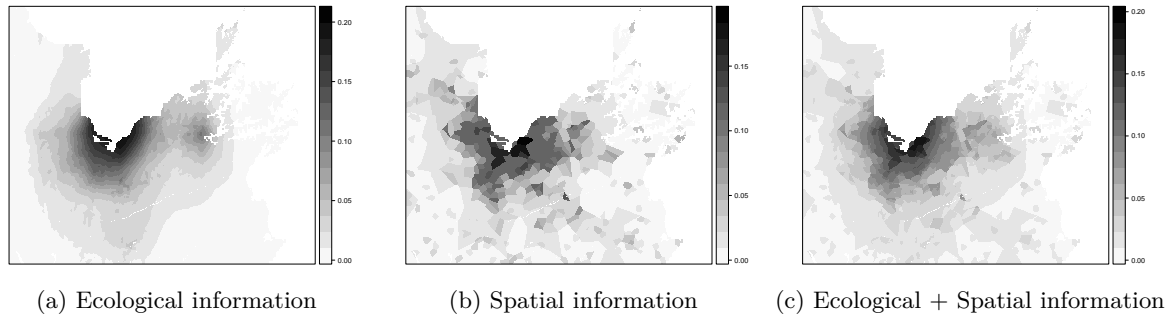


Figure 6.1: Illustration of inclusion probabilities set during adaptive monitoring using ecological and spatial information. Figure 6.1.a shows the inclusion probabilities, π_i^{eco} s, set using the ecological information. Figure 6.1.b shows the inclusion probabilities, π_i^{spat} s, set using the spatial information. Figure 6.1.c shows how π_i^{eco} and π_i^{spat} s are combined into the inclusion probabilities π_i^{both} using both types of information.

one of New Zealand’s Crown Research Institutes. This data set was used for cross validating the distribution of the virtual invasion. The virtual data set was created as part of a GWB project at AgResearch, Lincoln, New Zealand and the bio-protection unit of Lincoln University, New Zealand. We are very grateful for the assistance of Professor Sue Warner and Senait Senay from Lincoln University, and Dr John Kean and Dr Craig Phillips from AgResearch for providing access to this data.

To generate the data of the virtual invasion a dispersal model was used to simulate and predict future distribution and dispersion of the GWB over the greater Nelson district area. Since the invasion of the GWB butterfly is likely to escape the Nelson city area to the greater Nelson district area, an arbitrary area around Nelson city was selected for this case study. This area is shown in *Figure 6.2*. The dispersal model allowed for simulated dispersal over a heterogeneous habitat suitability layer. This means that each pixel was given a habitat suitability with a value between zero and one. This habitat suitability was then used as a survival probability for the GWB if it would invade that unit. This habitat suitability layer was obtained by species distribution modelling using 17 ecological covariate data layers. These covariates were geological, climatological or ecological covariates such as altitude, rainfall or type of habitat that could be used to predict the species habitat suitability. Details of these ecological data layers can be

found in Appendix A. The locations of the initial dispersers in year one were chosen according to reports of observations of the GWB in the year 2010. Species specific parameters such as birth and mortality rates were used to specify the rate of local spread of the species using population growth and spread rates were retrieved from the literature. Both short distance and long distance dispersal were included in the dispersal model.

Details on the species distribution modelling and dispersal modelling used to generate the distribution maps can be found in Senay (2014a) and Senay (2014b). The simulation studies of the invasion of the GWB started in 2010 and the invasion by the GWB was predicted until the year 2020. *Figure 6.2* illustrates one example of a simulated invasion of the GWB, starting with the dispersal seeds in the year 2010 until the year 2018.

The local eradication strategy that already has been implemented by the Department of Conservation mainly focuses on Nelson city. The study area chosen for this simulation study includes a larger section of the Nelson district. For surveying purposes the selected area around Nelson was tessellated into square units of 100 by 100 meters. The dimensions of the sampling domain are 207 rows by 278 columns, in total 57546 units. Excluding the units outside the sampling domain (for example in the ocean) there are $N = 48303$ units to be sample from.

6.3 Simulation study: Stationary GWB population

A simulation study was performed using the GWB data set to evaluated whether the detection rate can be increased by combining both types of information compared with using only one type of information. Adaptive monitoring was performed for seven consecutive years. Each year 5% of the units are sampled. Monte Carlo simulations were performed with $m = 500$ (the number of simulations) using the following methods:

1. Adaptive monitoring using ecological information only (AMBAS, as illustrated in chapter 4): Similar to Chapter 4, we will use logistic regression as the SDM of choice. As a rescaling



Figure 6.2: Illustration of the study area surrounding Nelson, South Island, New Zealand, selected for monitoring/eradication of the GWB in the performed simulation studies. The location of Nelson city is illustrated by the dark grey point and lies central in the selected study area.

method the minimum suitability approach with $l = 0.05$ (LIM05) was used.

2. Adaptive monitoring using spatial information only (PANUT, as illustrated in Chapter 5): Since the GWB data are scattered and has only presence/absence data PANUT were selected as our NUT method of choice (using the cumulative data for both presence and absence data).
3. Adaptive monitoring using both ecological and spatial information: To be able to compare this method with the previous two methods logistic regression was used the set p_i^{eco} and PANUT was used to set p_i^{spat} . A lower limit of $l = 0.05$ was used.

For each method and for each survey the following statistics were calculated: The detection rate at survey k is given by

$$\bar{D}_k = \frac{1}{m} \sum_{j=1}^m D_{j,k}, \quad (6.4)$$

as well as the simulated variance of the calculated detection rates per year $\hat{V}_{\text{SIM}}(D_k)$ is

$$\hat{V}_{\text{SIM}}(D_k) = \frac{1}{m} \sum_{j=1}^m (D_{j,k} - \bar{D}_k)^2. \quad (6.5)$$

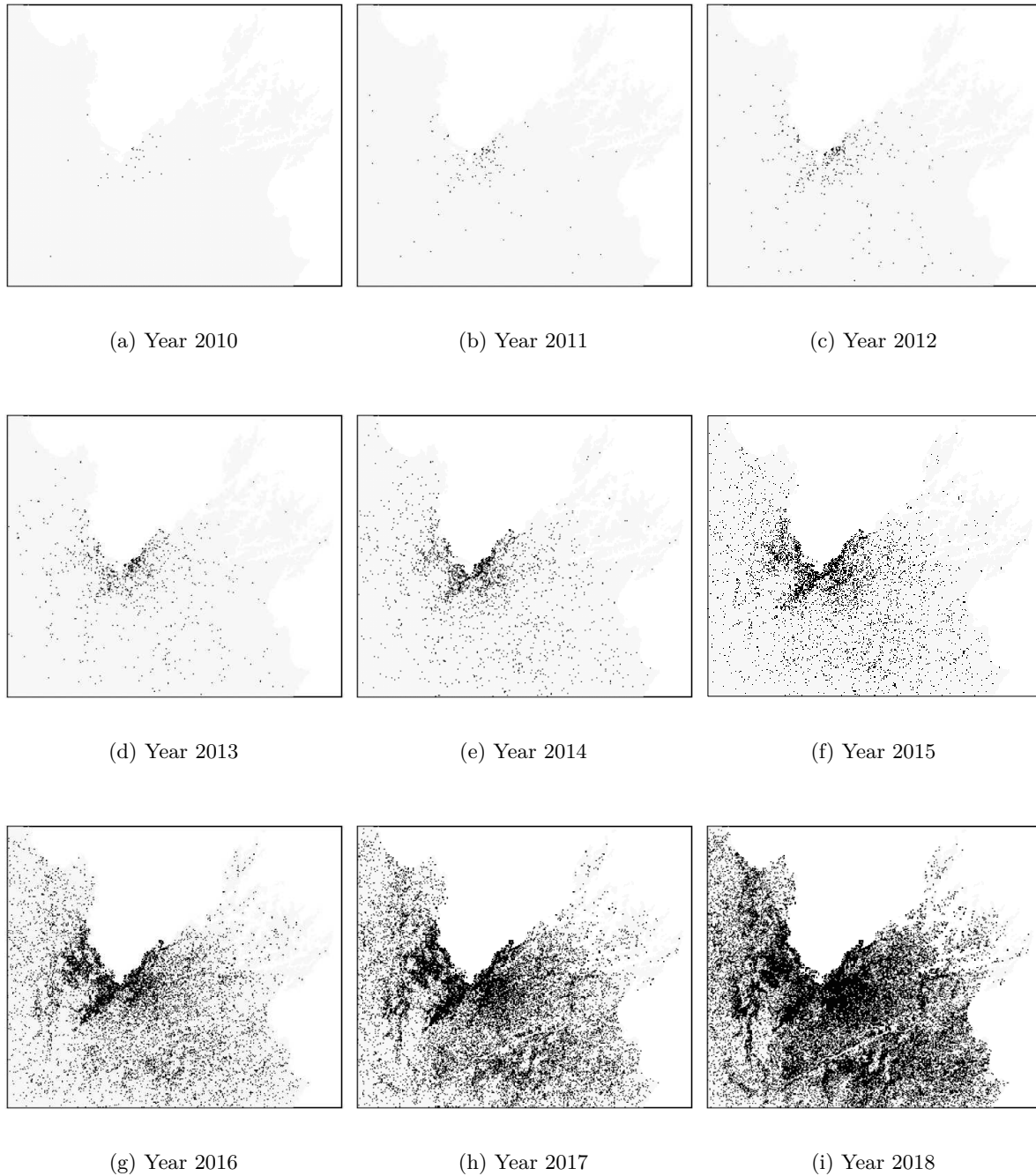


Figure 6.3: Illustration of a potential invasion by the GWB, starting in 2010 until 2019. Note that the occupied units in the year 2010 (the initial seeds of invasion) are based on the observed occurrences of the GWB in the year 2010

Similar to the simulation studies performed in Chapter 4, we evaluate adaptive monitoring using different types of auxiliary information for three different GWB populations:

1. **Low density stationary population:** The low density scenario chosen for this simulation study is similar to the year 2014 GWB population illustrated in *Figure 6.3*. In this low density scenario the population is stationary which means that the population does not change over time.
2. **High density stationary population:** The high density scenario chosen for this simulation study is similar to the year 2018 GWB population illustrated in *Figure 6.3*.
3. **Increasing density population:** Where, over time the GWB population increases. For this simulation study the species distributions starting in the year 2014 until the year 2020, as illustrated in *Figure 6.3* are used.

6.3.1 Results

The results of the simulation study for each scenario are shown in *Figure 6.4*. Since an equiprobable sample was selected in year one, all methods achieved similar detection rates. After the initial survey, adaptive monitoring using ecological and spatial information always gave higher mean detection rates compared with adaptive monitoring using only spatial or only ecological information. This was the case for all years after the initial survey and for all three scenarios. Adaptive monitoring using spatial information only, generally achieved higher detection rates compared with adaptive monitoring using ecological information with the exception of year four to year seven in the high density scenario. This confirms that, in the case when a species is not in equilibrium with its habitat, which is often the case for invasive species, spatial information can often be as important or even more important than ecological information to predict the species distribution. The GWB population in the second scenario is more in equilibrium with the environment (invaded more sites that are suitable for the species) compared with the first scenario, this explains why the advantage of using spatial information over ecological was less obvious in the second scenario.

In the third scenario the differences between the three methods for adaptive monitoring were

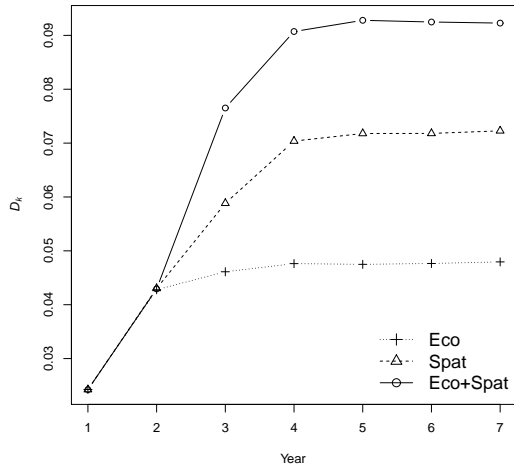
smaller compared with the first two scenarios. This may be able to be explained as the population increases in size over time. Because the GWB invades new units in addition to the units already occupied, it is difficult to predict which units will be occupied in the next year. Therefore, using a combination of ecological and spatial information performed best and resulted in the highest mean detection rates. It is interesting to notice that in this third dynamic scenario the highest detection rates are reached by 2018 and that these detection rates are higher than the highest observed detection rate for scenario 2 which was stationary at the level of 2018. This can be explained by the fact that since the population expands over time for certain hotspots (initial seeds) the model will focus the sampling allocation more on the already invaded areas and areas near these hotspots (especially since the model incorporates spatial information). In the second scenario the model has a relatively larger surveillance samples since there are more areas invaded and hence more areas that are close to these invaded areas. It is therefore not unlikely to expect that if the simulation study would incorporate more surveys over time that then the detection rates for both scenarios would become similar.

In the first scenario adaptive monitoring using ecological and spatial information had the lowest observed variances compared with adaptive monitoring using only one type of information. Adaptive monitoring using ecological information had a larger observed variance in the detection rate compared with adaptive monitoring using spatial information. For the second and third scenarios the trends in the observed variances were less obvious. However, in general adaptive monitoring that uses both types of information performed best and had the smallest observed variances in the detection rates.

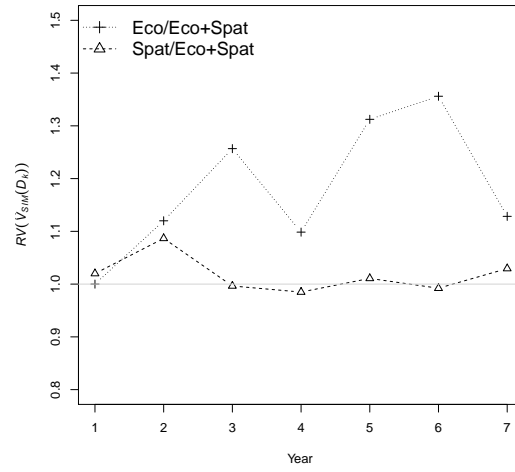
The results of the simulation study confirm our hypothesis that both types of auxiliary information are important. The use of a combination of both types of information allow for more accurate prediction of the distribution of the invasive species and can therefore lead to higher detection rates.

Note that, as discussed in Chapter 4, the level of the lower limit l defines the size of the surveillance sample. In this chapter we have used a lower limit of $l = 0.05$. A higher lower limit will increase the size of the surveillance sample, and hence, the monitoring method will be

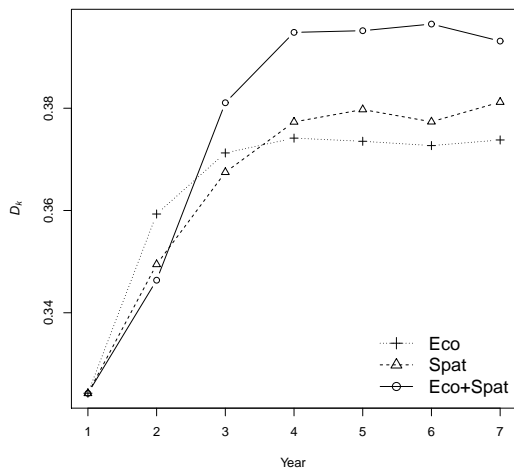
more sensitive to new or undetected hotspots of the invasive species. However, in that case, the proportion of the sampling effort that can be used in assumed high density areas will be reduced. Whereas, a smaller lower limit will reduced the size of the surveillance sample. A smaller surveillance sample could mean that new hotspots of the invasive species remain undetected. This could have many negative consequences for the control of the invasive species. If a researcher is confident that the spatial information (and thus NUT) or the ecological information (and thus the AMBAB), or a combination of both, will provide enough information to accurately predict the species distribution, then a small lower limit could be considered. On the other hand, if the dispersal of a species is unpredictable and cannot be captured well by spatial or ecological information, a higher lower limit should be considered.



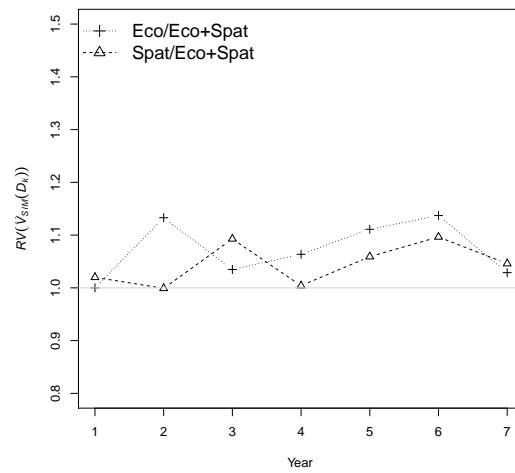
(a) Low GWB density: \bar{D}



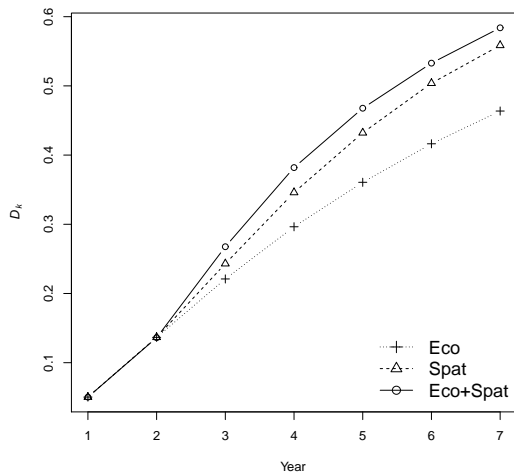
(b) Low GWB density: $RV(\hat{V}_{SIM}(\bar{D}))$



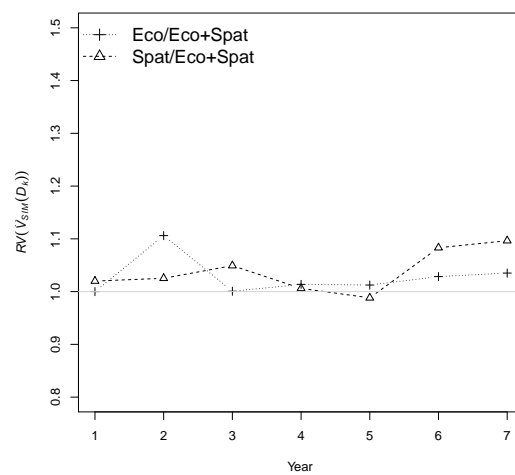
(c) High GWB density: \bar{D}



(d) High GWB density: $RV(\hat{V}_{SIM}(\bar{D}))$



(e) Increasing GWB density: \bar{D}



(f) Increasing GWB density: $RV(\hat{V}_{SIM}(\bar{D}))$

Figure 6.4: Results of the simulation study on stationary GWB Populations

6.4 Adaptive Monitoring to set up an Eradication Strategy

Achieving high detection rates while monitoring a species can be useful to improve the efficiency of an eradication strategy. One aspect of having a more efficient eradication strategy is that a higher proportion of the occupied units are being selected and thus more of the invasive species can be eradicated. The proportion of units occupied by the species in survey k is defined by

$$E_k = \frac{N'_k}{N}, \quad (6.6)$$

where N'_k is the number of units occupied by the species at time k . We call an eradication strategy positive when E_k decreases over time. In that case, the population of the invasive species will decrease, and the eradication strategy achieves its goal of reducing the invasive species population. An eradication strategy is called negative when E_k increases over time. In that case the eradication strategy has failed and the species population continues to expand. A neutral eradication strategy has on average a constant E_k over time.

The method that we suggest is to use adaptive monitoring as to direct an eradication strategy to reduce E_k over time. In the subsequent simulations it will, for simplicity, be assumed that each time a unit is selected all the individuals of the invasive species in that unit are eradicated. This is in practise of course rarely a realistic assumption and hence further research is necessary to test the effect for when eradication is not 100 % succesful. However, to be able to set the inclusion probabilities based on ecological and spatial information for the consecutive surveys, an 'eradicated' unit remains registered as a presence unit. The reason for this is as follows. Suppose that only spatial information is used to set the inclusion probabilities. If a sampled unit is occupied by a species the species in that unit should be eradicated. However, the chances are relatively high that the units surrounding the sampled unit are occupied. Hence, by modelling the sampled unit as a presence unit, the inclusion probabilities of surrounding units will increase. If the sampled unit would be modelled as an absence unit then the surrounding units will be assigned smaller inclusion probabilities. In the case that ecological information is used then changing the value of all eradicated units from a zero to one would make the results of SDMs

unreliable and biased, since the species habitat suitability from a certain unit would change from high to low even though the actual habitat remains the same.

6.5 Simulation study: Non-stationary GWB Populations

In the previous simulation study, as well as in previous chapters, the species population was assumed to be stationary over time or non-stationary without any form of eradication implemented. This was done because it allows for straightforward comparison of the suggested methods and for simple interpretation of the results. Nevertheless, making this assumption is highly unrealistic. In reality, the population of an invasive species is highly non-stationary. The population size of the invasive species can increase or decrease, and the distribution of the population can change often. Moreover, eradication strategies can have an important impact on the species population. Therefore, a simulation study was performed to evaluate our suggested method for adaptive monitoring and to illustrate its use as an eradication strategy.

In this simulation study we apply our method for adaptive monitoring together with the eradication method as discussed above on a non-stationary GWB population. This simulation is performed as follow: Begin by choosing one initial GWB population distribution. Next, an equiprobable sample is selected using BAS for the initial survey. To implement the eradication strategy, we update our initial GWB population distribution by eradicating all the species in the sampled units. Because of this the number of units occupied by the GWB will usually decrease. Next, a dispersal model is applied to the updated GWB population distribution. The dispersal model we used here is the same dispersal model as was used to generate the virtual dispersal maps as for example illustrated in *Figure 6.3*. Complete details about this dispersal model can be found in (Senay, 2014b,a). After this dispersal step, generally the number of occupied units by the GWB would increase again. Based on this eradication step and dispersal step the species distribution for the next year is set.

To select a sample in the next survey, the inclusion probabilities are set based on the collected information in the initial/previous survey. This is based on the cumulatively collected ecological and spatial information as previously discussed. A sample is selected using BAS based on the

set inclusion probabilities. Then follows a second eradication phase and a second dispersal phase. These steps are repeated for seven consecutive surveys.

Each year 5 percent of the units are sampled. The simulation study was repeated for two different initial GWB population distributions. The first GWB distribution is the low density distribution predicted for the year 2014 as shown in *Figure 6.3*. The second GWB distribution is the high density distribution predicted for the year 2018 as shown in *Figure 6.3*.

Since the population size changes, it would be infeasible to compare the mean detection rates per survey over all simulations for different methods. Furthermore, since the goal was to eradicate the species, we computed the eradication efficiency E as a test statistic to compare the different methods. The eradication efficiency at each survey k , \bar{E}_k is given by

$$\bar{E}_k = \frac{1}{m} \sum_{j=1}^m E_{j,k}, \quad (6.7)$$

as well as the simulated variance of the calculated eradication efficiency per year $\hat{V}_{\text{SIM}}(E_k)$ and is given by

$$\hat{V}_{\text{SIM}}(E_k) = \frac{1}{m} \sum_{j=1}^m (E_{j,k} - \bar{E}_k)^2. \quad (6.8)$$

6.5.1 Results

Figure 6.5 shows the results of the second simulation study for both the low density initial GWB distributions and the high density initial GWB distribution. Both the low and the high density initial distribution showed a similar pattern in \bar{E}_k over time. After the initial survey the number of occupied units increased and thus more units were occupied by the GWB compared with the first year. This can be explained because an equal probability sample was selected in the initial survey. This means that the sample did not focus on any potential high density areas, since no auxiliary information was available to set unequal inclusion probabilities. Because of this, the detection rate was relatively low in the initial survey. The number of unit that were occupied after the first dispersal step was generally higher than the number of units that have been eradicated during the initial survey. Hence, the proportion of invaded units increased. From the second survey onwards auxiliary information was available to set the inclusion prob-

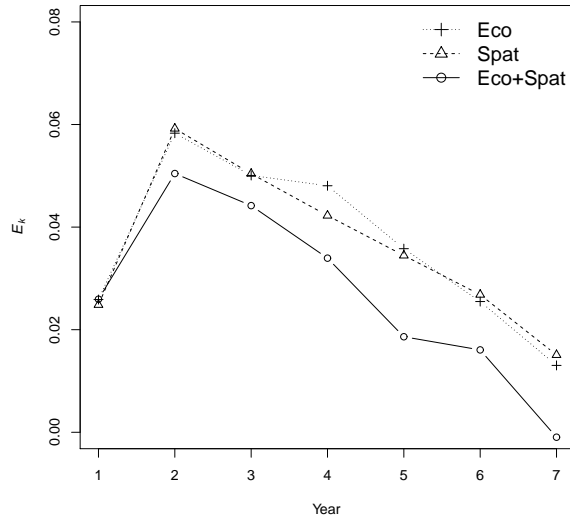
abilities. Because of this, it was possible to more accurately predict the species distribution and thus to focus the sampling effort on expected high species density areas. As a result, the detection rate from year two onwards increased and more units were eradicated. Since the number of occupied units decreased from survey two onwards it seems that more units were eradicated than the new units were invaded.

For the low initial GWB distribution the proportion of occupied units became on average lower than in year one from year 5 onwards. For the high initial GWB distribution this number of occupied units was lower than year one from year 7 on. This suggests that when an eradication strategy is implemented when the size of the invasive species population is small, the eradication strategy will more quickly have a positive effect on the control of the invasive species population.

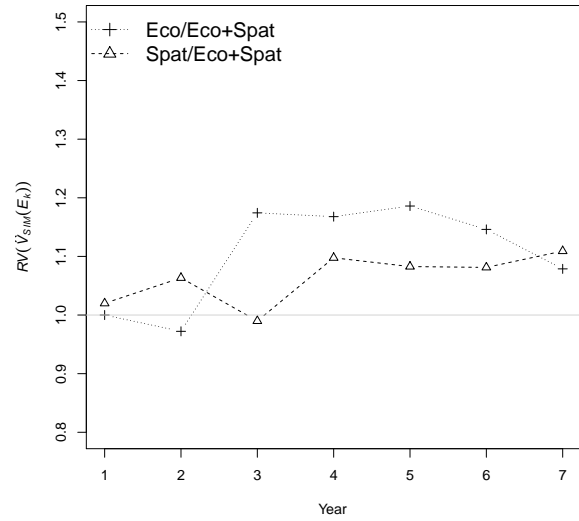
6.5.2 Alternative Eradication Method

In the previous simulation study, if a species was detected in a unit, then all individuals were eradicated in that unit. However, the method for adaptive monitoring can be used in an alternative way to define an eradication strategy. This method assumes that after each sample another batch of units are selected in which all the individuals are eradicated. This second set is selected based on the set inclusion probabilities. Let n' be the sample size of this second set of units. Then select the n' units with the highest assigned inclusion probabilities. Visit these n' units and eradicate the invasive species if present. Note that, the information gained from visiting these n' units cannot be used for setting the inclusion probabilities for the next survey(s). This is due to no straightforward method being available to incorporate the information with the proposed method for adaptive monitoring. As an example of this eradication method, assume that *Figure 6.1* illustrates the computed inclusion probabilities after the first survey. The selected n' units with the highest inclusion probabilities are shown in black in *Figure 6.6*. The n units in survey 2 were selected using an unequal probability BAS sample, that are illustrated in grey.

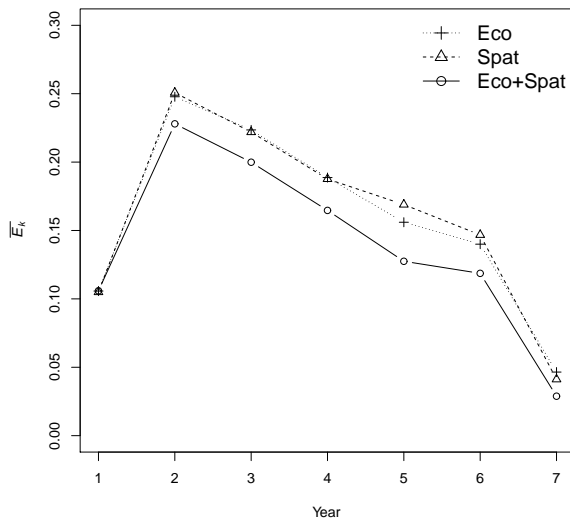
However, more research is needed on this type of eradication. Therefore, we will not go into further details or test the performance of the eradication method using a simulation study.



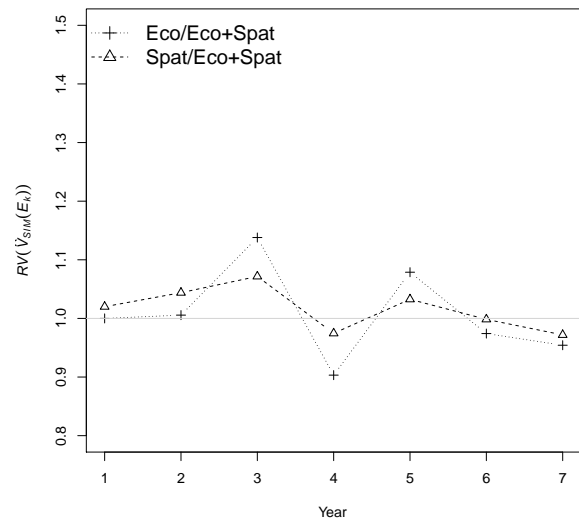
(a) Low initial GWB density \bar{E}



(b) Low initial GWB density $RV(\hat{V}_{SIM}(\bar{E}))$



(c) High initial GWB density \bar{E}



(d) High initial GWB density $RV(\hat{V}_{SIM}(\bar{E}))$

Figure 6.5: Results of the simulation study on non-stationary GWB Populations

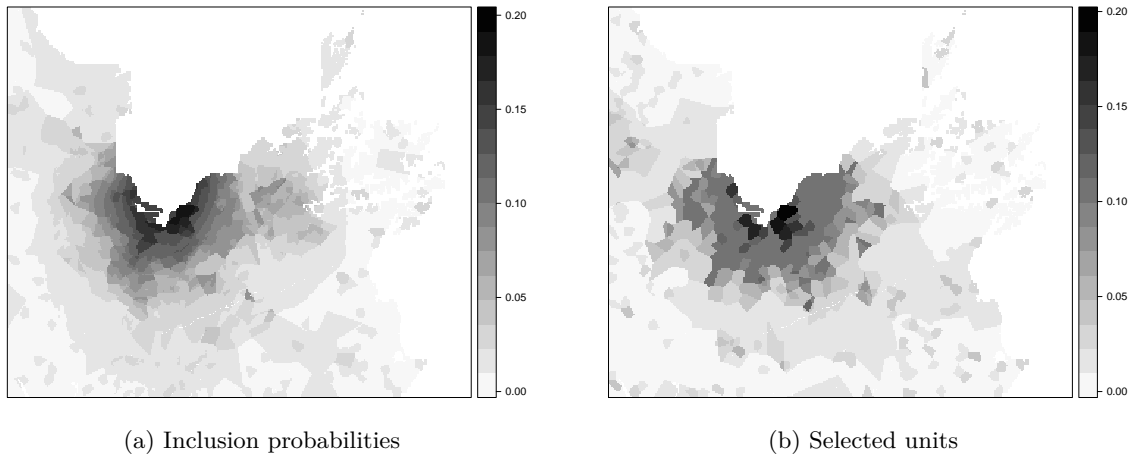


Figure 6.6: Illustration of alternative eradication method. In this case the n' units are selected based that have the highest inclusion probabilities. Figures *a* shows the set inclusion probabilities based on the ecological and spatial information gained from the first survey. Figure *b* illustrates the n' units with the highest inclusion probabilities (in black) and a sample of size n' selected by an unequal probability sample selected using BAS based on the set inclusion probabilities (in grey).

6.6 Discussion

In this chapter we have demonstrated how to combine ecological and spatial information to set the inclusion probabilities for adaptive monitoring. This was illustrated using the GWB invasion of the Nelson district. The results have shown that combining both ecological and spatial information improves the detection rate of an adaptive monitoring strategy compared to a strategy that makes use of only spatial information or ecological information alone at least in the case of the GWB example.

Moreover, we also gave other uses for our method, which saw it implemented as an eradication strategy. The efficiency of a potential eradication method has been illustrated using the GWB data. The results of the simulation study demonstrated that the proposed method performed most efficiently when the size of the population under scrutiny was small as opposed to pop-

ulations that are larger and are more dispersed. In addition, we noted that the use of both ecological and spatial information outperformed adaptive monitoring methods which used only one type of auxiliary information.

The results from this chapter have further reiterated the three key aspects set as our objectives in the introduction. These aspects were so that the existing methods for adaptive monitoring can be further improved by: (i) increasing the detection rate by the use of a spatially balanced probability sample, (ii) combining spatial and ecological information and (iii) incorporating an eradication strategy to an adaptive monitoring design.

Chapter 7

Discussion and Future Work

In this thesis, I have developed the ideas of for an iterative monitoring method, which was first proposed in Stohlgren and Schnase (2006) and was extended by others, to introduce a new and improved general framework for adaptive monitoring. In Chapter 1, we provided a motivation for the need for the development of adaptive monitoring designs and we outlined three potential points for improvement: (1) selection of the sampling design, (2) using spatial and ecological information to update the monitoring strategy and (3) exploring the idea of using the proposed idea of adjusting unit specific inclusion probabilities to set up an eradication strategy. In the chapters that followed these points were addressed in more detail. Note that, by selecting these three points, we identified what we believe to be the most important aspects that could be used to improve adaptive monitoring. However, it can be further noted that there exists other aspects that can also improve adaptive monitoring, and therefore should be explored, see *Section 1.1*. The algorithm for our proposed framework for adaptive monitoring has been introduced in Chapter 2. The general concept is that, after each survey the inclusion probabilities are adjusted using the collected information. This adjustment is done so that units with an expected higher species density are assigned a higher inclusion probability.

Additionally in this second chapter, a basic background to ecological sampling has been provided. The concepts of probability sampling and of spatially balanced sampling were discussed. Moreover, a motivation on why adaptive monitoring would benefit from probability sampling and spatially balanced sampling was given. This reinforced our motivation that the use of spatially

balanced probability sampling can be expected to increase the detection rate, whilst selecting a surveillance sample to ensure that newly invaded areas will be detected.

In the third chapter we introduced a selection of probability sampling designs. These sampling designs were: SRS, StratRS, and CP sampling. We also introduced several spatially balanced sampling methods. Since GRTS can be seen as the gold standard of spatially balanced sampling and BAS is a recently introduced alternative to GRTS, we mainly focused on BAS and GRTS. BAS is a new spatially balanced sampling design, based on a relatively simple algorithm (Robertson et al., 2013). The performance of the spatially balanced sampling designs for adaptive monitoring was evaluated by using a simulation study. We illustrated that BAS generally achieved a higher spatial balance compared with SRS and GRTS. This was the case when selecting a whole batch of units at the same time, as well as when adding units one by one to the sample. Moreover, BAS achieved a higher spatial balance with and without non-response areas compared with GRTS and SRS.

We evaluated each of the introduced sampling designs for selecting a one off sample, which led to selection of the initial sample in a monitoring strategy. The main criteria for a suitable sampling design is to achieve high detection rates. This was evaluated by performing a simulation study on the Rockfish population. The results showed that spatially balanced sampling design achieved similar detection rates compared with SRS, CP and a stratified sampling design but the observed variance in the detection rates was generally lower. StratRS performed good as well as long as the selected strata were small enough to increase the spatial balance of the stratified sample. Overall, BAS achieved the highest detection rates.

The second criteria was to achieve precise population estimates. We demonstrated this by conducting a simulation study which used two different virtual target populations. Results from the study showed that BAS generally performed at least as well as the other methods. However, StratRS has the important advantage of having an unbiased variance estimator compared with the other tested spatially balanced sampling designs. Additional simulation studies were also conducted to illustrate the importance of having a spatial autocorrelation in the response and setting the inclusion probabilities proportional with the response variable when selecting a

probability sample. These simulation studies illustrated that in case the inclusion probabilities are not set proportional with the response there is no advantageous in selecting a spatially balanced sample, and estimates can even become worse compared with SRS or StratRS. The problem of setting poor inclusion probabilities can become the case for example when not enough information is available to set the inclusion probabilities. This is sometimes the case before the initial sample.

Ultimately, in chapter 3, we explored some practical features of BAS, that allowed it to be used for long term monitoring. In addition, BAS is computationally cheaper than other spatially balanced sampling designs that were considered.

Since overall BAS scored best for both criteria and has several desirable practical features, we are of the opinion that BAS is the best sampling design to select the initial sample for adaptive monitoring. Based on these results, BAS was used in the following chapters to select the probability samples based on the pre-set inclusion probabilities.

Adaptive monitoring means that, for each additional survey, the monitoring design is updated based on the obtained information from the previous years. In the fourth chapter we illustrated how to use ecological auxiliary information to update the monitoring strategy. Based on the collected ecological information the species habitat suitability is estimated using a SDM. Next, the inclusion probabilities of all units in the study area were set proportional to the estimated species habitat suitability. A probability sample is then selected based on the set inclusion probabilities. This method was called AMBAS, as it uses BAS to select the samples in each survey. Previous studies such as Guisan et al. (2006) followed a similar method, however they used the estimated habitat suitability to partition the study area in a suitable stratum and unsuitable stratum. Based on the stratification of the study area, a sample was selected using StratRS. We named this method AMStrat.

We compared AMStratRS with AMBAS and SRS by conducting a simulation study using a virtual data set about a possible invasion of the Asian tiger mosquito in the South Island of New Zealand. In most of our simulation studies the detection rates for AMBAS were at least twice as high than those with SRS. When comparing AMBAS with AMStratRS the mean detection

rates were similar for AMBAS and AMStratRS, but AMBAS generally gave smaller variances in the observed detection rates.

To avoid problems with (near) zero inclusion probabilities, three different rescaling approaches were proposed, and simulation study was performed to compare the results of the approaches. The minimum suitability approach, with a lower limit of $l = 0.05$ (LIM05) gave the best results in our simulation studies. Setting a lower limit means that the expected low density area can still be sampled and a surveillance sample is ensured. This surveillance sampling is necessary especially when monitoring an invasive species to hopefully avoid new outbreaks. The main issue with selecting a lower limit is that the size of the surveillance sample depends from situation to situation. For example is only a small surveillance sample is required and one would like to increase the detection rate then a smaller lower limit would work best. However, in many cases a larger proportion of the sample should be used for surveillance sampling to ensure new hotspot of the invasive species can be found or the detect new outbreaks. In that case a larger lower limit should be used. This will then likely come at a cost of higher a lower detection rate. The results for population estimation using AMBAS were very poor. Because of these result it was decided not to use AMBAS or any similar method for adaptive monitoring for population estimation.

A second simulation study was conducted to compare the differences between using spatially balanced sampling designs and using non-spatially balanced sampling designs for adaptive monitoring. The spatially balanced sampling designs achieved higher mean detection rates and higher precision. Adaptive monitoring using BAS had marginally better results over adaptive monitoring using GRTS.

All the simulation studies were performed by using three different species density scenarios. The results of these simulation studies showed that AMBAS worked better for species populations that are in equilibrium with the habitat, compared with species populations that are not yet in equilibrium. Finally, in Chapter 4, an additional simulation study illustrated that AMBAS does not works as well in the case of count data, In which case AMStratRS performed better.

To be able to set the inclusion probabilities based on the observed spatial information we intro-

duced the NUT methods in Chapter 5. The NUT methods are a set of geographical techniques that uses the observed spatial distribution of a species during a survey to update the inclusion probabilities for adaptive monitoring. NUT was developed since the existing geographical methods were not suitable for our proposed method for adaptive monitoring. Furthermore, the NUT methods are intuitive and can easily be applied in practice.

Using a simulation study on a Rockfish population and also the Asian tiger mosquito population, we illustrated that the NUT methods can lead to high detection rates. We also illustrated that the detection rate of a monitoring strategy can be increased for both clustered and scattered populations. We recommend using CNUT for clustered data when count data is available. In the case of presence/absence data, we recommend using PNUT or PANUT. For the scattered Asian tiger mosquito population PNUT-0.5 was the most reliable method.

Although this chapter is a good introduction of the NUT methods, further development of the NUT methods are recommended. This includes additional extensive analyses that may provide us with a better understanding of the use of each NUT method. In addition, the application of NUT to several other case studies would illustrate the robustness of the methods. Nevertheless, we have illustrated that NUT can be a useful geographical method, especially when applied to adaptive monitoring.

In Chapter 6, it was demonstrated how to combine ecological and spatial information to set the inclusion probabilities for adaptive monitoring. This was illustrated using the GWB invasion of the Nelson district. The results have shown that combining both ecological and spatial information improves the detection rate of an adaptive monitoring strategy compared with a strategy that makes use of only spatial information or ecological information.

Moreover, we also gave other uses for our method, which saw it implemented as an eradication strategy. The efficiency of our eradication method has been illustrated using the GWB data. The results of the simulation study demonstrated that the proposed method performed most efficiently when the size of the population under scrutiny was small as opposed to populations that are larger and are more dispersed. In addition, we noted that the use of both ecological and spatial information outperformed adaptive monitoring methods when only one type of auxiliary

information is used.

The results from this final chapter have further reiterated the three key points set as our objectives in the introduction. These points were defined so that the existing methods for adaptive monitoring can be further improved by (i) increasing the detection rate by the use of a spatially balanced probability sample, (ii) combining spatial and ecological information and (iii) incorporating an eradication strategy to an adaptive monitoring design.

7.1 Future Work

We have illustrated that our proposed algorithm for adaptive monitoring performs well by increasing the detection rate in comparison with previously applied monitoring methods. Despite these improvements we acknowledge that we can improve our method further, and suggestions for future research are numerous .

Here, the most relevant future research questions and remarks are briefly outlined:

- In chapter one, we assumed that throughout this thesis we will have perfect detectability. This means that if a species is present in a unit and that unit is sampled, then the species is always observed. However, this is a simplification of most real life applications and many studies have discussed issues with imperfect detectability. Therefore further research is needed to understand what the effects of imperfect detectability are on our proposed method for adaptive monitoring.
- In this thesis, we selected an equal sample size in each survey. However, the efficiency of distributing the sample size equally over subsequent surveys has not been evaluated. A possible scenario is that this strategy may be suboptimal over other sampling allocation strategies over time. For example it maybe the case that, allocating the majority of the sampling effort in the first survey would give better results. This is a simple research question that will not likely have a simple answer.

- A useful extension of BAS is to use BAS in a stratified context, which there are several applications. For example, selecting a BAS sample in each strata or generating Halton points over the strata until a sufficient number of points in each strata has been selected. This research question is already work in progress. However, it was not straightforward to include these results into this thesis without compromising the logical structure of this thesis.
- The importance of a surveillance sample has been mentioned several times in this thesis, however we failed to provide specific details. This is because studying the importance of the size and distribution of a surveillance sample can be complex and would require a great number of additional simulation studies, which required more time than what we had.
- Further work on NUT needs to be undertaken to ensure it works well in different situations and for different applications. Although the result of NUT look to be very promising, more research is required to make it suitable for publication. For example, more research is needed on the effect of the used lower limit. Additionally, more research is needed on the cumulative or non-cumulative use of collected information.
- A potential extension of the proposed method for adaptive monitoring would be the inclusion of practical covariates. For example, adding a cost function to the monitoring strategy. Question such as, how would our monitoring strategy change when certain units would be twice as expensive or more difficult to visit compared to other units, could be addressed. This very interesting research question could easily be to subject of one or many new PhD theses.

These research questions, and many other questions that are not given in this list, illustrate the need for future research in this interesting and challenging field of ecological statistics.

Chapter 8

Appendix

8.1 Appendix A: Data case studies

8.1.1 Ecological Covariates

The auxiliary information used in this thesis, are based on the standard 19 bioclimatic variables (BIOCLIM) used by Worldclim, an online databank for global climate data (<http://www.worldclim.org/>). On there website they discribe these 19 ecological covariates as:

1. **Annual Mean Temperature:** The mean of all the weekly mean temperatures. Each weekly mean temperature is the mean of that week's maximum and minimum temperature.
2. **Mean Diurnal Range(Mean(period max-min)):** The mean of all the weekly diurnal temperature ranges. Each weekly diurnal range is the difference between that week's maximum and minimum temperature.
3. **Isothermality 2/7:** The mean diurnal range (parameter 2) divided by the Annual Temperature Range (parameter 7).
4. **Temperature Seasonality:** the temperature Coefficient of Variation as the standard deviation of the weekly mean temperatures expressed as a percentage of the mean of those temperatures (i.e. the annual mean). For this calculation, the mean in degrees Kelvin

is used. This avoids the possibility of having to divide by zero, but does mean that the values are usually quite small.

5. **Max Temperature of Warmest Period:** The highest temperature of any weekly maximum temperature.
6. **Min Temperature of Coldest Period:** The lowest temperature of any weekly minimum temperature.
7. **Temperature Annual Range (5-6):** The difference between the Max Temperature of Warmest Period and the Min Temperature of Coldest Period.
8. **Mean Temperature of Wettest Quarter:** The wettest quarter of the year is determined (to the nearest week), and the mean temperature of this period is calculated.
9. **Mean Temperature of Driest Quarter:** The driest quarter of the year is determined (to the nearest week), and the mean temperature of this period is calculated.
10. **Mean Temperature of Warmest Quarter:** The warmest quarter of the year is determined (to the nearest week), and the mean temperature of this period is calculated.
11. **Mean Temperature of Coldest Quarter:** The coldest quarter of the year is determined (to the nearest week), and the mean temperature of this period is calculated.
12. **Annual Precipitation:** The sum of all the monthly precipitation estimates.
13. **Precipitation of Wettest Period:** The precipitation of the wettest week or month, depending on the time step.
14. **Precipitation of Driest Period:** The precipitation of the driest week or month, depending on the time step.
15. **Precipitation Seasonality:** The Coefficient of Variation is the standard deviation of the weekly precipitation estimates expressed as a percentage of the mean of those estimates (i.e. the annual mean).

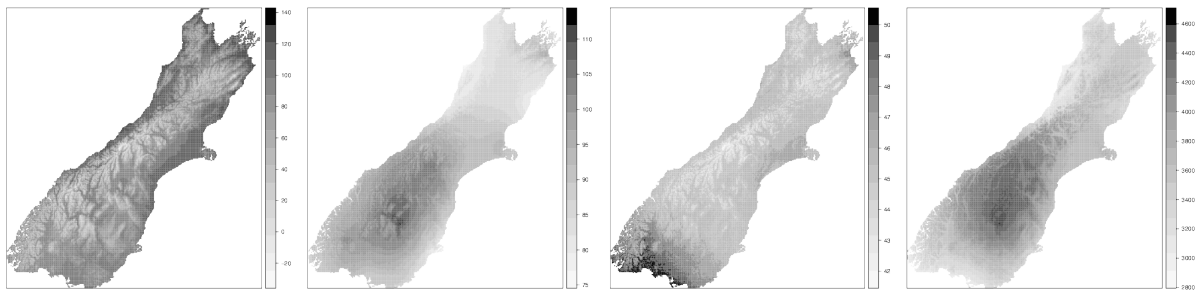
16. **Precipitation of Wettest Quarter:** The wettest quarter of the year is determined (to the nearest week), and the total precipitation over this period is calculated.
17. **Precipitation of Driest Quarter:** The driest quarter of the year is determined (to the nearest week), and the total precipitation over this period is calculated.
18. **Precipitation of Warmest Quarter:** The warmest quarter of the year is determined (to the nearest week), and the total precipitation over this period is calculated.
19. **Precipitation of Coldest Quarter:** The coldest quarter of the year is determined (to the nearest week), and the total precipitation over this period is calculated.

Asian tiger mosquito

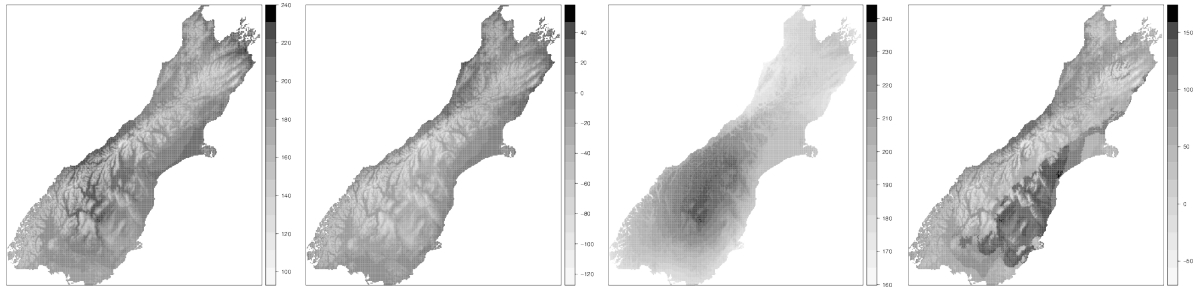
The ecological layer/covariates that were used to create the artificial distribution of the Asian tiger mosquito are illustrated in *Figure 8.1*. The resolution of this data is on the original scale. Each pixel is five by five kilometre. These layers were used both in Chapter 4 and Chapter 5 and the pixel size was reduced for simulation speed purposes.

Great White butterfly

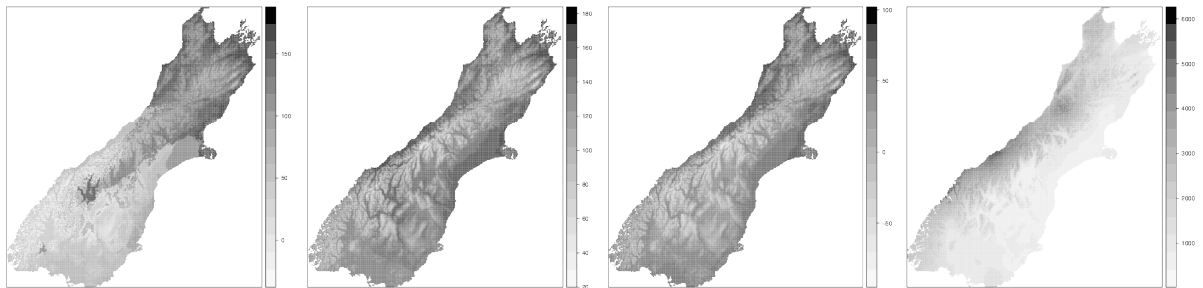
The ecological layer/covariates that were used to create the artificial distribution of the Great White butterfly are illustrated in *Figure 8.2*. The resolution of this data is on the original scale. Each pixel is one by one kilometre. These layers were used Chapter 6.



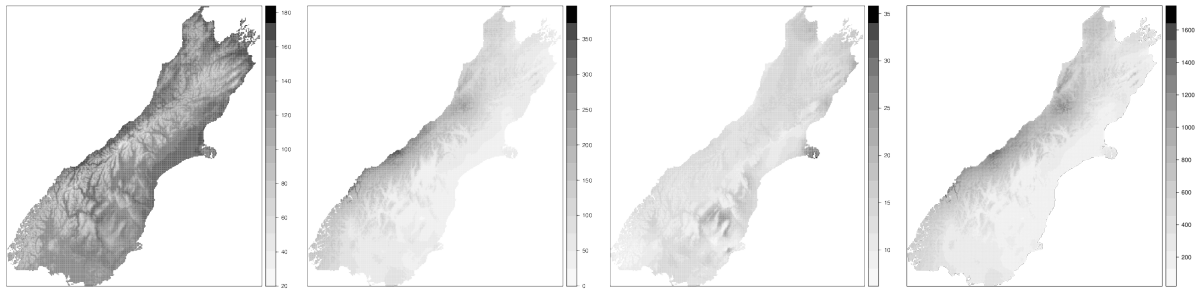
(a) Annual mean temperature (°C) (b) Mean diurnal temperature range (°C) (c) Max temperature of warmest week (°C) (d) Mean temperature of driest quarter (°C)



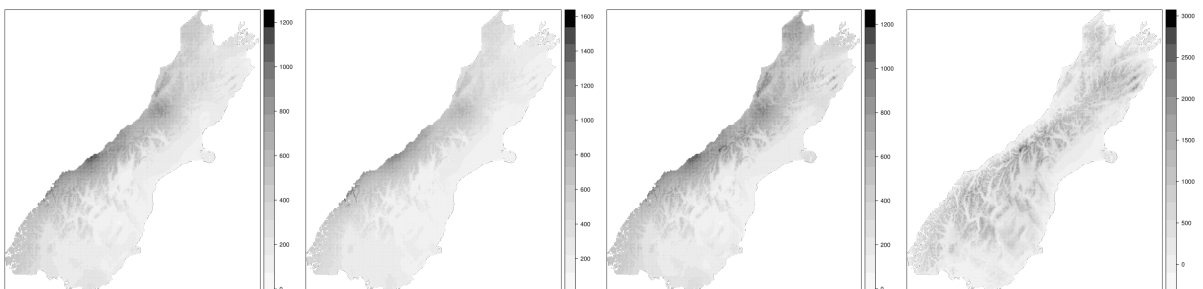
(e) Mean temperature of warmest quarter (°C) (f) Precipitation of driest week (mm) (g) Precipitation of driest quarter (mm) (h) Annual mean radiation (W m-2)



(i) Lowest weekly radiation (W m-2) (j) Radiation seasonality (C of V) (k) Radiation of driest quarter (W m-2) (l) Radiation of warmest quarter (W m-2)

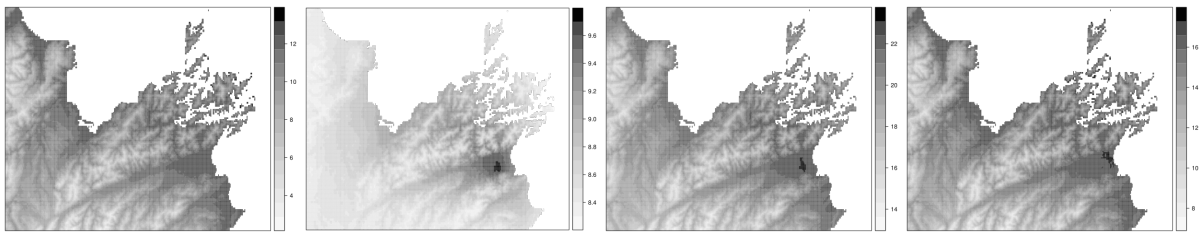


(m) Radiation of coldest quarter (W m-2) (n) Annual mean moisture index (o) Lowest weekly moisture index (p) Mean moisture index of driest quarter

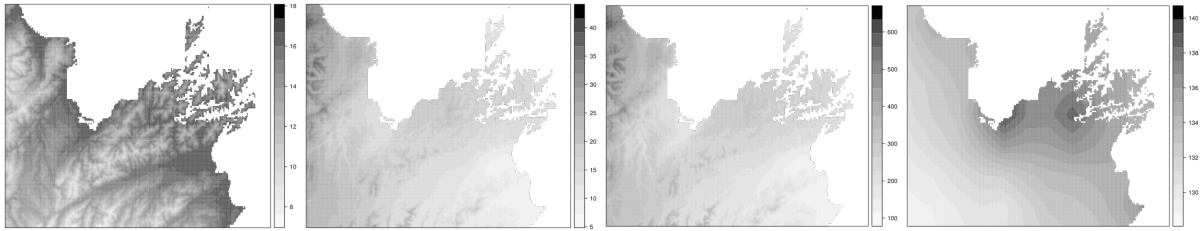


(q) Mean moisture index of coldest quarter (r) Mean moisture index of coldest quarter (s) Mean moisture index of coldest quarter (t) Altitude (m)

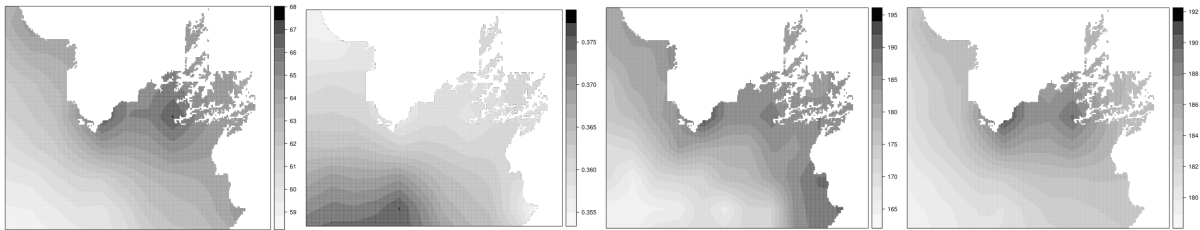
Figure 8.1: covariates used for the Asian tiger mosquito case studies.



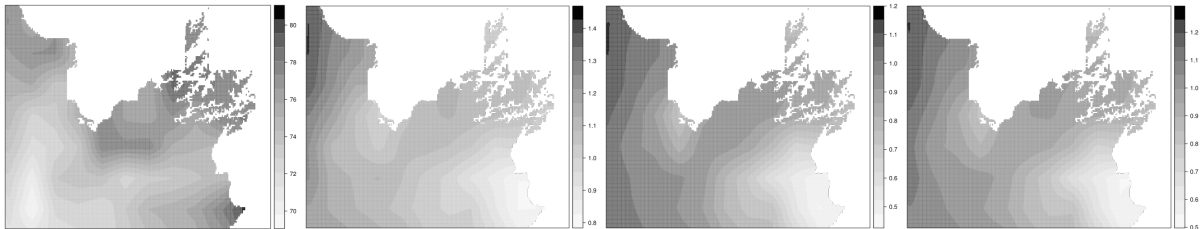
(a) Annual mean temperature (°C) (b) Mean diurnal temperature range (°C) (c) Max temperature of warmest week (°C) (d) Mean temperature of driest quarter (°C)



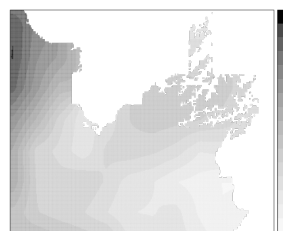
(e) Mean temperature of warmest quarter (°C) (f) Precipitation of driest week (mm) (g) Precipitation of driest quarter (mm) (h) Annual mean radiation (W m-2)



(i) Lowest weekly radiation (W m-2) (j) Radiation seasonality of V (k) Radiation of driest quarter (W m-2) (l) Radiation of warmest quarter (W m-2)



(m) Radiation of coldest quarter (W m-2) (n) Annual mean moisture index (o) Lowest weekly moisture index (p) Mean moisture index of driest quarter



(q) Mean moisture index of coldest quarter

Figure 8.2: covariates used for the Great White butterfly case studies.

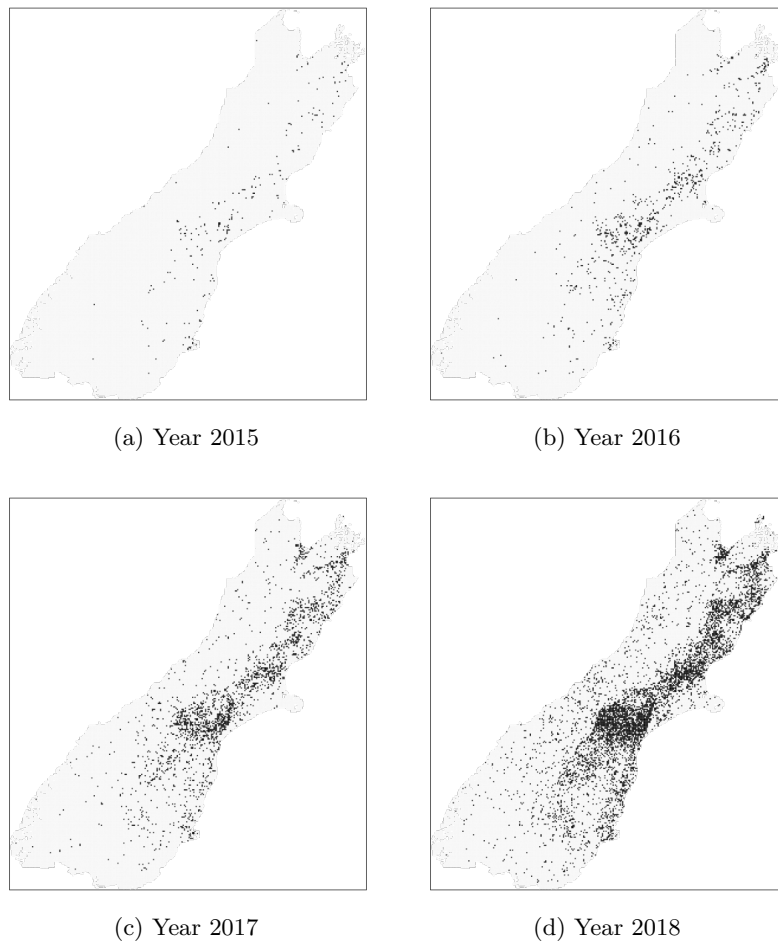


Figure 8.3: Distribution maps Asian tiger mosquito.

8.1.2 Distribution maps Asian tiger mosquito

The artificial distribution of the Asian tiger mosquito from year 2015 till 2018 are illustrated in *Figure 8.3*. The resolution of this data is on the original scale. Each pixel is one by one kilometre. In this thesis, for simulation purpose (mainly speeds related issues), the resolution of the data was reduced to three by three kilometres (Chapter 4) sized pixels or two by two kilometres (Chapter 5) sized pixels.

8.2 Appendix B: R code

Each time BAS was used the R script was rewritten such that it was most convenient to use in that particular context. The following R-script illustrates the R code used for the non-stationary simulation study on the GWB (Chapter 6). This R-script was chosen since it gives an illustration of the use of BAS, NUT and AMBAS together with the usages of ecological raster layers all in one script. Note that more details to the code can be provided if required.

```
#####
##### script settings ###
#####
##### read start up file #####
source("H:/R_codes/functions/startup.R")
i <- c(NA)
if(is.na(i))
{
  startup(par.ini=TRUE, read.library=TRUE, working.directory="s:/pja101/R_
    output")
}
i<-1 # makes sure function startup will only be read once

#####
##### parameter settings #####
#####
factor.aggregate <- 1 # reduce pixel size
rejection.type <- "both" # eco,geo, both (rejection sampling based
  on )
memory <- "both" # no, both, presc, abs
combi.method <- "add" # add or times
n.sims <- 1 # number of simulations
fixed.seed <- "no" # "yes" or "no"
```

```

perc.sampled      <- 0.1      # percentage units sampled
method            <- "PANUT"  # "NUT", "PANUT" or "CNUT"
combi_method      <- "add"    # "add" or "ratio" # for PANUT only
response         <- "binary" # binary or counts
plot.map        <- "yes"    # plot yes or no
limit             <- 1        # value of lower limit NUT
critical.value.geo <- 0.01     # range goes from 0 till 1
critical.value.eco <- 0.01     # range goes from 0 till 1
years             <- c("yr2015", "yr2015", "yr2015", "yr2015", "yr2015", "yr2015", "yr2015", "yr2015") # select years

#####
#####      read in data      ###
#####

source("H:/R_codes/GWB/layers2.R")
dev.off()

##### plot #####
#myTheme <- rasterTheme(region=rev(grey_pal(start = 0.0, end = 0.99)(2)))
#levelplot(stack_R[["year2020"]], par.settings=myTheme, xlab="", margin=F,
  scales=list(x=list(draw=FALSE), y=list(draw=FALSE)))
levelplot(stack_R[["yr2015"]], xlab="", margin=F, scales=list(x=list(draw=
  FALSE), y=list(draw=FALSE)))

# raster to point dataset
stack_R[is.na(bio.1_R)]<-NA
#stack_R[is.na(GWB.2018_R)]<-NA
#stack_R[is.na(GWB.2021_R)]<-NA
stack_R[is.na(surv.global_R)]<-NA
stack_P<-rasterToPoints(stack_R)
# make zero counts are zero instead of NA
for(i in unique(years))
{
  stack_P[is.na(stack_P[, i]), i]<-0
}

```

```

}

#####
#####          Functions #####
#####
##### NBS variance estimator
source("H:/R_codes/functions/NBSvar.R")

##### CNUT function
CNUT<-function()
{
  ### count cells ###
  cell_count<-rep(0,ncells)
  select.list<-which(!is.na(samplepresc))
  voro.box<-rep(0,ncells)
  for(d in 1:ncells)
  {
    dist_M <- rdist(stack_P[c(d,1),c("x","y")], stack_P[which(!is.na(
      samplepresc)),c("x","y")])
    vec<-dist_M[1,]
    select <- which(vec==min(vec))
    voro.box[d]<-select[1]
    cell_count[d]<-stack_P[select.list[select[1]],years[y]]
  }

  cutoff<-quantile(cell_count[cell_count[]!=0],probs=0.99, na.rm=T)
  cell_count[cell_count[]>cutoff]<-cutoff
  #cell_count<-log10(cell_count+1)
  pi_counts<-cell_count*(n.samples/sum(cell_count,na.rm=T))

  pi_equal<-n.samples/ncells
  pi_equal<-rep(pi_equal,ncells)

```

```
geoX_pred<-(0.25*pi_equal+1.75*pi_counts)/2
geoX_pred<-geoX_pred/max(geoX_pred,na.rm=T)
geoX_pred[is.na(geoX_pred)]<-0
geoX_pred[geoX_pred<critical.value.geo]<-critical.value.geo

geoHT_pred<-geoX_pred*(n.samples/sum(geoX_pred))

return(cbind(geoHT_pred,geoX_pred))
}

##### PANUT function #####
PANUT<-function()
{
pres_abs<-samplepresc[!is.na(samplepresc)]
lower.limit<-ncells/(n.samples)*limit
### presence points ###
voro.box<-rep(0,ncells)
cell_type<-rep(0,ncells)

for(d in 1:ncells)
{
dist_M <- rdist(stack_P[c(d,1),c("x","y")], stack_P[which(samplepresc
  []==1),c("x","y")])
vec<-dist_M[1,]
select <- which(vec==min(vec))
voro.box[d]<-select[1]
}

vor.size<-function(variable)
{
return(length(voro.box[which(voro.box[]==variable)]))
}

voro.size<-sapply(voro.box, vor.size)
```

```

voro.size[which(voro.size<lower.limit)]<-lower.limit
geoX_pred_pres<- (1/voro.size)/(max(1/voro.size))

### absence points ###
voro.box<-rep(0,ncells)
cell_type<-rep(0,ncells)

for(d in 1:ncells)
{
  dist_M <- rdist(stack_P[c(d,1),c("x","y")], stack_P[which(samplepresc
    []==0),c("x","y")])
  vec<-dist_M[1,]
  select <- which(vec==min(vec))
  voro.box[d]<-select[1]
  #cell_type[d]<-pres_abs[select[1]]
}

voro.size<-sapply(voro.box, vor.size)
voro.size[which(voro.size<(lower.limit))]<-lower.limit
geoX_pred_abs<- (1/voro.size)/(max(1/voro.size))

### presence and absence ###
HT_pred_pres<-geoX_pred_pres*(n.samples/sum(geoX_pred_pres))
HT_pred_abs<-geoX_pred_abs*(n.samples/sum(geoX_pred_abs))

### ratio ###
if(combi_method == "ratio")
{
  geoX_pred_ratio<-geoX_pred_pres/geoX_pred_abs
  HT_pred<-geoX_pred_ratio*(n.samples/sum(geoX_pred_ratio))
  geoX_pred<-geoX_pred_ratio/max(geoX_pred_ratio)
}
### add ###

```

```
if (combi_method == "add")
{
  geoX_pred_ratio <- geoX_pred_pres - geoX_pred_abs
  geoX_pred_ratio <- geoX_pred_ratio + 1
  geoX_pred_ratio <- geoX_pred_ratio / 2
  geoX_pred <- geoX_pred_ratio / max(geoX_pred_ratio)
  geoX_pred[geoX_pred < critical.value.geo] <- critical.value.geo

  geoHT_pred <- geoX_pred * (n.samples / sum(geoX_pred))
}
return(cbind(geoHT_pred, geoX_pred))
}
```

```
##### GLM function
```

```
GLMfunction <- function()
{
  data.glm_P <- as.data.frame(stack_P[!is.na(samplepresc[ ]),])

  response <- samplepresc[!is.na(samplepresc[ ])]
  data.glm_P <- cbind(data.glm_P, response)

  glm1 <- glm(response ~ # cov01 +
               cov02 +
               # cov05 +
               cov09 +
               # cov10 +
               cov14 +
               # cov17 +
               cov20 +
               # cov22 +
               cov23 +
               # cov25 +
```

```

        cov26 +
      # cov27 +
        cov28 +
      # cov30 +
        cov33
      # cov35
      ,family=binomial , data=data.glm_P)

ecoX_pred<-predict(glm1 , newdata=data.frame(stack_P) , type="response")

ecoX_pred<-ecoX_pred/max(ecoX_pred)
ecoX_pred[is.na(ecoX_pred)]<-0
ecoX_pred[ecoX_pred<critical.value.eco]<-critical.value.eco

ecoHT_pred<- ecoX_pred*(n.samples/sum(ecoX_pred ,na.rm=T))

return(cbind(ecoHT_pred ,ecoX_pred))
}

##### plot maps
plotmaps<-function()
{
  ### species ###
  inputs<-stack_P[,years[y]]
  inputs[inputs[]!=0]<-1
  #inputs[samplepresc[]==1]<-0
  r_R<- rasterize(stack_P[,1:2] , empty_R, inputs)
  myTheme <- rasterTheme(region=c("gray99" ,"black" ,"grey60"))
  print(levelplot(r_R,xlab="" ,par.settings=myTheme, margin=F, scales=list(x=
    list(draw=FALSE) ,y=list(draw=FALSE)) ,
  colorkey =FALSE))
  ### sample ###
  r_R<- rasterize(stack_P[,1:2] , empty_R, sample, fun=sum)

```



```
myTheme <- rasterTheme(region=rev(grey_pal(start = 0.01, end = 0.97)(2)))
print(levelplot(r_R, xlab="", par.settings=myTheme, margin=F, scales=list(x
  =list(draw=FALSE), y=list(draw=FALSE)),
  colorkey =FALSE))
### geo ###
if(rejection.type=="geo")
{
  inputs<-geoX_pred*n.samples/(sum(geoX_pred))
  r_R<- rasterize(stack_P[,1:2], empty_R, inputs, fun=sum)
  print(levelplot(r_R, xlab="", margin=F, scales=list(x=list(draw=FALSE), y=
    list(draw=FALSE))))
}
### eco ###
if(rejection.type=="eco")
{
  inputs<-ecoX_pred*n.samples/(sum(ecoX_pred))
  r_R<- rasterize(stack_P[,1:2], empty_R, inputs, fun=sum)
  print(levelplot(r_R, xlab="", margin=F, scales=list(x=list(draw=FALSE), y
    =list(draw=FALSE))))
}
### both ###
if(rejection.type=="both")
{
  inputs<-((ecoX_pred+geoX_pred)/2)*n.samples/
  (sum(((ecoX_pred+geoX_pred)/2)))
  r_R<- rasterize(stack_P[,1:2], empty_R, inputs, fun=sum)
  print(levelplot(r_R, xlab="", margin=F, scales=list(x=list(draw=FALSE), y
    =list(draw=FALSE))))
}
### species ###
# r1_R<- rasterize(stack_P[,1:2], empty_R, sample)
# inputs<-((ecoX_pred+geoX_pred)/2)*n.samples/
(sum(((ecoX_pred+geoX_pred)/2)))
```

```

# r2_R<- rasterize(stack_P[,1:2], empty_R, inputs, fun=sum)
# r1_R[r2_R>0.16]<-0.5
# r_R<-r1_R
# myTheme <- rasterTheme(region=c("gray97", "black", "grey60"))
# print(levelplot(r1_R, par.setting=myTheme, xlab="", margin=F, scales=list(
  x=list(draw=FALSE), y=list(draw=FALSE))
, colorkey =F))

# myTheme <- rasterTheme(region=rev(grey_pal(start = 0.01,
end = 0.97)(2)))
# inputs<-((ecoX_pred+geoX_pred)/2)*n.samples/
(sum(((ecoX_pred+geoX_pred)/2)))
# r_R<- rasterize(stack_P[,1:2], empty_R, inputs, fun=sum)
# print(levelplot(r_R, par.settings=myTheme, xlab="", margin=F, scales=list(
  x=list(draw=FALSE), y=list(draw=FALSE))))

# r_M<-as.matrix(r_R)
# wireframe(r_M, screen = list(z = 290, x = -65, y = 3)
, drape = TRUE,
# par.settings = list(axis.line = list(col = "
transparent")), zlim=c(0,1), shade=F,
scales = list(arrows = FALSE, cex=1.2),
# xlab=list("Y-coordinate", rot = -60, cex=1.3), ylab=
list("X-coordinate", rot = 15, cex=1.3),
zlab=list("Rescaled_inclusion_Probability", rot = 90, cex=1.3), colorkey=
F)
# plot3D(GWB.2022_R*100, col=heat.colors)
}

#####
##### preparing data #####
#####

```

```
# point dataset
empty_R<-stack_R[["cov01"]]
empty_R[!is.na(empty_R)]<-0
dim_R<-dim(empty_R)
ncells<-dim_R[1]*dim_R[2]
ext_R<-extent(empty_R)
cell<-1:ncells
vals<-getValues(empty_R)
cell.id<-cell[which(vals[]==0)]
stack_P<-cbind(stack_P, cell.id)
# number of cells
ncells<-length(stack_P[,1])
# sample size
n.samples<-round(perc.sampled*ncells,0)
# halton sequence
dat.halton.all<-runif.halton(n=5000000,dim=3)
# results matrix
dat.result<-matrix(nrow=0,ncol=6)

# select a type of dispersal
my<-0
mx<-0
mv<-c(my,mx)
sdy<-0.25
sdx<-0.25
rho<-0.0
Sigma<-matrix(c(sdy^2,sdy*sdx*rho,sdy*sdx*rho,sdx^2),2,2)

# make species raster
species_R<-empty_R
species1_R<-stack_R[[years[1]]]
species_R[species1_R[]==1]<-1
```

```
#####
##### start simulations ##
#####
for (s in 1:n.sims)
{
  print("simulation")
  print(s)
  y<-1

  stack_P[stack_P[, years [y]]>0, years [y]]<-1

  halton.seed<-runif(1,1,1000000)      #if (fixed.seed=="yes")
  {halton.seed<-1}
  dat.halton<-dat.halton.all[-c(1:halton.seed),]

  x.raster <- ceiling(dat.halton[,1]*dim.R[2])
  y.raster <- ceiling(dat.halton[,2]*dim.R[1])
  cell.select<-dim.R[2]*(y.raster-1) + x.raster
  cell.select<-cell.select[which(cell.select%in% cell.id)]
  halton.select<-dat.halton[which(cell.select%in% cell.id),3]

  #####
  ##### sample year 1 #####
  HT_pred<-rep(n.samples/ncells, ncells)

  ##### take sample #####
  sample.select_V<-cell.select[1:n.samples]

  sample<-rep(0, ncells)
  sample[which(stack_P[, "cell.id"]%in% sample.select_V)]<-1

  samplepresc<-rep(NA, ncells)
  samplepresc[ which(sample[]==1 & stack_P[, years [y]]==0 )] <-0
}
```

```
samplepresc [ which(sample[]==1 & stack_P[,years [y]]>0 )]<-1

# combine presence points
samplepresc_previous<-samplepresc

# estimation
d <- length(which(samplepresc==1))/n.samples
HT_pred<-rep(n.samples/ncells ,ncells)
HT<- sum(sample*stack_P[,years [y]]/(n.samples/ncells))
nbs<-1
#res <- EST(stack_P[,years [y]],HT_pred ,stack_P[,1] ,stack_P[,2] ,sample ,
vartype='Local')
#nbs<-res$SE

species<- stack_P[,years [y]]
species<- species [sample [] !=1]
count_species<-sum(species)

new.entry<-c(y,s,d,HT,nbs ,count_species)
dat.result<-rbind(dat.result ,new.entry)

##### spatial information #####
if(rejection.type!="eco")
{
##### CNUT #####
if(method=="CNUT")
{
output<-CNUT()
geoHT_pred<-output [,1]
geoX_pred<-output [,2]
}
}

##### PANUT #####
```

```

if(method=="PANUT")
  {
  output<-PANUT()
  geoHT_pred<-output[,1]
  geoX_pred<-output[,2]
  }
}

##### ecological information #####
if(rejection.type!="geo")
  {
  output<-GLMfunction()
  ecoHT_pred<-output[,1]
  ecoX_pred<-output[,2]
  }

##### plot #####
if(plot.map=="yes")
{plotmaps()}

#####
##### sample year X #####

n.halt<-n.samples

for(y in 2:length(years))
  {
  print("year")
  print(y)
  ## dispersal #####
  eradication_R<- rasterize(stack_P[,1:2], empty_R, sample, fun=sum)
  species_R[eradication_R]==1<-0
  levelplot(species_R)
  levelplot(eradication_R)
  presence.units<-which(species_R==1)

```

```
list.units<-c(presence.units)

for( i in presence.units)
{
  print(i)
  n.gwb<-rpois(1,0.5)

  if(n.gwb>0)
  {
    xy<-cbind(rlaplace(n.gwb, location = 0, scale = 2),
              rlaplace(n.gwb, location = 0, scale = 2))
    xy[xy[] < -15]<- -15
    xy[xy[] > 15]<- 15
    xy<-xy*50
    xy.new<-xyFromCell(species_R, i, spatial=F)
    xy[,1]<-xy.new[1]+xy[,1]
    xy[,2]<-xy.new[2]+xy[,2]

    new.presence<-cellFromXY(species_R, xy=data.frame(xy))
    list.units<-c(list.units,new.presence)
  }
}

list.units<-list.units[!is.na(list.units)]
accept.reject<-runif(length(list.units))*100
surv<-surv.global_R[list.units]
list.units2<-list.units[surv>accept.reject]
species_R[list.units2]<-1
species_points<-rasterToPoints(species_R)
stack_P[,years[y]]<-species_points[,3]
species_R<- rasterize(stack_P[,1:2], empty_R, stack_P[,years[1]]
,fun=sum)

print(levelplot(species_R))
```

```
#####

stack_P[stack_P[, years [y]] > 0, years [y]] <- 1

sample.select_V <- c()
while (length (sample.select_V) < n.samples)
{
  n.halt <- n.halt + 1
  if (rejection.type == "geo")
  {
    HT_pred <- geoHT_pred
    if (halton.select [n.halt] < geoX_pred [which (stack_P [, "cell.id"
] == cell.select [n.halt])])
    {
      sample.select_V <- c (sample.select_V, cell.select [n.halt])
    }
  }
  if (rejection.type == "eco")
  {
    HT_pred <- ecoHT_pred
    if (halton.select [n.halt] < ecoX_pred [which (stack_P [, "cell.id"
] == cell.select [n.halt])])
    {
      sample.select_V <- c (sample.select_V, cell.select [n.halt])
    }
  }
  if (rejection.type == "both")
  {
    if (combi.method == "add")
    {
      bothX_pred <- geoX_pred + ecoX_pred / 2
      bothX_pred <- bothX_pred / max (bothX_pred, na.rm = T)
      HT_pred <- (geoHT_pred + ecoHT_pred) / 2
    }
  }
}
```



```

    if(halton.select[n.halt]<bothX_pred[which(stack_P[,"cell.id"]
==cell.select[n.halt])])
    {
      sample.select_V<-c(sample.select_V,cell.select[n.halt])
    }
  if(combi.method=="times")
  {
    bothX_pred<-geoX_pred * ecoX_pred
    bothX_pred<-bothX_pred/max(bothX_pred,na.rm = T)
    HT_pred<-(geoHT_pred+ecoHT_pred)/2
    if(halton.select[n.halt]<bothX_pred[which(stack_P[,"cell.id"]
==cell.select[n.halt])])
    {
      sample.select_V<-c(sample.select_V,cell.select[n.halt])
    }
  }
}
}
}

sample<-rep(0,ncells)
sample[which(stack_P[,"cell.id"]%in% sample.select_V)]<-1

if(memory=="no")
{
  samplepresc<-rep(NA,ncells) # no memory
}
if(memory=="both")
{
  samplepresc<-samplepresc_previous # memory presence and absence
}

samplepresc[ which(sample[]==1 & stack_P[,years[y]]==0 ) ] <-0

```

```

samplepresc [ which(sample[]==1 & stack_P[, years [y]] >0 )]<-1

# combine presence points
samplepresc_previous<-samplepresc

# estimation
d <- length(which(sample==1 & stack_P[, years [y]] >0))/n.samples
HT<- sum(sample*stack_P[, years [y]]/HT_pred)
#res <- EST(stack_P[, years [y]],HT_pred, stack_P[, 1], stack_P[, 2], sample,
  vartype='Local ')
#nbs<-res$SE

species<- stack_P[, years [y]]
species<- species [sample [] !=1]
count_species<-sum(species)

new.entry<-c(y, s, d, HT, nbs, count_species)
dat.result<-rbind(dat.result, new.entry)

##### spatial information #####
if(rejection.type!="eco")
{
  ##### CNUT #####
  if(method=="CNUT")
  {
    output<-CNUT()
    geoHT_pred<-output [, 1]
    geoX_pred<-output [, 2]
  }

  ##### PANUT #####
  if(method=="PANUT")
  {

```

```
    output<-PANUT()
    geoHT_pred<-output[,1]
    geoX_pred<-output[,2]
  }
}
##### ecological information #####
if(rejection.type!="geo")
{
  output<-GLMfunction()
  ecoHT_pred<-output[,1]
  ecoX_pred<-output[,2]
}

##### plot #####
if(plot.map=="yes")
{plotmaps()}

#####
##### end simulations ####
} # bracket years loop
} # bracket simulations loop
#####
##### end simulations ####

plot(dat.result[,1],dat.result[,6])
points(dat.result[,1],dat.result[,3],col=2)
aggregate(dat.result[,6],by=list(dat.result[,1]),mean)

warnings()
#write.csv(dat.result,"GWB-geo.csv")
```

Bibliography

- Albert, C. H., N. G. Yoccoz, T. C. Edwards, C. H. Graham, N. E. Zimmermann, and W. Thuiller (2010, December). Sampling in ecology and evolution - bridging the gap between theory and practice. *Ecography* 33(6), 1028–1037.
- Alexander, H. M., N. A. Slade, and W. D. Kettle (1997, June). Application of mark-recapture models to estimation of the population size of plants. *Ecology* 78(4), 1230–1237.
- Armstrong, A. J. and S. L. Louw (2013, February). Monitoring of the eggs of the Karkloof blue butterfly, *Orachrysops ariadne*, for its conservation management. *Koedoe - African Protected Area Conservation and Science* 55(1), 1–11.
- Aukema, J. E., D. G. McCullough, B. Von Holle, A. M. Liebhold, K. Britton, and S. J. Frankel (2010, December). Historical Accumulation of Nonindigenous Forest Pests in the Continental United States. *BioScience* 60(11), 886–897.
- Bahn, V. and B. J. McGill (2007, November). Can niche-based distribution models outperform spatial interpolation? *Global Ecology and Biogeography* 16(6), 733–742.
- Benedetti, R., M. Bee, G. Espa, and F. Piersimoni (2010, April). *Agricultural Survey Methods*. Chichester, UK: John Wiley & Sons, Ltd.
- Bickford, C. A., C. E. Mayer, and K. D. Ware (1963). An Efficient Sampling Design for Forest Inventory: The Northeastern Forest Resurvey. *Journal of Forestry* 61(11), 8–18.
- Bogich, T. L., A. M. Liebhold, and K. Shea (2008, August). To sample or eradicate? A cost minimization model for monitoring and managing an invasive species. *Journal of Applied Ecology* 45(4), 1134–1142.
- Brewer, K. R. W. and M. Hanif (1983). *Sampling with unequal probabilities*. New York: Springer-Verlag.

- Brown, J. A., M. M. Salehi, M. M. Moradi, G. Bell, and D. Smith (2008). An adaptive two-stage sequential design for sampling rare and clustered populations. *Population Ecology* 50(3), 239–245.
- Brus, D. and J. de Gruijter (1997, October). Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). *Geoderma* 80(1-2), 1–44.
- Buckley, Y. M. (2008, April). The role of research for integrated management of invasive species, invaded landscapes and communities. *Journal of Applied Ecology* 45(2), 397–402.
- Burgman, M. A. and J. C. Fox (2003, February). Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. *Animal Conservation* 6(1), 19–28.
- Burgman, M. A., M. A. McCarthy, A. Robinson, S. M. Hester, M. F. McBride, J. Elith, and F. Dane Panetta (2013, May). Improving decisions for invasive species management: reformulation and extensions of the Panetta-Lawes eradication graph. *Diversity and Distributions* 19(5-6), 603–607.
- Caughlan, L. and K. L. Oakley (2001). Cost considerations for long-term ecological monitoring. *Ecological Indicators* 1(2), 123–134.
- Cochran, W. G. (1977). *Sampling techniques*. Wiley, New York.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20, 37–40.
- Colwell, R. K. and T. F. Rangel (2009, November). Hutchinson’s duality: the once and future niche. *Proceedings of the National Academy of Sciences of the United States of America* 106 Suppl(Supplement_2), 19651–8.
- Comte, L. and G. Grenouillet (2013, March). Species distribution modelling and imperfect detection: comparing occupancy versus consensus methods. *Diversity and Distributions* 19(8), 996–1007.
- Crall, A. W., C. S. Jarnevich, B. Panke, N. Young, M. Renz, and J. Morissette (2013, January). Using habitat suitability models to target invasive plant species surveys. *Ecological Applications* 23(1), 60–72.
- Dalenius, T. and J. L. Hodges (1957). The choice of stratification points. *Skandinavisk Aktuarietidskrift* 40, 198–203.

- de Gruijter, J. J. and C. J. F. ter Braak (1990, May). Model-free estimation from spatial samples: A reappraisal of classical sampling theory. *Mathematical Geology* 22(4), 407–415.
- Deville, J. C. (1998, March). Unequal probability sampling without replacement through a splitting method. *Biometrika* 85(1), 89–101.
- Dodder, N. G., K. A. Maruya, G. G. Lauenstein, J. Ramirez, K. J. Ritter, and K. C. Schiff (2012, October). Distribution and sources of polybrominated diphenyl ethers in the Southern California Bight. *Environmental toxicology and chemistry* 31(10), 2239–45.
- Dunn, R. and A. Harrison (1993). Two-dimensional systematic sampling of land use. *Journal of the Royal Statistical Society* 42(4), 585–601.
- Edwards, T. C., D. R. Cutler, N. E. Zimmermann, L. Geiser, and J. Alegria (2005, May). Model-based stratification for enhancing the detection of rare ecological events. *Ecology* 86(5), 1081–1090.
- Ekman, G. (1959). Approximate expressions for the conditional mean and variance over small intervals of a continuous distribution. *The Annals of Mathematical Statistics* 30(4), 1131–1134.
- Elith, J. and C. H. Graham (2009, April). Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. *Ecography* 32(1), 66–77.
- Elith, J. and J. R. Leathwick (2009, December). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics* 40(1), 677–697.
- Engler, R., A. Guisan, and L. Rechsteiner (2004, April). 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.
- Epanchin-Niell, R. S., R. G. Haight, L. Berec, J. M. Kean, and A. M. Liebhold (2012, August). Optimal surveillance and eradication of invasive species in heterogeneous landscapes. *Ecology letters* 15(8), 803–12.
- Foxcroft, L. C. (2004, December). An Adaptive Management Framework for Linking Science and Management of Invasive Alien Plants 1. *Weed Technology* 18(sp1), 1275–1277.

- Gaeta, J. W., J. S. Read, J. F. Kitchell, and S. R. Carpenter (2012, April). Eradication via destratification: whole-lake mixing to selectively remove rainbow smelt, a cold-water invasive species. *Ecological Applications* 22(3), 817–827.
- Giera, N. and B. Bell (2008). Economic Costs of Pests to New Zealand. Technical report, MAF Biosecurity New Zealand.
- Gitzen, R. A., J. J. Millspaugh, A. B. Cooper, and D. S. Licht (2012). Spatial Sampling Designs for Long-Term Monitoring. In *Design and Analysis of Long-term Ecological Monitoring Studies*, Chapter 5, pp. 101–126. Cambridge University Press.
- Glasby, T. M. (2012, October). *Caulerpa taxifolia* in seagrass meadows: killer or opportunistic weed? *Biological Invasions* 15(5), 1017–1035.
- Grafström, A. and N. L. P. Lundström (2013). Why Well Spread Probability Samples Are Balanced. *Open Journal of Statistics* 3(1), 36–41.
- Grafström, A., N. L. P. Lundström, and L. Schelin (2012a, June). Spatially balanced sampling through the pivotal method. *Biometrics* 68(2), 514–20.
- Grafström, A., N. L. P. Lundström, and L. Schelin (2012b, June). Spatially balanced sampling through the pivotal method. *Biometrics* 68(2), 514–20.
- Grafström, A. and Y. Tillé (2013, March). Doubly balanced spatial sampling with spreading and restitution of auxiliary totals. *Environmetrics* 24(2), 120–131.
- Guisan, A., O. Broennimann, R. Engler, M. Vust, N. G. Yoccoz, A. Lehmann, and N. E. Zimmermann (2006, April). Using Niche-Based Models to Improve the Sampling of Rare Species. *Conservation Biology* 20(2), 501–511.
- Guisan, A. and W. Thuiller (2005, September). Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8(9), 993–1009.
- Hagstrum, D. W. and B. Subramanyam (2010). Immature Insects: Ecological Roles of Mobility. *American Entomologist* 56(4), 231.
- Hajek, J. (1981). *Sampling from a finite population*. New York: Dekker.
- Halton, J. H. (1960). On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik* 2(1), 84–90.

- Harvey, C. T., S. A. Qureshi, and H. J. MacIsaac (2009, May). Detection of a colonizing, aquatic, non-indigenous species. *Diversity and Distributions* 15(3), 429–437.
- Hazard, J. W., B. E. Law, K. E. Associates, and C. E. R. Laboratory (1989). *Forest survey methods used in the USDA Forest Service*. Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency.
- Hijmans, R. J. (2012, March). Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology* 93(3), 679–688.
- Hijmans, R. J., S. J. Phillips, J. R. Leathwick, and J. Elith (2013). R package : Dismo: Species distribution modeling.
- Hill, N., E. Lawrence, J. Dambacher, N. Barrett, J. Hulls, A. Williams, B. Barker, S. Nichols, V. Lucieer, F. Althaus, J. Kool, and K. Hayes (2013, August). Developing long-term monitoring programs in offshore waters with little prior knowledge: Applying a novel sampling design to inventory biological assets.
- Homans, F. and T. Horie (2011, April). Optimal detection strategies for an established invasive pest. *Ecological Economics* 70(6), 1129–1138.
- Horvitz, D. G. and D. J. Thompson (1952, December). A Generalization of Sampling Without Replacement from a Finite Universe. *Journal of the American Statistical Association* 47(260), 663–685.
- Hulme, P. E. (2009, February). Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology* 46(1), 10–18.
- Kean, J. and C. Phillips (2013a). Phenology and diapause research for great white butterfly (*Pieris brassicae*) (No. RE400/2012/445, 10 October 2013). Technical report, AgResearch, Lincoln, New Zealand.
- Kean, J. and C. Phillips (2013b). Phenology and diapause research for great white butterfly (*Pieris brassicae*) (No. RE400/2012/445, 2 July 2013). Technical report, AgResearch, Lincoln, New Zealand, Lincoln.
- Kean, J. M., C. S. B. Phillips, and M. R. McNeill (2008). Surveillance for early detection: lottery or investment? *New Zealand Plant Protection* 61, 11–17.

- Kerrya, R., P. Goovaerts, I. Smit, and I. B.R. (2013). A comparison of multiple indicator kriging and area-to-point Poisson kriging for mapping patterns of herbivore species abundance in Kruger National Park, South Africa. *International Journal of Geographical Information Science* 27(1), 47–67.
- Knotters, M., D. Brus, and J. Oude Voshaar (1995, August). A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. *Geoderma* 67(3-4), 227–246.
- Kovacs, K. F., R. G. Haight, D. G. McCullough, R. J. Mercader, N. W. Siegert, and A. M. Liebhold (2010). Cost of potential emerald ash borer damage in U.S. communities, 2009–2019. *Ecological Economics* 69(3), 569–578.
- Krige, D. G. (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Chemical Metal and Mining Society of South Africa* 52(6), 119–139.
- Lackey, L. G. and E. D. Stein (2013, May). Evaluation of Design-Based Sampling Options for Monitoring Stream and Wetland Extent and Distribution in California. *Wetlands* 33(4), 717–725.
- Lavallée, P. and M. Hidiroglou (1988). On the stratification of skewed populations. *Survey Methodology* 14, 33–43.
- Le Lay, G., R. Engler, E. Franc, and A. Guisan (2010, December). Prospective sampling based on model ensembles improves the detection of rare species. *Ecography* 33(6), 1015–1027.
- Legendre, P. (1993, September). Spatial Autocorrelation: Trouble or New Paradigm? *Ecology* 74(6), 1659–1673.
- Lemke, D., C. Schweitzer, I. Tazisong, Y. Wang, and J. Brown (2013, August). Invasion of a mined landscape: what habitat characteristics are influencing the occurrence of invasive plants? *International Journal of Mining, Reclamation and Environment* 27(4), 275–293.
- Levy, P. S. and S. Lemeshow (2013). *Sampling of Populations: Methods and Applications*. John Wiley & Sons.
- Lindenmayer, D. B. and G. E. Likens (2009, September). Adaptive monitoring: a new paradigm for long-term research and monitoring. *Trends in ecology & evolution* 24(9), 482–6.
- Lindenmayer, D. B., G. E. Likens, A. Haywood, and L. Mieziš (2011, December). Adaptive monitoring in the real world: proof of concept. *Trends in ecology & evolution* 26(12), 641–6.

- Lister, A. J. and C. T. Scott (2009, February). Use of space-filling curves to select sample locations in natural resource monitoring studies. *Environmental monitoring and assessment* 149(1-4), 71–80.
- Lobo, J. M., A. Jiménez-Valverde, and J. Hortal (2010, February). The uncertain nature of absences and their importance in species distribution modelling. *Ecography* 33(1), 103–114.
- Lodge, D. M., S. Williams, H. J. MacIsaac, K. R. Hayes, B. Leung, S. Reichard, R. N. Mack, P. B. Moyle, M. Smith, D. A. Andow, J. T. Carlton, and A. McMichael (2006, December). Biological invasions : Recommendations for U.S. policy and management. *Ecological Applications* 16(6), 2035–2054.
- Lohr, S. L. (2010). *Sampling: Design and Analysis: Design And Analysis*. Cengage Learning.
- Lyet, A., W. Thuiller, M. Cheylan, and A. Besnard (2013, January). Fine-scale regional distribution modelling of rare and threatened species: bridging GIS Tools and conservation in practice. *Diversity and Distributions* 19(7), 651–663.
- Manel, S., J.-M. Dias, and S. J. Ormerod (1999, August). Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecological Modelling* 120(2-3), 337–347.
- McArdle, B. H. and R. Blackwell (1989). Measurement of density variability in the bivalve *Chione stutchburyi* using spatial autocorrelation. *Marine ecology progress series* 52(3), 245.
- Medlock, J. M., K. M. Hansford, F. Schaffner, V. Versteirt, G. Hendrickx, H. Zeller, and W. Van Bortel (2012, June). A review of the invasive mosquitoes in Europe: ecology, public health risks, and control options. *Vector borne and zoonotic diseases (Larchmont, N.Y.)* 12(6), 435–47.
- Mehta, S. V., R. G. Haight, F. R. Homans, S. Polasky, and R. C. Venette (2007, March). Optimal detection and control strategies for invasive species management. *Ecological Economics* 61(2-3), 237–245.
- Moore, J. L., C. E. Hauser, J. L. Bear, N. S. G. Williams, and M. A. McCarthy (2011, March). Estimating detection-effort curves for plants using search experiments. *Ecological Applications* 21(2), 601–607.
- Nelder, J. and R. Wedderburn (1972). Generalized Linear Models. *Journal of the Royal Statistical Society*. 135(3), 370–384.
- Nenzén, H. and M. Araújo (2011, September). Choice of threshold alters projections of species range shifts under climate change. *Ecological Modelling* 222(18), 3346–3354.

- Neyman, J. (1934). On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection. *Journal of the Royal Statistical Society* 97(4), 558–625.
- Olea, R. A. (1984, May). Sampling design optimization for spatial functions. *Journal of the International Association for Mathematical Geology* 16(4), 369–392.
- Olea, R. A. (2006, May). A six-step practical approach to semivariogram modeling. *Stochastic Environmental Research and Risk Assessment* 20(5), 307–318.
- Oliver, M. and R. Webster (2014, February). A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *Catena* 113, 56–69.
- Parnell, S., T. R. Gottwald, C. A. Gilligan, N. J. Cunniffe, and F. van den Bosch (2010, July). The effect of landscape pattern on the optimal eradication zone of an invading epidemic. *Phytopathology* 100(7), 638–44.
- Parnell, S., T. R. Gottwald, F. van den Bosch, and C. A. Gilligan (2009, December). Optimal strategies for the eradication of asiatic citrus canker in heterogeneous host landscapes. *Phytopathology* 99(12), 1370–6.
- Payandeh, B. (1970). Relative Efficiency of Two-Dimensional Systematic Sampling. *Forest Science* 16(6), 271–276.
- Peterman, W. E., J. A. Crawford, and A. R. Kuhns (2013, April). Using species distribution and occupancy modeling to guide survey efforts and assess species status. *Journal for Nature Conservation* 21(2), 114–121.
- Phillips, C., K. Brown, C. Green, G. Walker, K. Broome, R. Toft, B. V. Lee, and M. King (2013). Great white butterfly interim report prepared for Ministry for Primary Industries external technical advisory group.
- Phillips, S. J., M. Dudík, and R. E. Schapire (2004, July). A maximum entropy approach to species distribution modeling. In *Twenty-first international conference on Machine learning - ICML '04*, New York, New York, USA, pp. 83. ACM Press.
- Pimentel, D., S. McNair, J. Janecka, J. Wightman, C. Simmonds, C. O'Connell, E. Wong, L. Russel, J. Zern, T. Aquino, and T. Tsomondo (2001). Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems & Environment* 84(1), 1–20.

- Pimentel, D., R. Zuniga, and D. Morrison (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics* 52(3), 273–288.
- Pitt, J. P. W., S. P. Worner, and A. V. Suarez (2009, June). Predicting Argentine ant spread over the heterogeneous landscape using a spatially explicit stochastic model. *Ecological Applications* 19(5), 1176–1186.
- Ramsey, D. S. L. and M. G. Efford (2010, June). Management of bovine tuberculosis in brushtail possums in New Zealand: predictions from a spatially explicit, individual-based model. *Journal of Applied Ecology* 47(4), 911–919.
- Ranjan, R., S. T. D Evelyn, N. Tarui, K. Burnett, and J. A. Roumasset (2008). Learning-by-catching: Uncertain invasive-species populations and the value of information. *Journal of Environmental Management* 89(4), 284–292.
- Ripley, B. D. (1991). *Statistical Inference for Spatial Processes*. Cambridge University Press.
- Robertson, B. L., J. A. Brown, T. McDonald, and P. Jaksons (2013, July). BAS: Balanced acceptance sampling of natural resources. *Biometrics* 69(3), 1–9.
- Robertson, M. P., M. H. Villet, and A. R. Palmer (2004, September). A fuzzy classification technique for predicting species distributions: applications using invasive alien plants and indigenous insects. *Diversity and Distributions* 10(5-6), 461–474.
- Rosén, B. (1997, August). Asymptotic theory for order sampling. *Journal of Statistical Planning and Inference* 62(2), 135–158.
- Rout, T. M., J. L. Moore, and M. A. McCarthy (2014, March). Prevent, search or destroy? A partially observable model for invasive species management. *Journal of Applied Ecology* 51(3), n/a–n/a.
- Royle, J. A., R. B. Chandler, C. Yackulic, and J. D. Nichols (2012, June). Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods in Ecology and Evolution* 3(3), 545–554.
- Salehi, M. M. and J. A. Brown (2010). Complete allocation sampling: an efficient and easily implemented adaptive sampling design. *Population Ecology* 52(3), 451–456.
- Seaman, D. E. and P. A. Roger (1996). An Evaluation of the Accuracy of Kernel Density Estimators for Home Range Analysis. *Ecology* 77(7), 2075–2085.

- Senay, S. D. (2014a). *Hybrid species distribution modelling (Chapter 6)*. Ph. D. thesis, Lincoln University.
- Senay, S. D. (2014b). *Landscape mapping for spatially explicit species dispersal models (Chapter 7)*. Ph. D. thesis, Lincoln University.
- Senay, S. D., S. P. Worner, and T. Ikeda (2013, August). Novel Three-Step Pseudo-Absence Selection Technique for Improved Species Distribution Modelling. *PLoS ONE* 8(8), e71218.
- Stevens, D. L. and A. R. Olsen (2003, September). Variance estimation for spatially balanced samples of environmental resources. *Environmetrics* 14(6), 593–610.
- Stevens, D. L. and A. R. Olsen (2004). Spatially Balanced Sampling of Natural Resources. *Journal of the American Statistical Association* 99(465), 262–278.
- Stockwell, D. R. and A. Peterson (2002, February). Effects of sample size on accuracy of species distribution models. *Ecological Modelling* 148(1), 1–13.
- Stohlgren, T. J. and J. L. Schnase (2006, February). Risk analysis for biological hazards: what we need to know about invasive species. *Risk analysis : an official publication of the Society for Risk Analysis* 26(1), 163–73.
- Strayer, D. L. (2010, January). Alien species in fresh waters: ecological effects, interactions with other stressors, and prospects for the future. *Freshwater Biology* 55, 152–174.
- Su, Z. and T. J. Q. II (2003, March). Estimator bias and efficiency for adaptive cluster sampling with order statistics and a stopping rule. *Environmental and Ecological Statistics* 10(1), 17–41.
- Suckling, D. M., P. C. Tobin, D. G. McCullough, and D. A. Herms (2012, February). Combining Tactics to Exploit Allee Effects for Eradication of Alien Insect Populations. *Journal of Economic Entomology* 105(1), 1–13.
- Theobald, D., D. Stevens, D. White, N. Urquhart, A. Olsen, and J. Norman (2007). Using GIS to Generate Spatially Balanced Random Survey Designs for Natural Resource Applications. *Environmental Management* 40(1), 134–146.
- Thompson, S. K. (1990, February). Adaptive Cluster Sampling. *Journal of the American Statistical Association* 85(412), 1050–1059.
- Thompson, S. K. (2006, December). Adaptive web sampling. *Biometrics* 62(4), 1224–34.

- Thompson, S. K. (2012, November). Adaptive web sampling in ecology. *Statistical Methods & Applications* 22(1), 33–43.
- Thompson, S. K. and G. A. F. Seber (1996). *Adaptive Sampling*. Wiley, New York.
- Tillé, Y. (2006). *Sampling Algorithms*. Springer Series in Statistics. Springer New York.
- Tillé, Y. (2011). Ten years of balanced sampling with the cube method : An appraisal. *Survey methodology* 3(1), 215–226.
- Tobin, P. C., L. Berec, and A. M. Liebhold (2011, June). Exploiting Allee effects for managing biological invasions. *Ecology letters* 14(6), 615–24.
- Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* 46(2), 234–240.
- Tombre, I. M. a., E. b. Eythórsson, and J. Madsen (2013). Stakeholder involvement in adaptive goose management; case studies and experiences from Norway. *Ornis Norvegica* 36, 17–24.
- Vapnik, V. (1999). *The Nature of Statistical Learning Theory*, Volume 1999. Springer.
- von Neumann, J. (1951). Various techniques used in connection with random digits. Monte Carlo methods. *Natural Bureau Standards* 12, 36–38.
- Wang, J.-F., C.-S. Jiang, M.-G. Hu, Z.-D. Cao, Y.-S. Guo, L.-F. Li, T.-J. Liu, and B. Meng (2013, February). Design-based spatial sampling: Theory and implementation. *Environmental Modelling & Software* 40, 280–288.
- Wang, J.-F., A. Stein, B.-B. Gao, and Y. Ge (2012, December). A review of spatial sampling. *Spatial Statistics* 2, 1–14.
- Webster, R. and M. A. Oliver (1992, March). Sample adequately to estimate variograms of soil properties. *Journal of Soil Science* 43(1), 177–192.
- Whittle, P. J. L., R. Stoklosa, S. Barrett, F. C. Jarrad, J. D. Majer, P. A. J. Martin, and K. Mengersen (2013, May). A method for designing complex biosecurity surveillance systems: detecting non-indigenous species of invertebrates on Barrow Island. *Diversity and Distributions* 19(5-6), 629–639.

- Widmer, A. M., L. L. Burckhardt, J. W. Kehmeier, E. J. Gonzales, C. N. Medley, and R. A. Valdez (2010, December). Detection and Population Estimation for Small-Bodied Fishes in a Sand-Bed River. *North American Journal of Fisheries Management* 30(6), 1553–1570.
- Williams, J. N., C. Seo, J. Thorne, J. K. Nelson, S. Erwin, J. M. O'Brien, and M. W. Schwartz (2009, July). Using species distribution models to predict new occurrences for rare plants. *Diversity and Distributions* 15(4), 565–576.
- Wright, D. H. (1991). Correlations Between Incidence and Abundance are Expected by Chance. *Journal of Biogeography* 18(4), 463–466.