The anthropogenic influences on the distribution of two orchid species in Xishuangbanna, China

**西双版纳**傣族自治州

ᦈᦹᧈᦈᦹᧈᦋᦵᦲᧁᦘᦱᦉᦱᦑᦺ᧑᧒ᦗᧃᦓᦱ

Samuel John Herniman

Being an honours project submitted to Bangor University in partial fulfilment of an honours degree in Applied Terrestrial and Marine Ecology

May 2016 500315403

# Declaration

I declare that this is the result of my own investigation and that it has not been submitted or accepted in whole or part for any degree, nor is it being submitted for any other degree.

Candidate: Samuel J Herniman

Signature: ……………………………….

# Acknowledgements

# Abstract

The tropical forests of Xishuangbanna are one of the most biodiverse regions of China. Within the last 20 years, 22% of the land has been converted to rubber plantation and tea resulting in large scale fragmentation and habitat loss. In addition, Xishuangbanna’s close proximity to Myanmar, Laos, and Vietnam means that it is a hub for wildlife trade, threatening many species in the region.

With over 400 orchid species of 115 genera, Xishuangbanna holds 31% of all Chinese orchids. Orchids are habitat specialists due to their specific pollinators and mycorrhizal associations, and are often vulnerable to many anthropogenic activities. Of these 400, 3 have been classified as possibly extinct in the wild, 15 Critically Endangered, 82 as Endangered and 124 as Vulnerable, regionally.

In this study we examine the distribution of two species that are thought to be threatened by both habitat loss and trade for horticultural use: *Luisia magniflora* and *Dendrobium thyrsiflorum*. Both species are classified as endangered in a regional red listing assessment yet little is known about their habitat requirements or distribution. Using existing presence data from digital herbarium records and ecological surveys in Xishuangbanna over the last five years, we produced habitat suitability models of these species using MAXENT. Environmental predictors such as distance to roads, water and settlements, as well as altitude, land cover, canopy cover, aspect, precipitation and daytime temperature were taken into account.

Our results show areas of high probability of presence and we identify environmental factors influencing the distribution of the species. We also highlight the areas of high conservation concern by modeling the future distribution of tea and rubber plantations with MAXENT. By overlaying the existing protected area network onto the distribution maps we can assess the efficacy of protected area coverage.

# Contents

[Declaration 1](#_Toc449029285)

[Acknowledgements 2](#_Toc449029286)

[Abstract 3](#_Toc449029287)

[Contents 5](#_Toc449029288)

[Tables 8](#_Toc449029289)

[Figures 8](#_Toc449029290)

[Introduction 10](#_Toc449029291)

[Orchidaceae 10](#_Toc449029292)

[**Diversity** of Orchids 10](#_Toc449029293)

[Use of orchids globally 10](#_Toc449029294)

[Orchids in Trade 11](#_Toc449029295)

[Sources of orchids 11](#_Toc449029296)

[Demand for orchids 11](#_Toc449029297)

[Horticulture or medicine? 11](#_Toc449029298)

[CITES 11](#_Toc449029299)

[Orchids in China 12](#_Toc449029300)

[Diversity 12](#_Toc449029301)

[Cultural significance 12](#_Toc449029302)

[Use in China 12](#_Toc449029303)

[Horticulture 12](#_Toc449029304)

[Food 12](#_Toc449029305)

[Medicine 12](#_Toc449029306)

[Threats to orchids in China 13](#_Toc449029307)

[Wild harvesting 13](#_Toc449029308)

[Deforestation 14](#_Toc449029309)

[Urbanization 14](#_Toc449029310)

[Cultivation 14](#_Toc449029311)

[Protected areas 14](#_Toc449029312)

[MAXENT 15](#_Toc449029313)

[Maximum Entropy 15](#_Toc449029314)

[MAXENT 17](#_Toc449029315)

[What can it be used for? 20](#_Toc449029316)

[Where it falls short… 21](#_Toc449029317)

[Conclusions 23](#_Toc449029318)

[ENM Tools 23](#_Toc449029319)

[Methods 24](#_Toc449029320)

[Description of the study site 24](#_Toc449029321)

[Fieldwork 26](#_Toc449029322)

[Data preparation 26](#_Toc449029323)

[Selection of species 26](#_Toc449029324)

[Presence points 27](#_Toc449029325)

[Environmental variables 27](#_Toc449029326)

[Habitat suitability modelling with MAXENT 32](#_Toc449029327)

[Model selection 32](#_Toc449029328)

[Final model 33](#_Toc449029329)

[Human-natural 33](#_Toc449029330)

[Analysis 33](#_Toc449029331)

[Results 34](#_Toc449029332)

[Orchid locations 34](#_Toc449029333)

[Environmental variables 34](#_Toc449029334)

[Model suitability 34](#_Toc449029335)

[Final models 34](#_Toc449029336)

[Final maps 35](#_Toc449029337)

[Discussion 36](#_Toc449029338)

[Conclusions 37](#_Toc449029339)

[References 38](#_Toc449029340)

[Appendices 39](#_Toc449029341)

Tables

[Table 1. The number and percent of species in a given conservation status in Xishuangbanna (Liu et al. 2015). 25](#_Toc449036737)

[Table 2. The wavelength and resolution of each band supplied by the OLI and TIRS. Bands 1 to 9 are supplied by OLI and Bands 10 and 11 are supplied by TIRS (NASA 2010). 28](#_Toc449036738)

[Table 3. The path, row, scene ID, and date of capture of each Landsat 8 OLI/TIRS tile. 28](#_Toc449036739)

[Table 4. The slope of each pixel with the corresponding reclassified value. 30](#_Toc449036740)

[Table 5. The aspect of each pixel with the original value and corresponding reclassified value. 31](#_Toc449036741)

[Table 6. Aspect-slope matrix showing all possible values within the aspect-slope grid. 31](#_Toc449036742)

[Table 7. The correlation between the number of factors used in each model and the test AUC, training AUC, and Δ (training AUC-test AUC). p-values indicate a significant positive correlation between the number of factors and test AUC and training AUC in Dendrobium thyrsiflorum models and a significant positive correlation between the number of factors and training AUC in Luisia magniflora models. 34](#_Toc449036743)

Figures

[**Figure 1.** Illustrates entropy vs information contained in the model. Where E is the greatest entropy value possible considering all the available information and I is the point where all available information is incorporated into the model. If a model contains more information than I it is making assumptions and if a model contains less information than I it is ignorantly not using all information available. The point of maximum entropy is where marginal information meets marginal uncertainty. 15](#_Toc448925004)

# Introduction

## Orchidaceae

### **Diversity** of Orchids

The Orchidaceae family is one of the largest families of flowering plants (sharing the top spot with the Asteraceae), composed of more than 25,000 species and 736 genera (Liu, Luo & Liu 2010; Joppa, Roberts & Pimm 2010; Xing *et al.* 2014; Zhang *et al.* 2015; Chase *et al.* 2015). Orchids have been found on all continents including Antarctic islands (Chen *et al.* 2014). Almost all orchid species rely on mycorrhizal fungi (known as rhizoctonias) in some or all of their life cycle (Rasmussen & Rasmussen 2009; Xing *et al.* 2014) and rely on specific pollinators (Zhang *et al.* 2015). Studies have shown that the distribution of orchids is affected on fine scales by soil moisture, light availability, and canopy size (Gravendeel *et al.* 2004; Huang *et al.* 2008; McCormick & Jacquemyn 2014; Zhang *et al.* 2015).

### Use of orchids globally

Horticulture

Medicine

Gastrodia elata, Dendrobium offcinale, Luisia discolor

Vanilla

## Orchids in Trade

The global orchid trade is vast and varied (Goh & Kavaljian 1989) supporting many livelihoods and supplying xxxx amount of the global horticultural and medicine trade.

### Sources of orchids

### Demand for orchids

Goh & Kavaljian (1989) outlined the factors determining the demand for orchids in the horticultural trade. These include: whether or not members of a country **consume** flowers regularly, consumer income, energy cost in production of the orchid, vaselife and quality, fashion, and predictions of the importers and distributors (Goh & Kavaljian 1989). **Demand for meds?**

### Horticulture or medicine?

### CITES

The primary mechanism in place to reduce wildlife trade is the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES). CITES has been successful in the case of intense farming and regulatory effort in combinations with the SARS epidemic of 2003 (Jiang *et al.* 2013), but its effectiveness has been called into question **in regards to…** (Phelps *et al.* 2010; Challender & MacMillan 2014; Nijman & Shepherd 2015). Trade bans can, in some cases, exacerbate overexploitation threats (Conrad 2012) and drive wildlife trade underground rather than reducing it (Veríssimo, Diogo; Challender, Daniel W.S. & Nijman 2012; Biggs *et al.* 2013).

While CITES makes it illegal to trade orchids across borders, it is not illegal to harvest orchids and trade them within China as long as they are not harvested from protected areas (Sang, Ma & Axmacher 2011). There is a need for research to determine whether the orchids bought and sold around the borders of China were harvested within China or one of its neighbors.

Where do orchids fall under CITES?

CITES lists all of the known wild Orchids (Tian, Chen & Xing 2013)

## Orchids in China

### Diversity

There are more than 1200 native orchid species and 173 genera of orchids in China (Liu, Luo & Liu 2010; Zhang *et al.* 2015). 35% of China’s orchids are endemic (Liu *et al.* 2014). Most orchid diversity occurs in the Southern areas of the country (Zhang *et al.* 2015)

### Cultural significance

Orchids have been highly admired in the Chinese culture since the time of Confucius (Goh & Kavaljian 1989).

### Use in China

Orchids are used horticulturally and medicinally in China

### Horticulture

### Food

G elata (Chen et al. 2014)

### Medicine

Orchids, especially dendrobiums, have been used in traditional Chinese medicine for more than 2 millennia (Liu *et al.* 2014). Roughly 25% (n=350) of the orchid species in China are used in Traditional Chinese Medicine, 27% (n=97) of which are endemics (Liu *et al.* 2014).

Examples: *Dendrobium officinale* had a large distribution throughout the south of China but is now rare in the wild (Ding *et al.* 2008) *G elata* (Chen *et al.* 2014)

It is tricky to regulate this use as some orchids do have medicinal uses (Ng *et al.* 2012; Chen *et al.* 2014). Although an argument can be made into sustainably harvesting these species so that medicines are available to future generations.

## Threats to orchids in China

The conservation of orchids is threatened by overharvesting and habitat destruction (Liu, Luo & Liu 2010; Chen *et al.* 2014). Orchids are sensitive to habitat disturbance due to their mycorrhizal and pollinator specificity (Zhang *et al.* 2015). There is an urgent need for protection (Liu, Luo & Liu 2010).

### Wild harvesting

Nearly all medicinal plants in China are wild harvested (López-Pujol, Zhang & Ge 2006; Sang, Ma & Axmacher 2011). This practice has threatened the existence of many orchid species including *Gastrodia elata* which underwent a near total population collapse in the wild in the 1960s (Zhang *et al.* 2005, 2015; Sang, Ma & Axmacher 2011; Chen *et al.* 2014).In the 1980s, harvested medicinal *Dendrobuim* volume was 600,000 kg annum-1. Harvests are now reduced due to the depletion of wild populations (Liu *et al.* 2014). Bans on collecting endangered species have been largely unenforced and subsequently ineffective (Sang, Ma & Axmacher 2011). This practice has now begun to make use of orchid populations in adjacent countries (Liu *et al.* 2014).

### Deforestation

Deforestation has been carried out in China for hundreds of years (Sang, Ma & Axmacher 2011). Consequently, the number of natural forests in China is highly reduced (Sang, Ma & Axmacher 2011). There is currently a ban on logging from natural forests, however, the impoverished rural areas of the country often take little heed to laws regarding protected areas (López-Pujol, Zhang & Ge 2006; Sang, Ma & Axmacher 2011).

### Urbanization

The Chinese government protects growth and urbanization has occurred rapidly in China (Sang, Ma & Axmacher 2011). This has led to an

### Cultivation

Very few native orchids in China have been cultivated on a large scale (Liu, Luo & Liu 2010; Chen *et al.* 2014). This is partly due to the mycorrhizal fungal associations many orchids rely on which makes the process of developing a cultivation method time consuming and expensive (Liu, Luo & Liu 2010; Xing *et al.* 2014) (citation needed). Even if cultivation were possible, Williams, Jones & Annewandter (2014) found that some cultivation programmes can increase the strain of wild harvesting. However there are successful examples of this technique (Liu, Luo & Liu 2010). Liu *et al.* (2014) demonstrated a method of planting orchids in natural forests and sustainably harvesting them.

### Protected areas

There were at least 2600 nature reserves in China in 2012. There was only one set up specifically to protect orchids (Liu *et al.* 2014), However 90% of China’s orchid species are found in nature reserves although this does not mean that they contain viable populations (Liu *et al.* 2014; Zhang *et al.* 2015).

Chinese authorities **struggle** to effectively manage protected areas (Zhou & Grumbine 2011) as little enforcement is carried out in these areas, they are largely underfunded, and rural human populations are often located in and around protected areas (Liu *et al.* 2014). It should also be noted that the selection of protected areas in China is often based on iconic species (especially pandas) and does not reflect habitats of conservation priority (Sang, Ma & Axmacher 2011).

While the presence of ineffective protected areas often causes people to feel that conservation progress is being made, the presence of a protected area is superior to no protected area at all (Zhou & Grumbine 2011)

The conservation status, economic value, and position within ecosystems make the Orchidaceae an ideal family to study (Zhang *et al.* 2015).

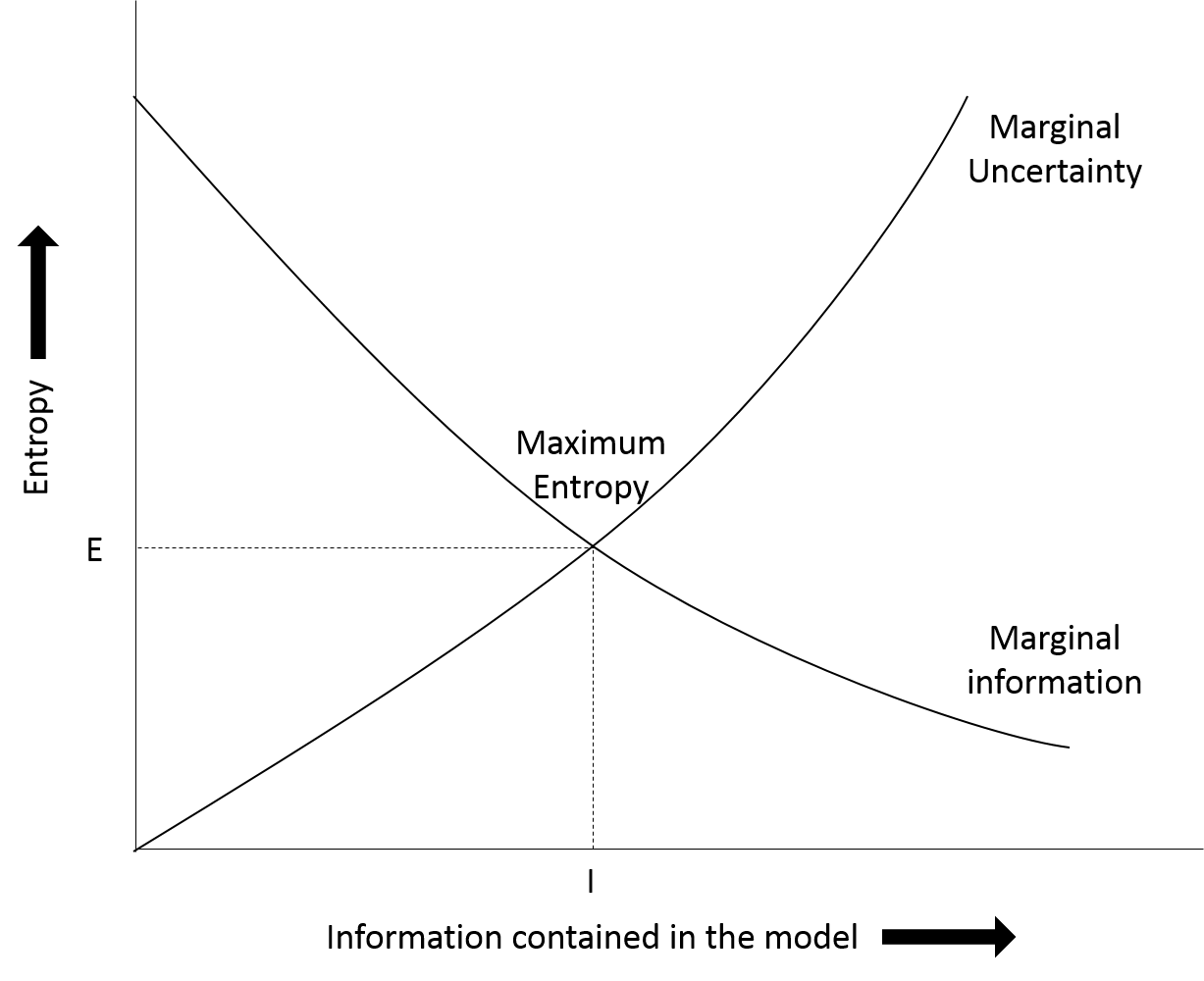
## MAXENT

### Maximum Entropy

To understand MAXENT we must first understand the concept of maximum entropy on which MAXENT is based.

Entropy is a measure of uncertainty or a lack of information (Penfield & Lloyd 2003). ˻When there is no information distinguishing one explanation over another, all the possible explanations can be given equal probabilities and entropy is at its greatest point (log *m* where *m* is the number of possible explanations)˼(Jaynes 1963). This is the principle of indifference (Conrad 2004). At its lowest, entropy has a value of 0, denoting that there is enough information to identify one explanation with no uncertainty over all other possibilities (Jaynes 1963; Penfield & Lloyd 2003; Phillips, Anderson & Schapire 2006).

The principle of maximum entropy is an elementary concept, although it is often explained in a counter-intuitive and convoluted manner: ˻the explanatory distribution (output) must take into account all of the provided data without making any assumptions not backed up by the data˼ (Lahoz-Monfort 2008). Essentially, when choosing between a number of explanations for a phenomenon, it is optimal to choose the one that maximises the entropy value while incorporating all the available information about the phenomenon, this is demonstrated in **Figure 1.** (Jaynes 1963; Conrad 2004). Therefore, the maximum entropy distribution is the most informed prediction one can make from the given information and the only prediction that can be made with reason. (Jaynes 1963; Penfield & Lloyd 2003; Conrad 2004).



**Figure 1.** Illustrates entropy vs information contained in the model. Where E is the greatest entropy value possible considering all the available information and I is the point where all available information is incorporated into the model. If a model contains more information than I it is making assumptions and if a model contains less information than I it is ignorantly not using all information available. The point of maximum entropy is where marginal information meets marginal uncertainty.

Thus, MAXENT produces a predicted species distribution over a landscape using all the information available without making any uninformed assumptions. A species is assumed to be present in an area unless there is information given which says otherwise. If no prior presence information was provided to MAXENT, a uniform species distribution would be produced – essentially the null hypothesis (Phillips, Dudík & Schapire 2004; Elith, Kearney & Phillips 2010).

### MAXENT

˻MaxEnt is a presence only, machine learning algorithm used to create habitat suitability maps based on the principle of maximum entropy (Phillips *et al.* 2009). MAXENT associates an input ofgeoreferenced presence points of a certain species with relevant environmental variables which have an influence on the distribution of a species (i.e. temperature, humidity, distance from rivers) to plot their predicted distributions across landscapes (Merow, Smith & Silander 2013; Särkinen, Gonzáles & Knapp 2013). These are referred to as *factors* in MAXENT (Lahoz-Monfort 2008). ˻When doing this, MAXENT combines the known presence points and environmental variables to determine what encompasses the species’ ecological niche, producing a potential distribution of the focus species˼ (Phillips, Dudík & Schapire 2004; Phillips, Anderson & Schapire 2006; Williams 2008; Elith, Kearney & Phillips 2010; Araújo & Peterson 2012).

**Receiver operating characteristic (ROC) analysis**

To tell whether a model produced by MAXENT is able to predict a presence point correctly, MAXENT conducts receiver operating characteristic (ROC) analysis (Williams 2008). A curve is produced with sensitivity (true positives) on the *y* axis and 1-specificity (false negatives) on the *x* axis (Phillips, Anderson & Schapire 2006). The area under the curve (AUC) is measured with a value of 1 denoting no incorrect predictions and a value of 0.5 denoting predictions that are no more accurate than random (Elith *et al.* 2006; Williams 2008).

**Presence Only Data**

Rather than looking at presence-absence data, MAXENT takes an input of presence-only data and converts it to presence-pseudoabsence data where the background areas (where no presence point is available) are considered to have an unknown presence – also known as a pseudoabsence (Phillips, Anderson & Schapire 2006; Phillips *et al.* 2009; Merow, Smith & Silander 2013). This presence-pseudoabsence data can then be used to determine the likelihood that an unknown space is occupied or not.

˻The documentation of accurate presence-absence data is relatively inefficient to carry out in most areas of the world. However, natural history museums and herbaria have a plethora of presence-only data˼ (Elith, Kearney & Phillips 2010; Merow, Smith & Silander 2013). Therefore, the use of presence only algorithms, like MAXENT, is realistic and relevant to conservation (Phillips, Dudík & Schapire 2004). MAXENT has gained a following due to its ability to produce accurate distributions based on very few (fewer than 100) presence only data points (Phillips, Dudík & Schapire 2004; Pearson *et al.* 2007; Merow, Smith & Silander 2013). However, Lozier, Aniello & Hickerson (2009) caution that one must ensure that the presence-only data used in the model are accurate. This presence-pseudoabsence data is combined with environmental variables to produce the MAXENT model.

**Environmental variables (features)**

Since MAXENT attempts to find patterns in the associations between the presence points and environmental variables, careful selection of which environmental variables should be included in the model is required. It is important to exclude variables which have little or no effect on the distribution of the species in question from a MAXENT model to reduce the likelihood of the algorithm finding associations which do not exist. Additionally it is wise to include all the environmental variables which do have an effect on the distribution to gain a complete understanding of the environmental envelopes the species reside in (Merow, Smith & Silander 2013; Hughes 2015).

It is important to note that when MAXENT is used for projecting the distribution of species into the future or studying environmental influences of a species, the feature classes (environmental variables) must be chosen carefully to create an accurate model (Merow, Smith & Silander 2013). However, when MAXENT is used for looking at the accuracy of presence points, MAXENT can determine the most appropriate features to include in the analysis using a machine learning algorithm and so it is possible to include all reasonable layers without affecting the output (Merow, Smith & Silander 2013). This process of determining which features are relevant or not is known as regularisation.

Regularisation

Regularisation is the process by which MAXENT removes the features which are not relevant to the model and maximises the use of the relevant ones (Phillips, Anderson & Schapire 2006; Hastie, Tibshirani & Friedman 2009; Merow, Smith & Silander 2013). The default regularisation settings are adequate for most species, however more accurate regularization may be achieved with more appropriate settings (Elith, Kearney & Phillips 2010; Merow, Smith & Silander 2013). A greater regularization value has the effect of removing more features from the analysis. This will create a simplified model but may lack influential features (Merow, Smith & Silander 2013).

### What can it be used for?

Once MAXENT produces a distribution model, it can be useful in many ecological applications. Among them:

Assessing the possible spread of invasive species

A distribution of the potential sites that an invasive species could move into is useful for conservationists who wish to take preventative measures to stop this spread from occurring (Araújo & Peterson 2012). It can also be useful to assess why an invasive species has not moved into a certain area when given enough time. This knowledge may lead researchers to a weakness of the invasive species.

Climate change outlooks for at risk species

It is common to use MAXENT and related algorithms to make predictions about how a species will be affected by climate change. MAXENT can be used to produce distributions that show where the species will no longer be able to inhabit once changes occur or it may show possible areas where the species may occupy in the future (Araújo & Peterson 2012). It is important to note that these predictions are only accurate assuming that the species makes no adaptations to climate change and is only capable of survival in the climates it already occupies.

Discovering new populations of a species

It is possible for researchers to use the known data about a species to determine where it may occur in addition to currently known populations and then visit these potential sites to confirm presence (Araújo & Peterson 2012).

Determining areas of conservation concern and the most appropriate reserve boundaries

Knowledge of where a species occurs when delimiting areas to conserve or reserve boundaries is very important since conservation funding is often limited. Making a reserve too large can be an unneeded financial burden while making a reserve too small can impact the survival of the species of concern (Araújo & Peterson 2012).

Assessing potential areas for reintroductions and translocations

MAXENT can determine new areas with suitable habitats for successful introductions. ˻As reintroductions are rarely successful˼, it is helpful to know with some certainty that there is a likelihood of success on a particular project (Araújo & Peterson 2012; Ewen, Soorae & Canessa 2014).

### Where it falls short…

Araújo and Peterson (2012) argue that while a clearly stated concept and purpose of using MAXENT can be informative and useful, it is not uncommon for the concept to be vague or unexplained leading to misinterpretation. It is simple to counteract these issues by carefully assessing the reasons for carrying out the modelling, assumptions, and what the potential results are prior to running MAXENT(Araújo & Peterson 2012).

Merow et al. (2013) claim that researchers are often unfamiliar with MAXENT to an appropriate level and rarely make informed decisions when choosing settings. Researchers generally choose the defaults while more appropriate settings (or even algorithms) are unknown (Merow, Smith & Silander 2013).

Sampling bias

Sampling bias is inherent in presence only data. It is very common for samplers to survey areas near roads and that are easily accessible. While sampling bias can be accounted for when the search effort is known, a process called Target Group Sampling (TGS) can be used to roughly account for sampling bias when search effort is unknown. TGS uses presence data from similar species to predict the presence of the target species over areas that have not been sampled (Phillips *et al.* 2009; Merow, Smith & Silander 2013).

Range

When setting up data before running MAXENT, it is important to choose the range of the analysis wisely. When MAXENT is run over a larger area than the known range of a species, MAXENT will identify suitable habitats for the species outside of the known range (Phillips *et al.* 2009; Webber *et al.* 2011; Barve *et al.* 2011; Merow, Smith & Silander 2013). Whereas running MAXENT on a smaller area than the known range of the species will not give a representative distribution of the suitable habitat of the species. These possibilities are adequate when carrying out appropriate analysis but it is important to consider the question being asked and the goals of the analysis before setting up MAXENT.

### Conclusions

The machine learning algorithm, MAXENT combines presence-only occurrence data of a particular species with relevant environmental layers (factors) to produce a habitat suitability model of the bioclimatic envelopes the species occupies or is capable of occupying. It is important to understand the limitations and capabilities of MAXENT before setting up the software and running the model and to carefully consider the methods and goals of a particular study. With a vast amount of presence-only species occurrence data available in herbaria and natural history museums, MAXENT is a strong tool in conservation.

## ENM Tools

ENM Tools is software that can be used to compare two or more species distribution models are more similar than expected by chance (Warren *et al.* 2008). This tool has been used to

Methods

## Description of the study site

Xishuangbanna Dai Autonomous Prefecture is an area of southern Yunnan, China (21°09’-22°36’N, 99°58’-101°50’E) taking up 19,220 square kilometres (0.2% of China’s total landmass) and supports between 12 and 18 percent of China’s flora (Shou-qing; Cao & Zhang 1997; Hongmao *et al.* 2002; Liu *et al.* 2015). Xishuangbanna is neighbors to both Myanmar and Laos (Zhu, Wang & Li 1998).

Xishuangbanna has a tropical monsoon climate, caused by the barrier to northern cold air formed by the Hengduan Mountains in the north of the region (Zhu, Wang & Li 1998; Zhu, Cao & Hu 2006). 80% of Xishuangbanna’s rainfall is between May and October and the temperature is 21.5˚C at 80% humidity on average (Hongmao *et al.* 2002). In the dry season, lingering fog, caused by the mountainous terrain (fluctuating between the altitudes of 430 and 2300 m), supplements the lack of precipitation (Zhu, Wang & Li 1998; Hongmao *et al.* 2002; Zhu, Cao & Hu 2006).

Yunnan is China’s most biologically and culturally rich area and Xishuangbanna is (Zhou & Grumbine 2011). Being in the transitional zone between South-East Asia, subtropical East Asia, the Sino-Japanese floristic region, and the Sino-Himalayan floristic region, Xishuangbanna is a biodiversity hotspot and China’s most biodiverse region (Cao & Zhang 1997; Zhu, Cao & Hu 2006; Li *et al.* 2008) Xishuangbanna is home to 13 recognized nationalities, most notably, the Dai people who make up a third of the population (Hongmao *et al.* 2002).

The rapidly growing rubber industry in the region has resulted in high deforestation (Li *et al.* 2008). In fact, over the last 20 years, 22% of the total land in Xishuangbanna has been converted to monoculture rubber plantations (Xu, Grumbine & Beckschäfer 2014; Liu *et al.* 2015). The expansion of rubber plantations is curtailed by little more than the 1000 m upper limit to rubber plants (gao paper, xu 2014). Tea gardens also use a considerable area of land in the region, however, they have been traditionally present and so have negligible effects on biodiversity at present (Liu *et al.* 2015)

**Orchids**

Xishuangbanna is an orchid richness hotspot (Zhang *et al.* 2015). Most orchid species are found above 1000 m (Liu *et al.* 2015).

A study by Liu *et al.* (2015) determined the conservation status of 410 orchids in Xishuangbanna these statuses are shown in table x

Table 1. The number and percent of species in a given conservation status in Xishuangbanna (Liu et al. 2015).

|  |  |  |
| --- | --- | --- |
| Status | Number of species | % studied species |
| Possibly extinct in the wild | 3 | 0.73% |
| Critically endangered | 15 | 3.66% |
| Endangered | 82 | 20.00% |
| Vulnerable | 124 | 30.24% |
| Least concern | 186 | 45.37% |

Many species of orchids from Xishuangbanna have been used in the horticultural and medicine sectors; collection from the wild has contributed to the decline in orchid populations in the region (Liu *et al.* 2015)

0 extinction policy (Liu *et al.* 2015)

## Fieldwork

**Field surveys**

Field surveys were conducted between April 2011 and December 2013. In total, 108 full days were spent in the field and a number of subsequent day trips were made. Surveys were planned to correspond with the flowering times and habitats of Xishuangbanna’s species to maximise the individuals found. For each individual found, location and species were recorded amongst other information. These data have been previously used by Liu *et al.* (2015). **The location of all individuals found is illustrated in fig**

**Market surveys**

Orchid experts from Xishuangbanna Tropical Botanical Garden (XTBG) were consulted in regards to the locations of orchid markets in Xishuangbanna. Market surveys were carried out in January 2015 visiting **X markets.** At each market the species for sale, number of individuals of each species, and price of each individual was recorded. Surveys were carried out by Han Chinese researchers from XTBG with oversight from afar by the author to ensure that the prices of orchids reflected those given to **local orchid enthusiasts.**

## Data preparation

### Selection of species

Lists of orchids found in field surveys and orchids found on market surveys were compared to determine the most commonly found species in the wild as well as found in markets. This step was taken to ensure that there would be efficient data to carry out MAXENT analysis and to ensure that these species were economically present.

***Dendrobium t***

**Picture**

***Luisia mag***

***Picture***

### Presence points

The locations collected in field surveys of each of the subject species were extracted and prepared in accordance with the requirements of MAXENT (Phillips, Anderson & Schapire 2006). X records were moved as they were in settlements and did not occur naturally.

### Environmental variables

**Selection**

**Remote sensing**

Night lights

Landsat 8

Landsat 8 (LS8) is a remote sensing satellite produced by Orbital ATK for the National Aeronautics and Space Administration (NASA) and the United States Geological Survey (USGS) (Orbital ATK 2015). LS8 was launched on 11/02/2013 18:02 UTC and commenced operations on 30/05/2013 (NASA 2013). LS8 has a sun-synchronous orbit and an altitude of 705 km which makes sure that the images taken by the satellite are never dark because they are always taken at the same time of day (NASA 2013; Packwood 2015). It takes LS8 98 minutes to orbit the earth and passes over the same point on earth every 16 days (NASA 2013). LS8 has two instruments, the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS), data collected by these instruments is made available in 8 bands which are defined in **table x**. LS8 images used in this research were obtained from the USGS (<http://earthexplorer.usgs.gov>) and distributed by the Land Processes Distributed Active Archive Center (LP DAAC), located at USGS/EROS, Sioux Falls, SD (<http://lpdaac.usgs.gov/>).

Table 2. The wavelength and resolution of each band supplied by the OLI and TIRS. Bands 1 to 9 are supplied by OLI and Bands 10 and 11 are supplied by TIRS (NASA 2010).

| Spectral band | Wavelength (µm) | Resolution (m) |
| --- | --- | --- |
| Band 1 - Coastal/Aerosol | 0.433-0.453 | 30 |
| Band 2 - Blue | 0.450-0.515 | 30 |
| Band 3 - Green | 0.525-0.600 | 30 |
| Band 4 - Red | 0.630-0.680 | 30 |
| Band 5 - Near Infrared | 0.845-0.885 | 30 |
| Band 6 - Short Wavelength Infrared | 1.560-1.660 | 30 |
| Band 7 - Short Wavelength Infrared | 2.100-2.300 | 30 |
| Band 8 - Panchromatic | 0.500-0.680 | 15 |
| Band 9 - Cirrus | 1.360-1.390 | 30 |
| Band 10 - Long Wavelength Infrared | 10.30-11.30 | 100 |
| Band 11 - Long Wavelength Infrared | 11.50-12.50 | 100 |

Six OLI/TIRS tiles were used in this study these are specified in **table x**.

Table 3. The path, row, scene ID, and date of capture of each Landsat 8 OLI/TIRS tile.

|  |  |  |  |
| --- | --- | --- | --- |
| Path | Row | Scene ID | Date |
| 129 | 45 | LC81290452015052LGN00 | 21/02/2015 |
| 130 | 44 | LC81300442015075LGN00 | 17/03/2015 |
| 130 | 45 | LC81300452015075LGN00 | 17/03/2015 |
| 131 | 44 | LC81310442015066LGN00 | 07/03/2015 |
| 131 | 45 | LC81310452015066LGN00 | 07/03/2015 |

All 5 tiles from each LS8 band used were stitched into one geoTIFF image with ERDAS IMAGINE 2015 and then clipped to the bounds of the Xishuangbanna border in ArcMap 10.3.1 (Hexagon Geospatial 2014; ESRI 2015). All work was done in the WGS1984 datum.

**Raw bands**

It was inferred that OLI bands 4, 5, and 6 would contain useful data in regards to the habitat preference of the two subject species (Lahoz-Monfort 2008). These bands were converted to ASCII files.

**Vegetation indices**

Vegetation indices were calculated using ArcMap 10.3.1 and the OLI bands. Each was calculated with the formulas described in **table x**.

Albedo

Raw bands

MSR

NDVI

RNDVI

Tasseled cap transformation

**DEM**

ASTER

**Aspect**

**Slope**

**Aspect-slope**

While most environmental variables work independently from one another, it is possible for the combination of two or more variables to have a significant combined effect while having an insignificant individual effect **(citation needed).**

Since aspect **(the *horizontal* direction a hillside is facing (revise))** and slope (the angle a hillside is at – almost the vertical direction a hillside faces) are related and essentially different dimensions of a three dimensional space, we combined them into an aspect-slope layer using the method described by Brewer & Marlow (1993).

This was done by using ArcGIS 10.3.1 to recalculate each layer to new values separately. The slope of each pixel was reclassified with the Reclassify tool as a percentage to correspond with the values dictated in **table x.** The aspect of each pixel was reclassified with the Reclassify tool to correspond to each of the values stipulated in **table x.**

Table 4. The slope of each pixel with the corresponding reclassified value.

|  |  |
| --- | --- |
| Slope (%) | Reclassified value |
| 0-5 | 10 |
| 6-20 | 20 |
| 21-40 | 30 |
| 41-maximum | 40 |

Table 5. The aspect of each pixel with the original value and corresponding reclassified value.

|  |  |  |
| --- | --- | --- |
| Aspect | Original value | Reclassified value |
| N | 0-22.5 | 1 |
| NE | 22.5-67.5 | 2 |
| E | 67.5-112.5 | 3 |
| SE | 112.5-157.5 | 4 |
| S | 157.5-202.5 | 5 |
| SW | 202.5-247.5 | 6 |
| W | 247.5-292.5 | 7 |
| NW | 292.5-337.5 | 8 |
| N | 337.5-360 | 1 |

By adding the value of each grid cell of the reclassified aspect grid to the corresponding value in the reclassified slope grid, using the Plus tool, a combined grid in which each pixel can describe both the aspect and slope of each pixel on the terrain was created. **Table x** gives a matrix for all possible values in the aspect-slope grid.

Table 6. Aspect-slope matrix showing all possible values within the aspect-slope grid.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Slope (%) |  | Aspect | | | | | | | |
|  | N | NE | E | SE | S | SW | W | NW |
| 0-5 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 6-20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| 21-40 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 |
| 41-maximum | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 |

**Distance to roads**

The distance of each pixel from the nearest road was calculated using a shapefile provided by OpenStreetMap (<https://www.openstreetmap.org/>).

**Direction of roads**

The direction from each pixel to the nearest road was calculated using a shapefile provided by OpenStreetMap (<https://www.openstreetmap.org/>).

**Resistance**

The cost distance tool in ArcMap was used to calculate a layer in which every pixel was given a value of the smallest accumulated travel cost to the nearest road. This layer was calculated using the slope layer described above and a shapefile provided by OpenStreetMap (<https://www.openstreetmap.org/>). This is essentially a layer which shows the relative amount of energy one must expel to get from the closest road to each pixel.

## Habitat suitability modelling with MAXENT

### Model selection

To determine the most accurate model possible a series of 63 test environmental niche models were made using MAXENT (version 3.3.3). The theory behind this software is explained in **section x.** All settings were set to their default values including maximum iterations (500), convergence threshold (0.00001), sample radius (0), default prevalence (0.5), regularization multiplier (1), background points (10000), and replicates (1) (Phillips 2008; Lahoz-Monfort 2008). A random test percentage of 25 was used in all test models. Each test model had a different combination of environmental variables detailed in **table x.** The models able to best predict suitable habitat for each species were selected using the test AUC value, which is generally preferred over the training AUC when comparing models since it uses independent data to test the model rather than data used to create the model (Phillips 2008; Lahoz-Monfort 2008).

### Final model

The test model of each species with the greatest test AUC value was selected used to create an additional model using the same environmental variables and a random test percentage of 0 to ensure that the model used as much data as possible in creating the model. This model produced jackknife tests and response curves to aid in determining the most informative environmental variables.

### Human-natural

## Analysis

# Results

## Orchid locations

## Environmental variables

Each of the environmental variables used in the creation of the test models and final models are displayed in **appendix x**.

## Model suitability

Models and their corresponding AUC values can be seen in **appendix x**

The correlation between the number of environmental variables (factors) and AUC values was tested using Pearson’s correlation coefficient (with the cor.test package in R). Δ was calculated as test AUC – training AUC. The results of this test are displayed in **table x.**

Table 7. The correlation between the number of factors used in each model and the test AUC, training AUC, and Δ (training AUC-test AUC). p-values indicate a significant positive correlation between the number of factors and test AUC and training AUC in Dendrobium thyrsiflorum models and a significant positive correlation between the number of factors and training AUC in Luisia magniflora models.

|  |  |  |  |
| --- | --- | --- | --- |
| *Dendrobium thyrsiflorum* | t | df | p |
| Number of factors - Test AUC | 3.1293 | 61 | <0.01 |
| Number of factors - Training AUC | 9.3065 | 61 | <0.01 |
| Number of factors - Δ | 0.62854 | 60 | 0.532 |
|  | | | |
| *Luisia magniflora* | t | df | p |
| Number of factors - Test AUC | 1.815 | 61 | 0.07444 |
| Number of factors - Training AUC | 13.876 | 61 | <0.01 |
| Number of factors - Δ | 1.2369 | 60 | 0.2209 |

This test shows that the greater number of environmental variables used in the model, the greater the AUC values generally are and the more predictive the model can be. This explains why both final models utilize 18 and 19 of the available 20 environmental variables.

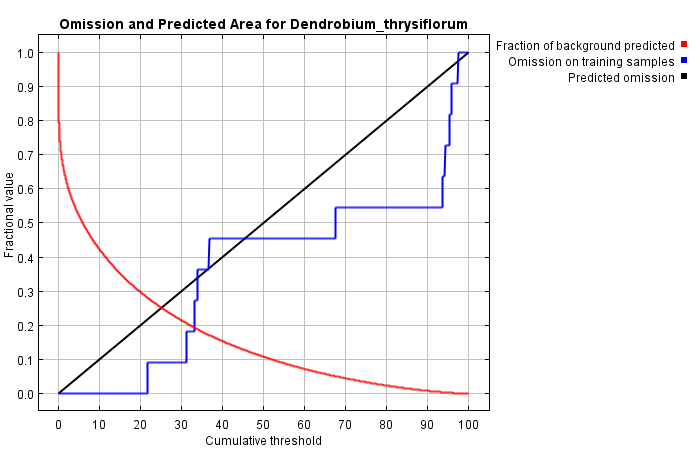
The highest test AUC for *Dendrobium thyrsiflorum* and *Luisia magniflora* were 0.8551 (model 37) and 0.9179 (model 35), respectively. The environmental variables used in the final models for each species are listed in **table x.**

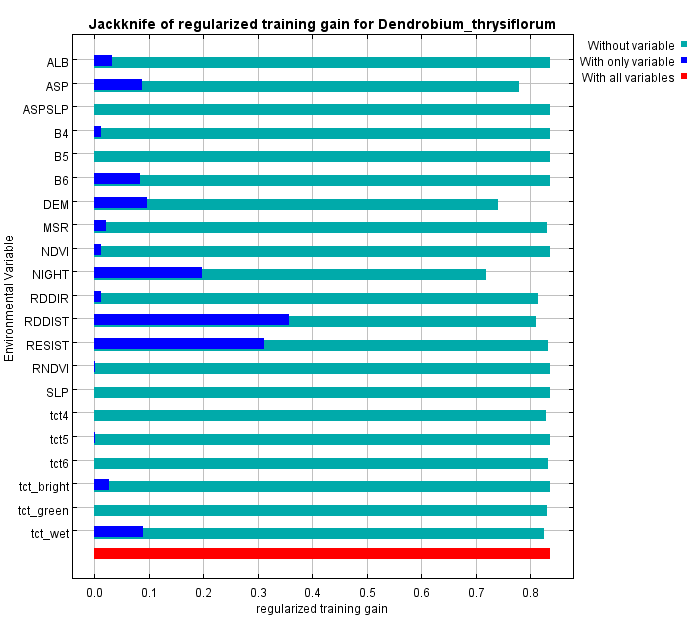
## Final models

The test models of each species with the greatest AUC values were run with all locations used and no random tests. The final AUC values are shown in **table x.**

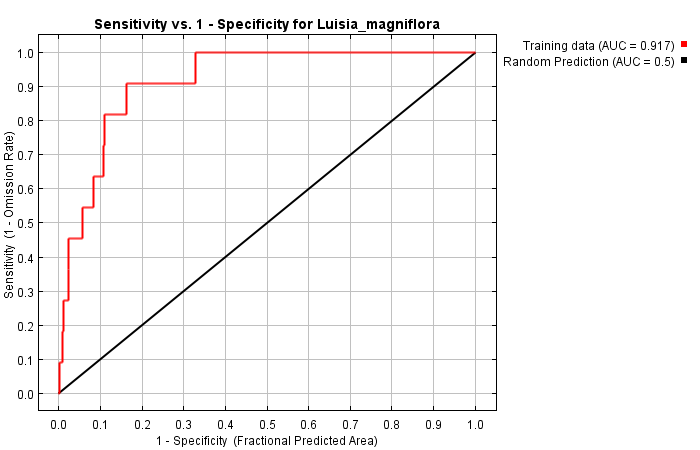
### *Dendrobium thyrsiflorum*

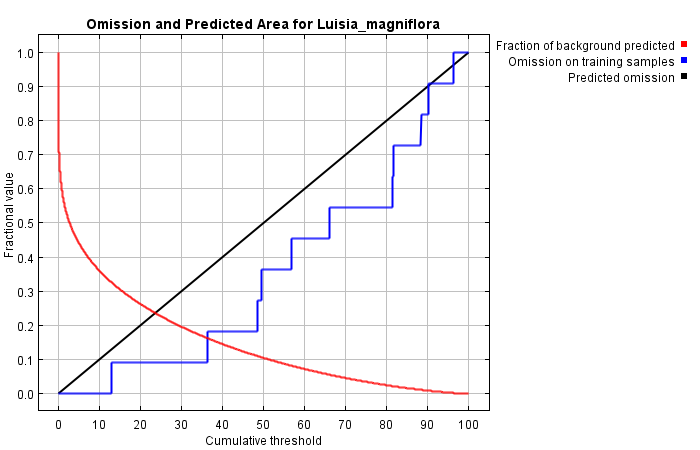


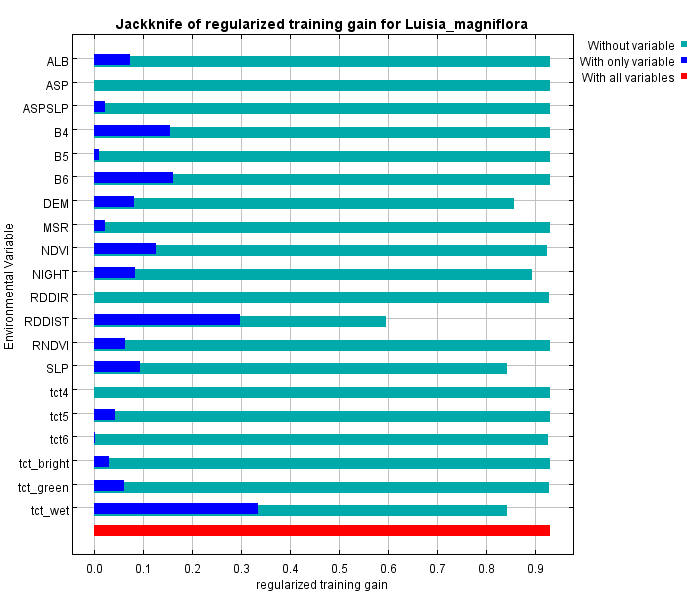




### *Luisia magniflora*



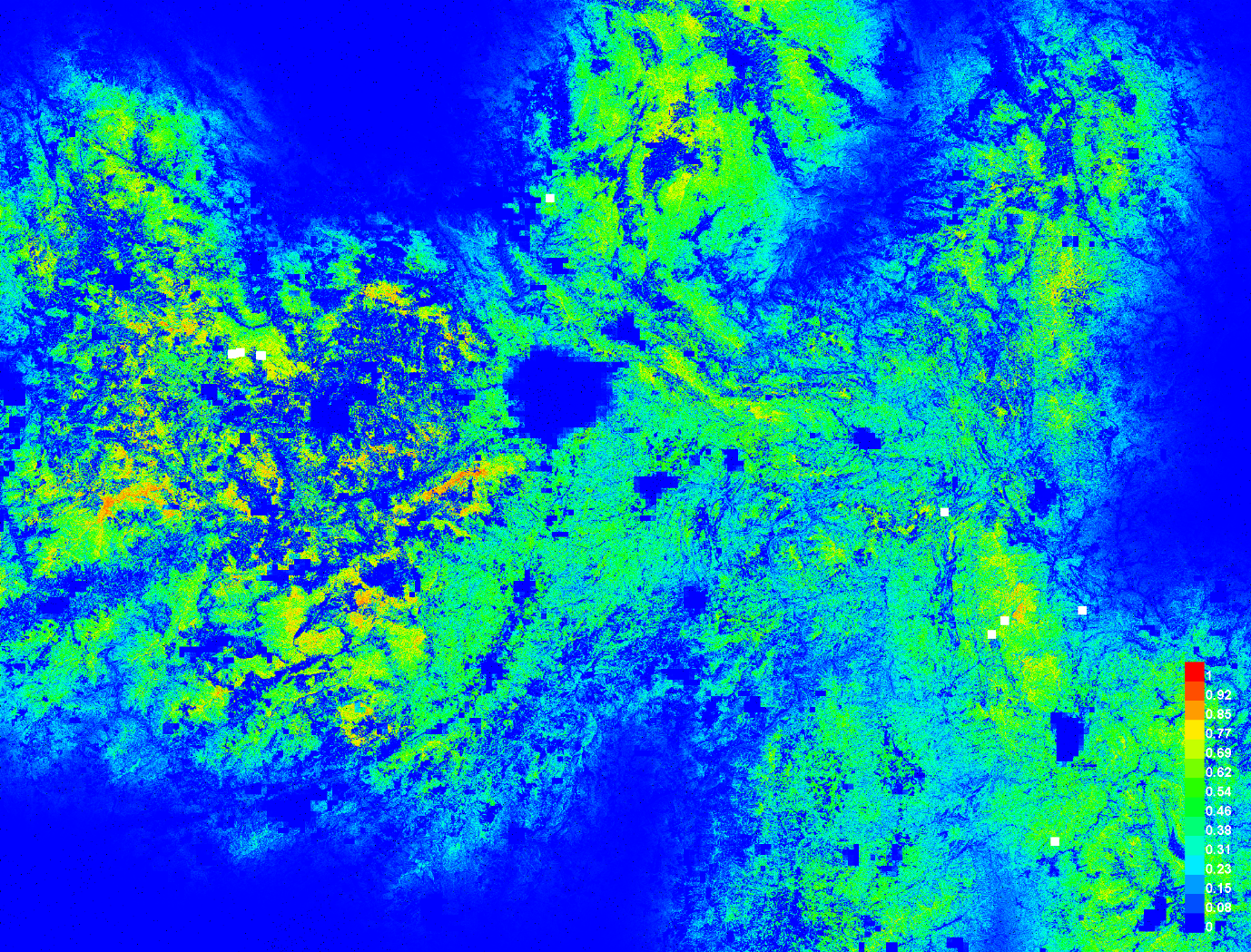




Roc analysis

## Final maps





# Discussion

Climate change rubber

# Conclusions

# References

# Appendices