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

西双版纳傣族自治州

[illegible]

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

Firstly, I must thank, my supervisors, Dr Sophie Williams and Andrew Packwood for teaching me an incredible amount about plants, people and computers. Thank you, Sophie, for introducing me to the wonders of research, taking risks, and the world. Without your mentoring, vigor, and foresight, this project would not have been possible and I would be a very different person. Andrew Packwood, thank you for your incredible guidance, enthusiasm and patience with me and for changing my mind about maps. Thanks to Prof Julia Jones and Dr Robert Annewandter for excellent feedback and for showing me what heroes look like.

Thanks to Prof. Chen Jin, Prof. Richard T. Corlett, and Prof. Gao JiangYun for allowing me to be a part of the XTBG family and to Nigel Brown for making me feel part of his.

This work was a part of the Gardd Dwy Ddraig / Two Dragons Garden project which is funded by the British Council and supported by Xishuangbanna Tropical Botanical Garden, the Chinese Union of Botanical gardens, the Confucius Institute, the Friends of Treborth Botanic Garden, Bangor University, the Royal Botanic Garden Edinburgh, and Botanic Gardens Conservation International.

Thanks to Dr Alice C Hughes, Liu Qiang, and Lu Zhilong for providing me with data, Neelam Notay for a photo and a lovely moussaka, and to Dr Kyle Tomlinson, Dr Christos Mammides, Hannah McFadyen, and Menno van Berkel for being fabulous friends and giving me a place to sleep at night. Thanks to the laowai for showing me how to party.

Thank you to the Friends of Treborth for giving me a home away from home, keeping me busy, and cheering me up when required.

Thanks to Wendy Herniman, John Herniman, Andrea Abegg, Nicole Barden and Faith Jones for editorial comments and for making me who I am today.

Finally, I would like to thank my imaginary friend, Penelope, for sticking by my side through it all.

This is for the freestyler.

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 species, 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. We examine the distribution of *Luisia magniflora* and *Dendrobium thyrsiflorum*, two species that have been classified as endangered in a regional red listing assessment. Little is known about their habitat requirements or distribution. Using existing presence data from surveys in Xishuangbanna and environmental variables, we produce habitat suitability models of these species using MAXENT. Jackknife tests highlight the distributional influence of two anthropogenic environmental variables on these species, lights at night (a proxy for settlement locations) and distance to roads (a proxy for wild harvesting ability), suggesting that these species are disturbed by wild harvesting. Jackknife tests also highlight the preference of both species for certain moisture levels, and *Dendrobium thyrsiflorum* for certain altitudes. We highlight that a multifaceted approach to reducing wild harvesting pressure and conserving these species is needed and demonstrate the robustness of MAXENT.

Contents

Tables

Figures

1.0 Introduction

The conservation of orchids within Xishuangbanna and the world is of high interest to researchers, policy makers, and the public due to orchids' high economic and cultural significance (Goh & Kavaljian 1989; Ding *et al.* 2008; Chen *et al.* 2014). Pimm & Raven (2000) showed that species extinctions are happening far more rapidly than is possible by the natural background extinction rate. All orchid species are listed under the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) and are vulnerable to human activities due to their high specificity in habitat as well as highly specific fungal associations (Rasmussen & Rasmussen 2009). The Xishuangbanna Tropical Botanical Garden, a research institution of the Chinese Academy of Sciences, has declared a zero extinction policy, and the extinction of an orchid species in Xishuangbanna would be catastrophic to this declaration (Liu *et al.* 2015).

Prior research has shown that the wild harvesting of orchids in China and beyond can drastically impact their populations and threaten them with extinction (Chen *et al.* 2014). Habitat suitability models have previously been used to inform researchers where protected areas should be placed to conserve both animals and plants (Williams 2008; Lahoz-Monfort 2008). Pearson *et al.* (2007) have demonstrated that it is possible to create accurate suitability maps with very little information as to where the species are found. However, existing research has not yet investigated the distribution of orchids in Xishuangbanna with habitat suitability models or investigated the ability for habitat suitability models to determine human disturbance in Xishuangbanna.

This study creates habitat suitability models of two epiphytic orchid species in Xishuangbanna, *Dendrobium thyrsiflorum* and *Luisia magniflora* and subsequently investigates the following aims

1. What factors determine the distribution of these orchids?
2. Do humans affect the distribution of these orchids?
3. Is it possible to create informative models with a small number of highly selected factors?

The next section gives an overview of the orchid family and the threats that orchids face in the world and in China, and this is followed by an overview of the habitat suitability model used in this work, MAXENT. Following this background information, the method is discussed, results are reported, findings are explained and finally conclusions are made.

2.0 Background

The following section gives an academic overview of orchid studies to date and the importance and threats to orchids and their economies. This is followed by an overview of the theory behind MAXENT, an algorithm for habitat suitability models.

2.1 Orchidaceae

This section will describe why the conservation status, economic value, and position within ecosystems of orchids make this an ideal family to study.

2.1.1 Global orchids

The Orchidaceae family is one of the largest families of flowering plants, 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 the 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).

The global orchid trade is vast and varied (Goh & Kavaljian 1989) Globally, orchids are used for many things by humans: the horticulture industry has made use of orchids for decoration, many medicines are derived from the plants, and orchids can be consumed for food (notably, vanilla is an orchid product) (Goh & Kavaljian 1989; Porras-Alfaro & Bayman 2007; Chen *et al.* 2014).

Goh & Kavaljian (1989) outlined the factors determining the demand for orchids in the horticultural trade. These include: whether or not members of a country purchase 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).

2.1.2 Orchids in China

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 these orchids are endemic (Liu *et al.* 2014). Most orchid diversity occurs in the southern areas of the country (Zhang *et al.* 2015). Orchids have been highly regarded in Chinese culture since the time of Confucius (Goh & Kavaljian 1989) and have been used horticulturally, medicinally and for consumption for two millennia (Liu *et al.* 2014; Chen *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).

2.1.3 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). Liu, Luo & Liu (2010) have declared an urgent need for protection.

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.* 2005a, 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). *Dendrobium officinale* had a large distribution throughout the south of China but is now rare in the wild (Ding *et al.* 2008). The overexploitation of this species has substantially contributed to this situation. 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). It is tricky to regulate this collection as some orchids have legitimate medicinal properties, although an argument can be made into sustainably harvesting these species so that medicines are available to future generations (Ng *et al.* 2012; Chen *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).

Urbanisation

The Chinese government prioritizes economic and urban growth (Sang, Ma & Axmacher 2011). This has led to extensive urbanization throughout the country. With rapid growth, little care has been made to manage deforestation and the subsequent impacts on threatened populations. Increased urbanisation has also improved access to forests that were previously inaccessible, enhancing opportunities for wild harvesting.

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). 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 *et al.* (2014) demonstrated a method of planting orchids in natural forests and sustainably harvesting them.

2.1.4 Protected areas

There were at least 2600 nature reserves in China in 2012 and 90% of China's orchid species are found in nature reserves, although this does not mean that they contain viable populations and only one of these protected areas was set up specifically to protect orchids (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 adjacent to 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 protected areas, no matter how ineffective, often causes people to feel that conservation progress is being made, the presence of any protected area is superior to no protected area at all (Zhou & Grumbine 2011).

2.1.5 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) which limits the unsustainable trade of plants and animals around the world (CITES 1973).

CITES lists all of the known wild Orchids under Appendix II except those from the genera *Paphiopedilum* and *Phragmipedium* and 6 specific species which are under Appendix I (CITES 1973; Tian, Chen & Xing 2013). Appendix II lists species which may face extinction if trade of them is not curtailed. Therefore individuals of appendix II species must be issued with an export permit, which states that the individual had been obtained legally and that the export does not pose a risk to the survival of the species, when traveling across borders (CITES 1973).

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). Additionally, trade bans can, in some cases, exacerbate overexploitation threats (Conrad 2012) and drive wildlife trade underground rather than reducing it (Verissimo, Challender & Nijman 2012; Biggs *et al.* 2013). There is a need for research to determine whether the orchids bought and sold around the borders of China are harvested from within China or its neighbors and to determine if trade bans are effective in Southeast Asia.

2.2 MAXENT

The following section gives a brief overview into the theory behind MAXENT, the habitat suitability model used in this study.

2.2.1 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 explains that 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 selection of explanations for a phenomenon, it is optimal to choose the one that maximises the entropy value (uncertainty) 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).

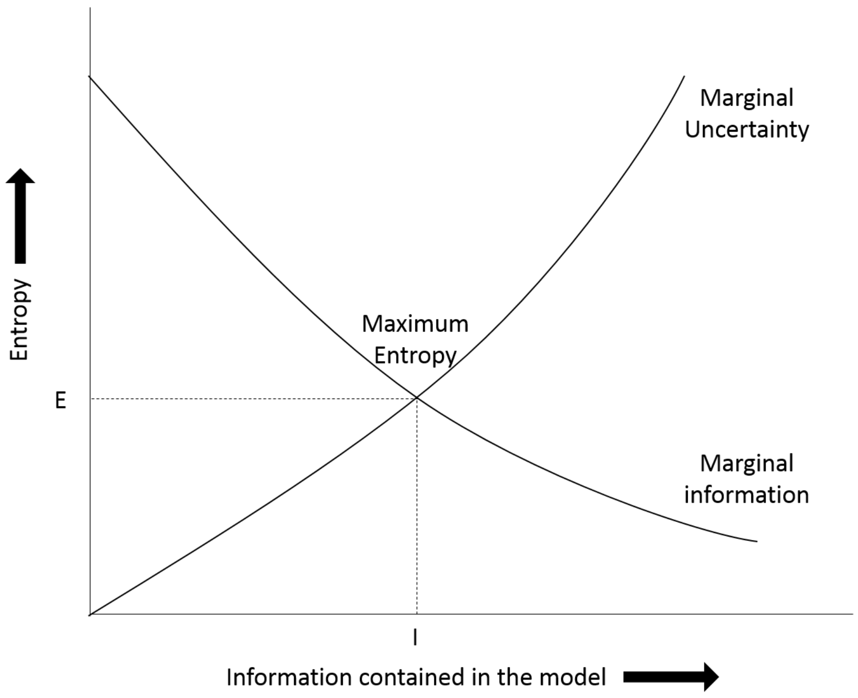

Figure 1. Illustrates entropy vs information contained in the model. Where E is the greatest entropy value possible considering all the available information and I is the point where all available information is incorporated into the model. If a model contains more information than I it is making assumptions and if a model contains less information than I it is 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).

2.2.2 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 of georeferenced presence points of a certain species with relevant environmental variables which have an influence on the distribution of a species (e.g. 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 (Phillips *et al.* 2009). 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, herbaria, and citizen science projects have a plethora of presence-only data (Elith, Kearney & Phillips 2010; Merow, Smith & Silander 2013; Matheson 2014). 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).

2.2.3 What can MAXENT be used for?

Once MAXENT produces a distribution model, it can be useful in many ecological applications. These include but are not limited to assessing the possible spread of invasive species, climate change outlooks for at risk species, discovering new populations of a species, and determining areas of conservation concern and the most appropriate reserve boundaries. These are discussed below:

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

MAXENT and other related algorithms are commonly used 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. This is not always a sensible assumption.

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 drawing 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).

2.2.4 Limitations of MAXENT

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. These issues can be counteracted 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 analyses but it is important to consider the question being asked and the goals of the analysis before setting up MAXENT.

2.2.5 Section summary

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.

3.0 Methods

3.1 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). Xishaungbanna's location within China is illustrated in `figure 2`. Xishuangbanna is neighbors to both Myanmar and Laos, this is shown in `figure 3` (Zhu, Wang & Li 1998).

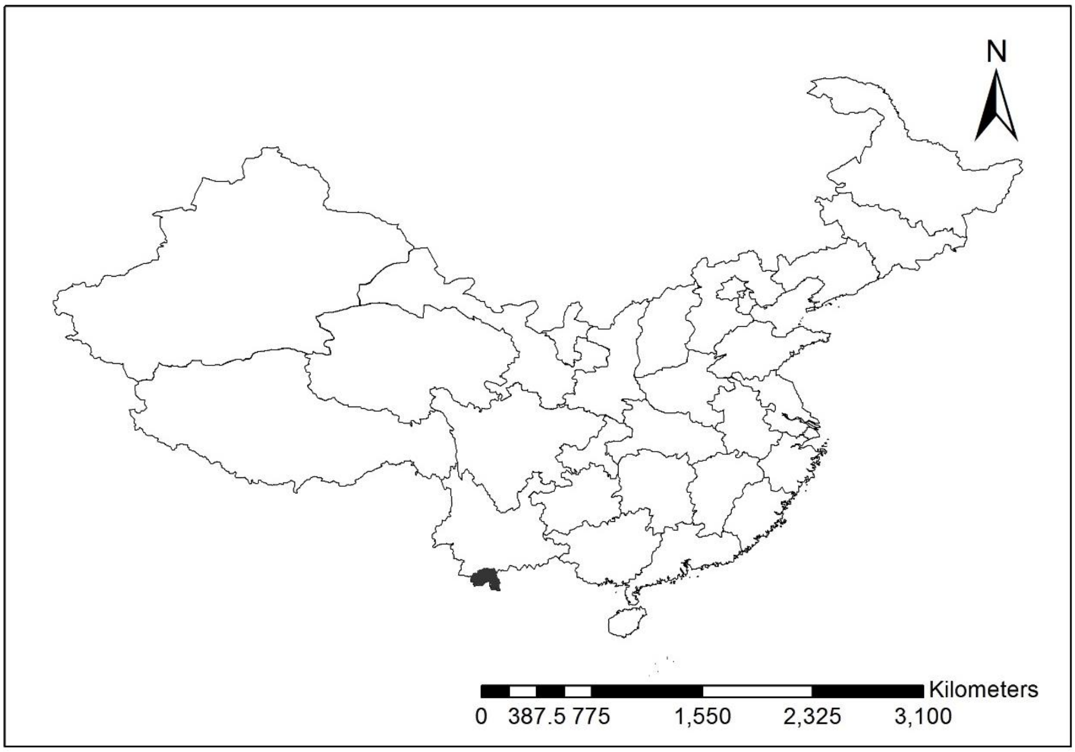

Figure 2. The location of Xishuangbanna within China. Xishuangbanna is shown in black. Data source: Esri, DeLorme Publishing Company, Inc. and Openstreetmap.

Xishuangbanna has a tropical monsoon climate, caused by the Hengduan Mountains in the north of the region which are a barrier to northern cold winds (Zhu, Wang & Li 1998; Zhu, Cao & Hu 2006). 80% of Xishuangbanna's rainfall occurs 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, supplements the lack of precipitation (Zhu, Wang & Li 1998; Hongmao *et al.* 2002; Zhu, Cao & Hu 2006).

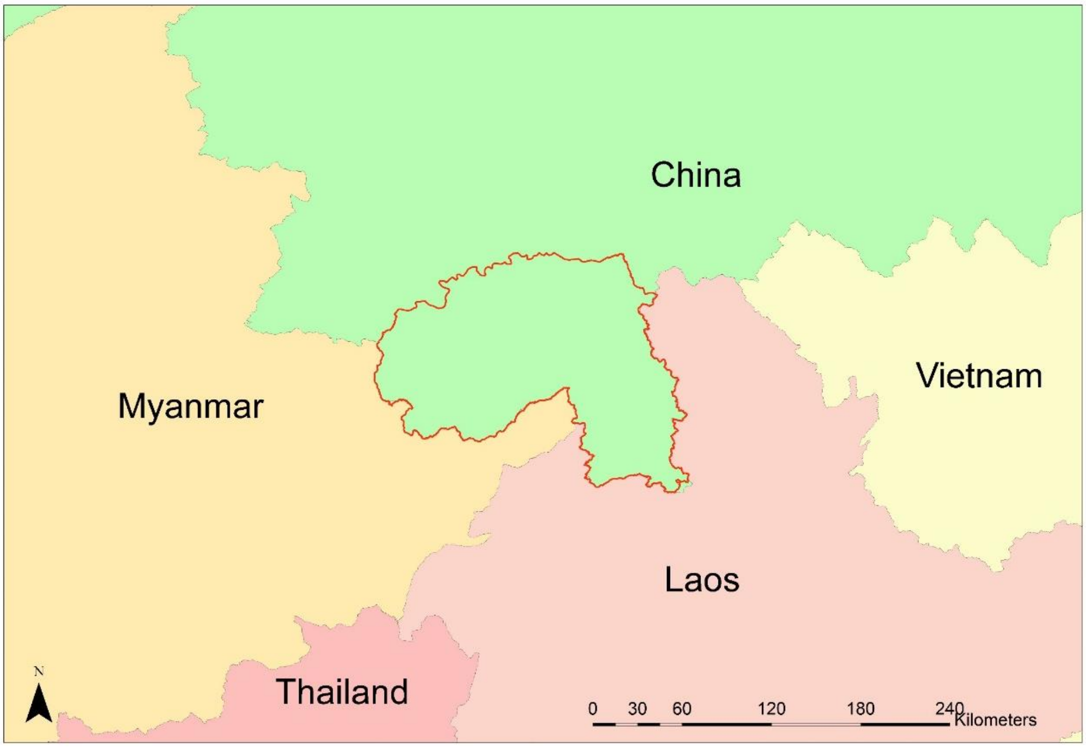

Figure 3. The location of Xishuangbanna (shown in red) in relation to the surrounding countries. The close proximity to Myanmar and Laos makes Xishuangbanna a hub of wildlife trade. Data source: Esri, DeLorme Publishing Company, Inc. and Openstreetmap.

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 in China's most biodiverse province, Yunnan (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 Xishuangbanna 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 altitude limit for rubber plants (Xu, Grumbine & Beckschäfer 2014; Liu *et al.* 2015). Tea gardens also use a considerable area of high altitude land in the region, however, orchids can be found in tea gardens and researchers are studying the possibility of planting orchids in tea gardens (Liu *et al.* 2015).

Orchids

Xishuangbanna is an orchid richness hotspot (Zhang *et al.* 2015). Most orchid species are found above 1000 m, the current ceiling to rubber habitat (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 1.

Table 1. The number and percent of species with 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)

3.2 Fieldwork

3.2.1 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 had previously been used by Liu *et al.* (2015).

3.2.2 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 9 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 rather than tourists. Figure 4 shows a typical orchid market in Xishuangbanna.

Figure 4. A typical orchid market on the side of a road in Xishuangbanna. Items for sale include dendrobiums in bundles for medicine, various intact horticultural orchids, papaya, bananas, and sugarcane. Taken in January 2015 by the author.

3.3 Data preparation

3.3.1 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 sufficient data to carry out MAXENT analysis and to ensure that these species were economically present. Each species is discussed below.

Dendrobium thyrsiflorum

Dendrobium thyrsiflorum is widely used in the horticulture industry and is not listed in the China Pharmacopoeia but it has been used as a substitute for Herba Dendrobii in Chinese traditional medicine and there is evidence that it has medicinal properties (Zhang *et al.* 2005b). *Dendrobium thyrsiflorum* is an epiphytic species of orchid which can be found throughout Southeast Asia (Yuan *et al.* 2011). It is commonly found in markets in Xishuangbanna. Figure 5 shows a botanical painting of the species and figure 6 shows a photo of it. In Chinese it is known as 球花石斛 qiu hua shi hu (eFloras 2016a).

Figure 5. Botanical painting of Dendrobium thyrsiflorum by Walter Hood Fitch in 1869. Public domain.

Figure 6.The raceme of *Dendrobium thyrsiflorum*. Source: Neelam Notay taken at Kew in 2016.

Luisia magniflora

Luisia magniflora is an epiphytic orchid of the *Luisia* genus. Little is known about the specific species, however the genus can be found throughout Southeast Asia and the Pacific islands (eFloras 2016b). *Luisias* are known as 钗子股属 chai zi gu shu in Chinese (eFloras 2016b). Figure 7 shows a botanical painting of *Luisia psyche* a closely related orchid to *Luisia magniflora.* They have similar morphologies.

Figure 7. Botanical illustration of *Luisia psyche* by Walter Hood Fitch in 1866. Public domain.

3.3.2 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). 9 records of *Dendrobium thyrsiflorum* and 6 records of *Luisia magniflora* were removed as they were in settlements and did not occur naturally.

3.3.3 Environmental variables

The following section explains the environmental variables used in this study and how they were prepared. Since little is known about these orchids, this study used all the easily accessible environmental variables at disposal. Each environmental variable was given an acronym (shown in table 7) when used as a factor in MAXENT.

Night lights

The lights at night layer was identified as a proxy for the location of settlements. Although it is possible for some settlements to have no available lighting, this is regarded as an acceptable measure of human presence (Small, Pozzi & Elvidge 2005). It was the only categorical factor in the study. This dataset comes from the 'Version 4 DMSP-OLS Nighttime Lights Time Series' (2013). If this information is important in the habitat suitability models but distance to roads (explained below) is not, it is possible to determine that humans are disturbing the orchid populations but they are not affected by wild harvesting.

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 11 bands which are defined in table 2. Each band is sensitive to a different wavelength of light and can be used to measure different characteristics of the earth's surface and air. 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 30m resolution OLI/TIRS tiles were used in this study these are specified in table 3.

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 (described in `table 2`) would contain useful data in regards to the habitat preference of the two subject species with little processing (Lahoz-Monfort 2008; Packwood 2015). These bands were converted to ASCII files and prepared for MAXENT.

Vegetation indices

Vegetation indices are environmental variables calculated from different bands of remote sensing data (Roujean & Breon 1995). They allow estimates to be made about the vegetation composition of an area (Roujean & Breon 1995).

Vegetation indices were used due to their ready availability and ease of use. Lahoz-Monfort (2008) has successfully used them to create habitat suitability models in the past, but it is unknown if they will contribute useful information to these species. Further information about each of these vegetation indices can be found in Roujean &

Breon (1995), Gong *et al.* (2003) and Lahoz-Monfort (2008). Vegetation indices were calculated using ArcMap 10.3.1 and the OLI bands. Each was calculated with the formulas described below using the bands from LS8 described in `table 2`.

- Albedo is a vegetation index of the amount of reflection a surface gives off. Understandably it is the combination of LS8 bands 1 through 7 which contain most of the reflected light from the earth's surface.

$$B1+B2+B3+B4+B5+B6+B7$$

- Modified Simple Ratio (MSR) was proposed by Chen (1996) as a modified version of several vegetation indices.

$$\frac{\left(B5/(B4-1)\right)}{\sqrt{(B5/B4)}+1} = MSR$$

- Normalized Difference Vegetation Index (NDVI) measures the amount of green vegetation in an area (Nouri *et al.* 2014).

$$\frac{(B5-B4)}{(B5+B4)} = NDVI$$

- Renormalized difference vegetation index (RNDVI or RDVI) is a modified vegetation index like the MSR. It linearizes the relationships between the biophysical elements of a surface and the vegetation index (Roujean & Breon 1995).

$$\frac{(B5-B4)}{\sqrt{(B5+B4)}} = RNDVI$$

- Tasseled cap transformation - Six environmental variables were created from the tasseled cap transformation. Each was calculated according to Baig *et al.* (2014) in which BX was replaced with the value stipulated.

$$(BX * B2) + (BX * B3) + (BX * B4) + (BX * B5) + (BX * B6) + (BX * B7)$$

Digital elevation model (DEM)

A digital elevation model of Xishuangbanna was provided by the ASTER project of NASA JPL (2009) and METI at 30m resolution. This environmental layer is used to investigate the altitude preference of each of the species and to calculate the aspect and slope layers below.

Aspect

Aspect is the direction on a compass that a slope faces. This was calculated from the DEM layer in ArcGIS. The aspect of a slope has been known to strongly influence the temperature of the microclimate due to the difference in the amount of sunlight each slope of a mountain or hill may receive (Bennie *et al.* 2006).

Slope

Similarly to aspect, slope is the gradient a hillside is set at. This was calculated from the DEM layer in ArcGIS. The slope can also affect the sun exposure on a hillside and therefore the microclimate (Bennie *et al.* 2006).

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.

Since aspect (the horizontal direction a hillside is facing) and slope (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 4. The aspect of each pixel was reclassified with the Reclassify tool to correspond to each of the values stipulated in table 5.

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 6 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. I.e. a north facing grid cell with a 1% slope would have a value of 11 and a west facing grid cell with a 7% slope would have a value of 27.

		Aspect							
		N	NE	E	SE	S	SW	W	NW
Slope (%)	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/). This factor was used as a measure of the ability of a wild harvester to get to a certain area to collect orchids and as a measure of human presence. The demonstration of this being an informative predictor of orchid distribution would show that wild harvesting has a negative effect on the distribution of these orchids.

Direction of roads

The direction from each pixel to the nearest road was calculated using a shapefile provided by OpenStreetMap (https://www.openstreetmap.org/). It may be that prevailing winds in the region blow exhaust from the roads into orchid habitats and therefore influence the distribution or roads in the region may be constructed on certain sides of valleys which may affect one's ability to travel into the forest as well as one's choice as to what direction to go in.

Resistance

The cost distance tool in ArcGIS 10.3.1 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.

Table 7. The acronyms and descriptions of each of the layers used as factors in MAXENT.

Acronym	Description of layer
ALB	Albedo
DEM	Digital elevation model
ASP	Aspect
ASPSLP	Aspect - slope combination
B4	Landsat band 4
B5	Landsat band 5
B6	Landsat band 6
RDDIST	Distance to roads
RDDIR	Direction of roads
MSR	Modified simple ratio
NDVI	Normalized difference vegetation index
NIGHT	Lights at night
RESIST	Resistance
RNDVI	Renormalized difference vegetation index
SLP	Slope
tct4	Tassled cap transformation 4
tct5	Tassled cap transformation 5
tct6	Tassled cap transformation 6
tct_bright	Tassled cap transformation for brightness
tct_green	Tassled cap transformation for greenness
tct_wet	Tassled cap transformation for wetness

3.4 Habitat suitability modelling with MAXENT

3.4.1 Model selection

To determine the most accurate model 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 2.2 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 appendix 3. 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).

3.4.2 Final models

The test model of each species with the greatest test AUC value was 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 influenced factors

The final model was calculated a second time with the factors that are directly influenced by the presence of humans (distance to roads, direction of roads, resistance, and night lights) removed. We shall call these factors *anthropogenic* factors and the factors that are not directly influenced by the presence of humans as *natural* factors. This was done to look at what factors would affect the distribution of each species in a habitat with little human disturbance. It should be noted that while the *anthropogenic* factors are always directly influenced by humans, the *natural* factors may also be directly influenced by humans. For example, the distance to roads of a grid square can only be changed by the construction of a road, but the albedo of a grid square will be changed if a road is placed there but will also be affected by the vegetation of the area.

3.5 Efficiency models

Finally, an investigation into the ability of MAXENT to produce a model with a very low number of environmental variables was created. This was carried out because the main expenditure in terms of time when creating habitat suitability models is the process of preparing factors for use in MAXENT. The identification of a small number of factors that predict the distribution of *Dendrobium thyrsiflorum* and *Luisia magniflora* would increase efficiency when modeling their distributions in other habitats. This investigation used MAXENT to model the distributions of both species using only the four most used factors in each of the final models (35A and 37A) in terms of percent contribution.

4.0 Results

The following section displays all relevant results to the study.

4.1 Orchid locations

14 *Dendrobium thyrsiflorum* and 16 *Luisia magniflora* locations were used in this study. Their locations have been omitted from this document to protect the locations of these orchid populations, although the data is available from Liu *et al.* (2015) on request.

4.2 Model suitability

Models and their corresponding AUC values can be seen in appendix 2.

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). The results of this test are displayed in table 8.

Table 8. The correlation between the number of factors used in each model and the test AUC, training 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

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

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 values for *Dendrobium thyrsiflorum* and *Luisia magniflora* were 0.8551 (model 37) and 0.9179 (model 35), respectively. These models were selected to be the final models.

4.3 Final models

The test models of each species with the greatest AUC values were run with all locations used and no random tests. Each was calculated twice, once with all factors used, the other with the *anthropogenic* factors removed.

4.3.1 *Dendrobium thyrsiflorum* – all factors (37A)

The following are the results captured from running MAXENT with the *Dendrobium thyrsiflorum* presence data and using all the factors stipulated in model 37 with a random test percentage of 0. A training AUC value of 0.903 was obtained. Table 9 shows the factors used in this model which will be known as model 37A.

Table 9. The environmental variables used in calculating model 37A with all factors.

Layer name	Description
B4, B5, B6	Raw LS8 bands
RDDIR, RDDIST	Distance and direction to roads
ASP-SLP, DEM	Aspect-slope combined layer and digital elevation model
NDVI, ALB, MSR, RNDVI	Vegetation indices
tct4, tct5, tct6	Tasselled cap transformations 4, 5, and 6
tct_bright, tct_green, tct_wet	Tasselled cap transformations brightness, greenness, and wetness
NIGHT	Lights at night

Figure 8 gives the ROC curve for model 37A. It is fair to say that the model is better than a random model at all points as the training data never falls below the random prediction.

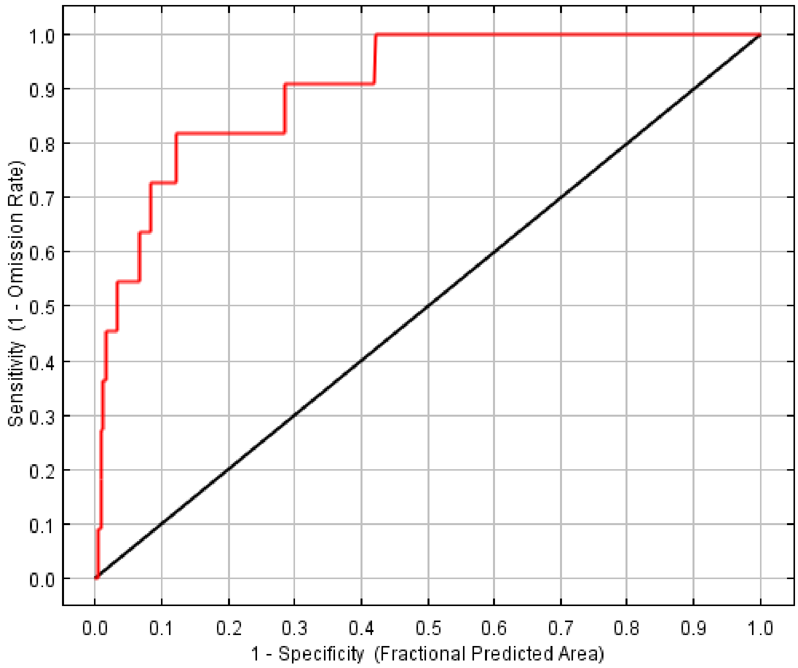

Figure 8. ROC curve for model 37A. Sensitivity vs 1-specificity. Training data is shown in red and the random prediction (AUC = 0.5) is shown in black.

The jackknife test illustrated in figure 9 found the factor that contributed most to the model by itself was the lights at night (NIGHT), the proxy for human habitation density, it subsequently had the most useful data. When the distance to roads (RDDIST) factor was removed, the model suffered the most, suggesting that this environmental variable has the greatest amount of information that none of the other variables possess. The gain contributed by all factors working together is greater than any of the individual factors, showing that the distribution is affected by a combination of variables rather than a single one.

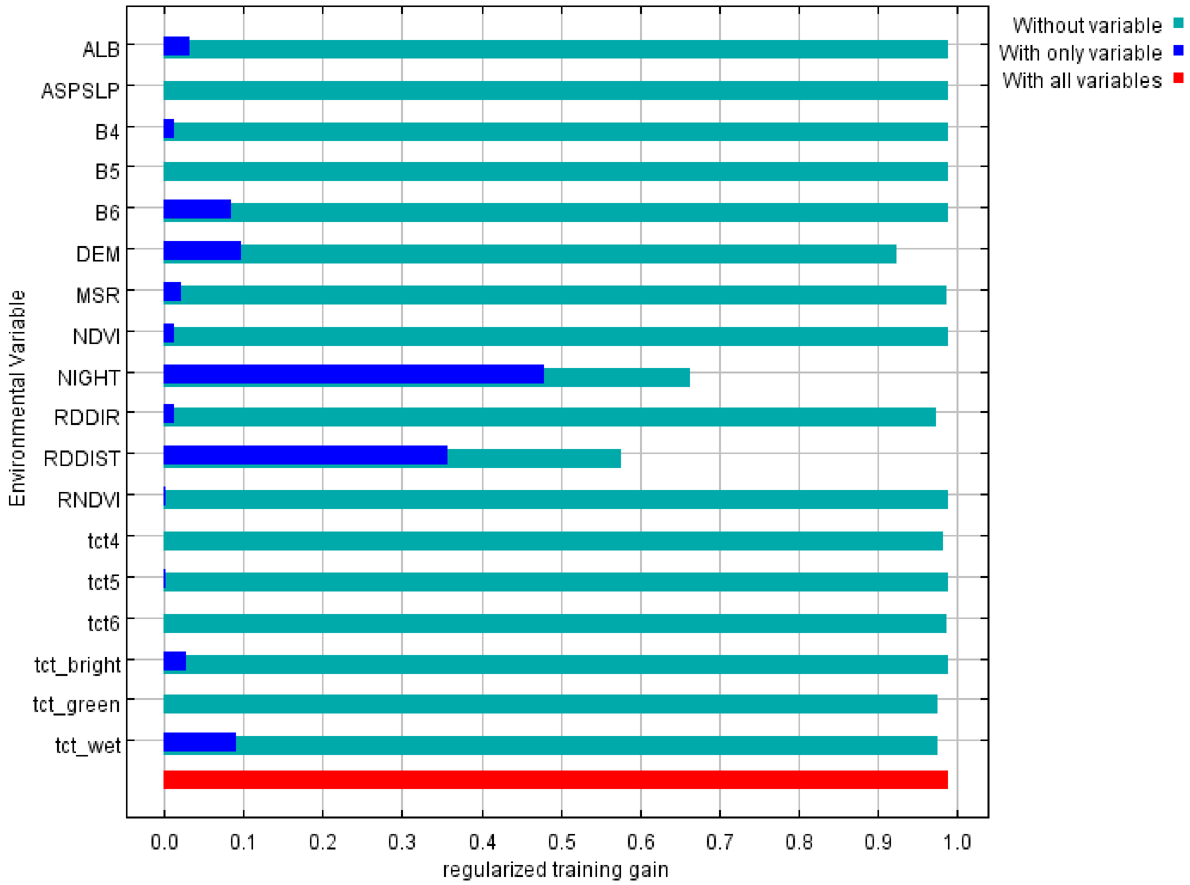

Figure 9. Jackknife of the regularised training gain for model 37A.

4.3.2 *Dendrobium thyrsiflorum* – natural factors only (37N)

The following are the results acquired from running MAXENT with the *Dendrobium thyrsiflorum* presence data and using only the natural factors stipulated in model 37 with a random test percentage of 0. *Anthropogenic* factors were not included in this model. A training AUC value of 0.811 was obtained. Table 10 shows the factors used in this model which will be known as model 37N.

Table 10. The environmental variables used in calculating model 37N with only natural factors.

Layer name	Description
B4, B5, B6	Raw LS8 bands
ASP-SLP, DEM	Aspect-slope combined layer and digital elevation model
NDVI, ALB, MSR, RNDVI	Vegetation indices
tct4, tct5, tct6	Tasseled cap transformations 4, 5, and 6
tct_bright, tct_green, tct_wet	Tasseled cap transformations brightness, greenness, and wetness

Figure 10 gives the ROC curve for model 37N. It is fair to say that the model is better than a random model at all points as the training data never falls below the random prediction.

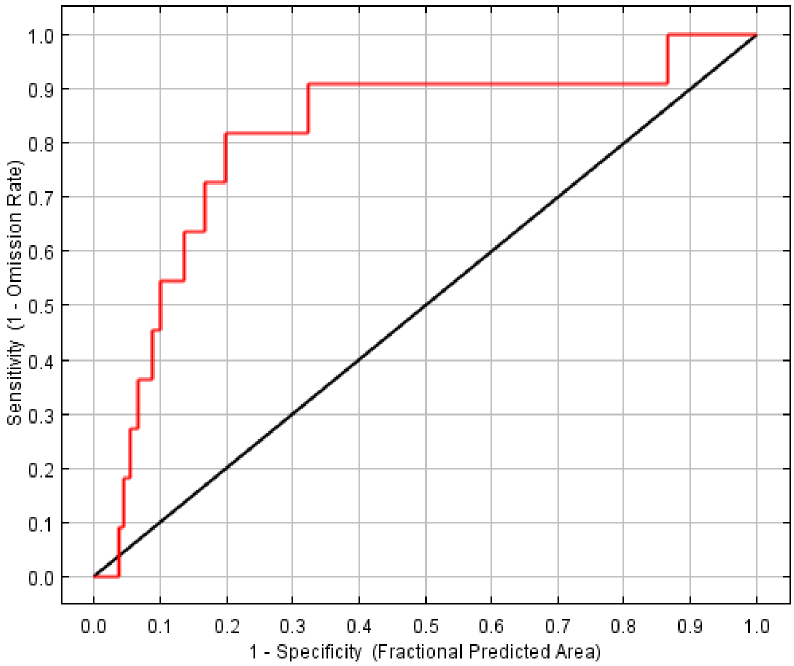

Figure 10. ROC curve for model 37N. Sensitivity vs 1-specificity. Training data is shown in red and the random prediction (AUC = 0.5) is shown in black.

The jackknife test illustrated in figure 11 found the factor contributing most to the model alone was the digital elevation map (DEM), it subsequently had the most useful data to the model. When the digital elevation map (DEM) factor was removed, the model suffered the most, suggesting that this environmental variable also has the greatest amount of information that none of the other variables possess. The gain contributed by all factors working together is greater than any of the individual factors, showing that the distribution is affected by a combination of variables rather than a single one.

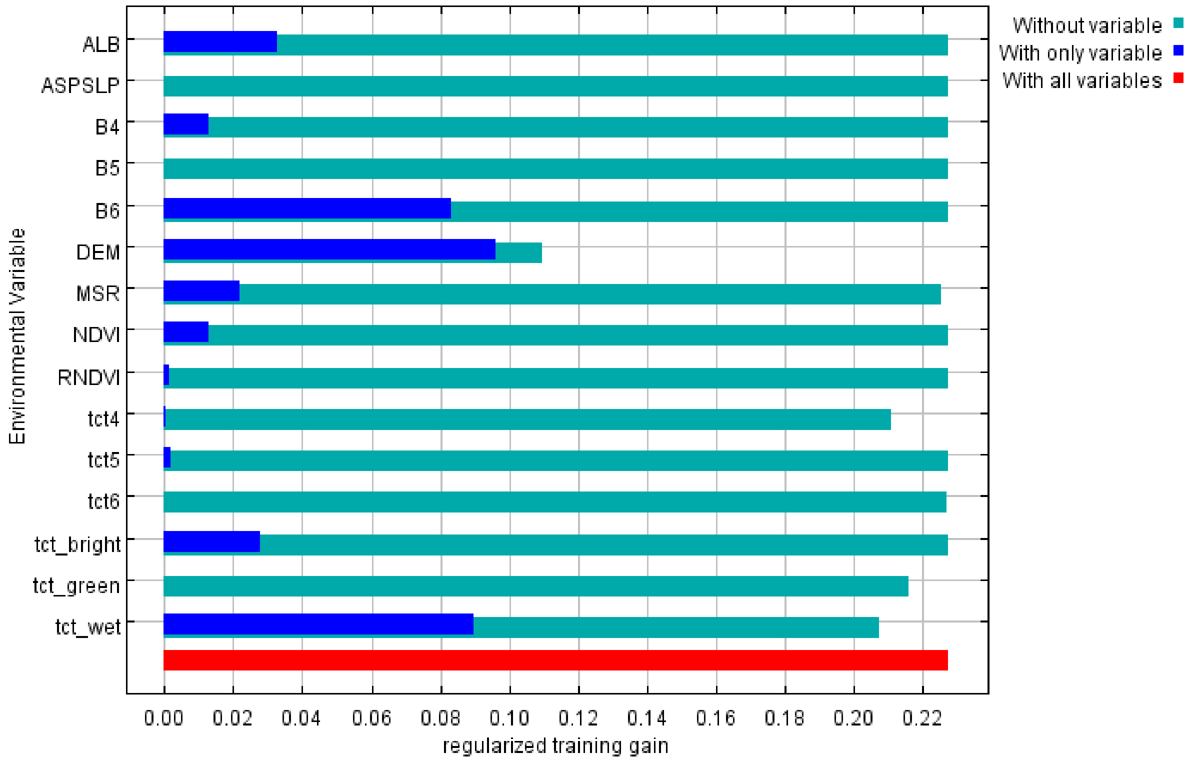

Figure 11. Jackknife of the regularised training gain for model 37N.

4.3.3 *Luisia magniflora* – all factors (35A)

The following are the results procured from running MAXENT with the *Luisia magniflora* presence data and using all the factors stipulated in model 35 with a random test percentage of 0. A training AUC value of 0.926 was obtained. Table 11 shows the factors used in this model which will be known as model 35A.

Table 11. The environmental variables used in calculating model 35A with all factors.

Layer name	Description
B4, B5, B6	Raw LS8 bands
RDDIR, RDDIST	Distance and direction to roads
ASP, SLP, DEM	Aspect, slope, and digital elevation model
NDVI, ALB, MSR, RNDVI	Vegetation indices
tct4, tct5, tct6	Tasselled cap transformations 4, 5, and 6
tct_bright, tct_green, tct_wet	Tasselled cap transformations brightness, greenness, and wetness
NIGHT	Lights at night

Figure 12 gives the ROC curve for model 35A. It is fair to say that the model is better than a random model at all points as the training data never falls below the random prediction.

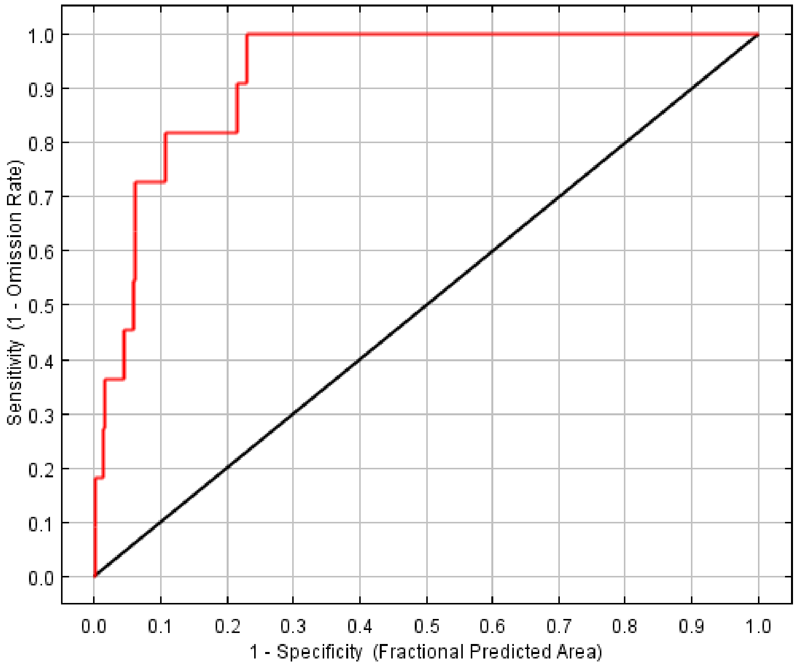

Figure 12. ROC curve for model 35A. Sensitivity vs 1-specificity. Training data is shown in red and the random prediction (AUC = 0.5) is shown in black.

The jackknife test illustrated in `figure 13` found the factor contributing most to the model by itself was the tasseled cap transformation for wetness (tct_wet), it subsequently had the most useful data. When the distance to roads (RDDIST) factor was removed, the model suffered the most, suggesting that this environmental variable has the greatest amount of information that none of the other variables possess. The gain contributed by all factors working together is greater than any of the individual factors, showing that the distribution is affected by a combination of variables rather than a single one.

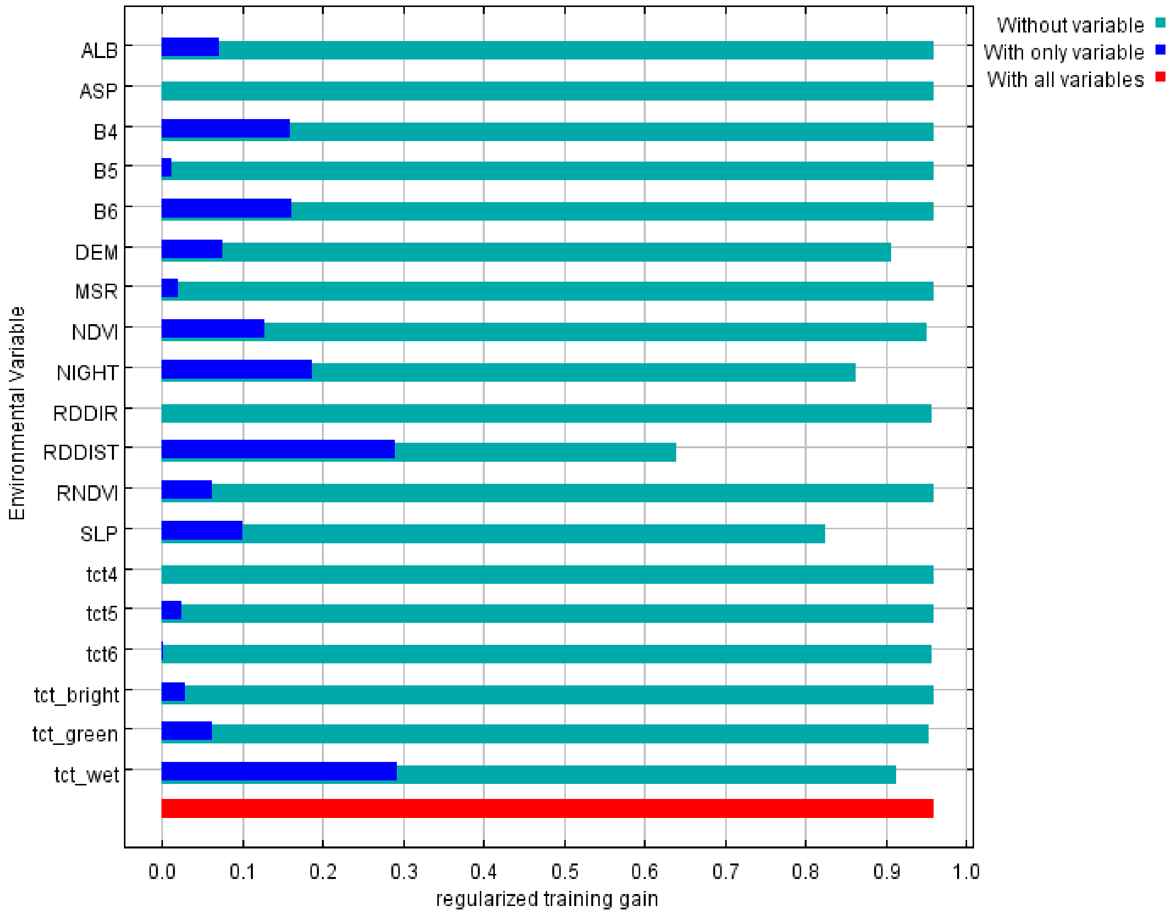

Figure 13. Jackknife of the regularised training gain for model 35A.

4.3.4 *Luisia magniflora* – natural factors only (35N)

The following are the results realised from running MAXENT with the *Luisia magniflora* presence data using only the natural factors stipulated in model 35 with a random test percentage of 0. *Anthropogenic* factors were not included in this model. A training AUC value of 0.878 was obtained. Table 12 shows the factors used in this model which will be known as model 35N.

Table 12. The environmental variables used in calculating model 35 with only natural factors.

Layer name	Description
B4, B5, B6	Raw LS8 bands
ASP, SLP, DEM	Aspect, slope, and digital elevation map
NDVI, ALB, MSR, RNDVI	Vegetation indices
tct4, tct5, tct6	Tasseled cap transformations 4, 5, and 6
tct_bright, tct_green, tct_wet	Tasseled cap transformations brightness, greenness, and wetness

Figure 14 gives the ROC curve for model 35N. It is fair to say that the model is better than a random model at all points as the training data never falls below the random prediction.

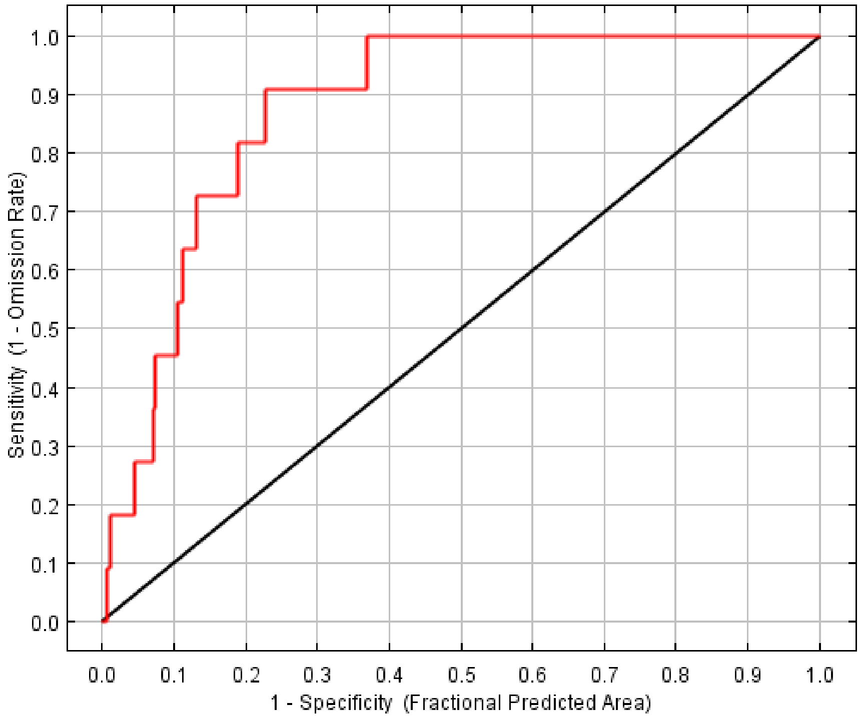

Figure 14. ROC curve for model 35N. Sensitivity vs 1-specificity. Training data is shown in red and the random prediction (AUC = 0.5) is shown in black.

The jackknife test illustrated in `figure 15` found that factor contributing most to the model by itself was the tasseled cap transformation for wetness (tct_wet), it subsequently had the most useful data. When the slope (SLP) factor was removed, the model suffered the most, suggesting that this environmental variable has the greatest amount of information that none of the other variables possess. The gain contributed by all factors working together is greater than any of the individual factors, showing that the distribution is affected by a combination of variables rather than a single one.

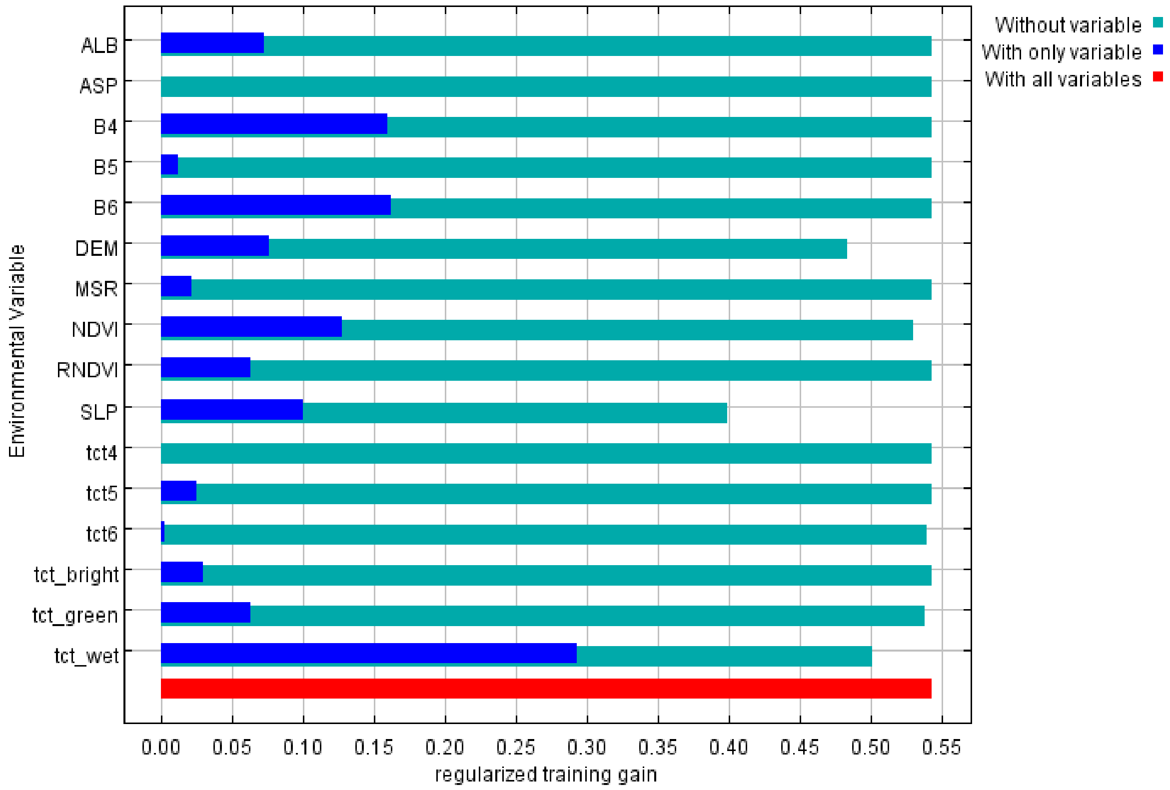

Figure 15. Jackknife of the regularised training gain for model 35N.

4.3.5 Percent contribution

Table 13 gives the percent contribution of each factor to each model. Out of all the final models the only factors used to any extent were DEM, MSR, NDVI, NIGHT, RDDIR, RDDIST, SLP, tct4, tct6, tct_green, and tct_wet. There were several factors that were selected for the final models but did not contribute to the overall model. These factors were those that were usually paired with other factors (like aspect with slope) in the test models. Table 13 can be used to tease out these paired factors and create future models with the most informative factors.

Table 13. The percent contribution of each factor to each model. Factors not used in a model are denoted by an x.

	Model (% contribution)			
Factor	35A	37A	35N	37N
ALB	0	0	0	0
ASP	0	x	0	x
ASPSLP	x	0	x	0
B4	0	0	0	0
B5	0	0	0	0
B6	0	0	0	0
DEM	6.9432	5.3489	16.8271	45.8767
MSR	0	0.0694	0	0.5728
NDVI	0.8986	0	7.0385	0
NIGHT	20.502	49.0186	x	x
RDDIR	0.124	1.4424	x	x
RDDIST	36.5557	39.0899	x	x
RNDVI	0	0	0	0
SLP	11.3862	x	23.2085	x
tct4	0	0.3807	0	8.7807
tct5	0	0	0	0
tct6	0.04	0.4252	0.3327	0.3672
tct_bright	0	0	0	0
tct_green	0.2303	0.5538	1.1132	1.769
tct_wet	23.32	3.6712	51.48	42.6338

4.3.6 Final maps

Below are the final habitat suitability maps created by MAXENT for each of the models 37A, 37N, 35A, and 35N in the logistic format. Figure 16 shows the suitable habitat of *Dendrobium thyrsiflorum* with all factors included (37A). Figure 17 shows the suitable habitat of *Dendrobium thyrsiflorum* with only natural factors included (37N). Figure 18 shows the suitable habitat of *Luisia magniflora* with all factors included (35A). Figure 19 shows the suitable habitat of *Luisia magniflora* with only natural factors included (35N).

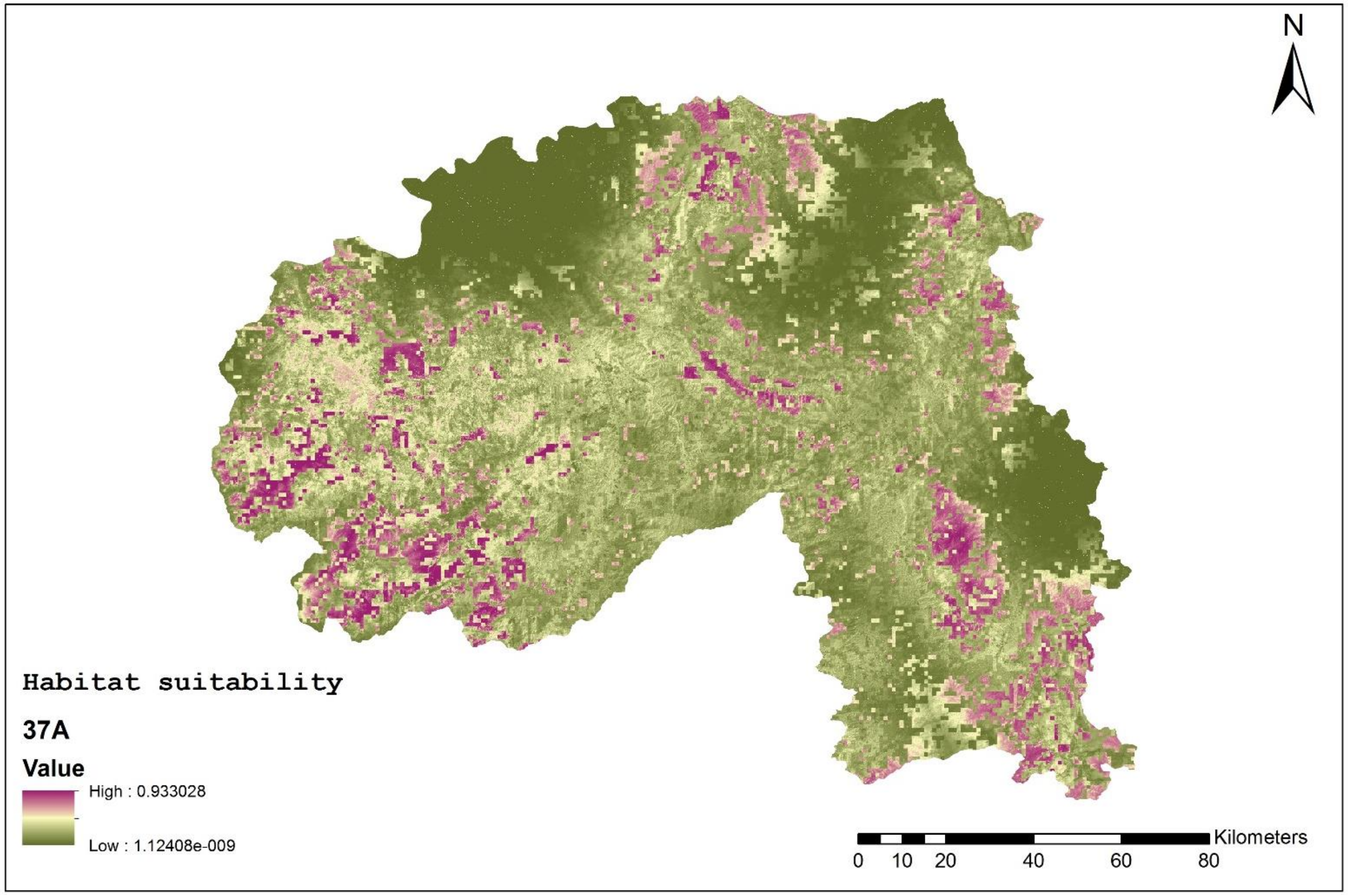

Figure 16. Habitat suitability map of *Dendrobium thyrsiflorum* in model 37A. High suitability is indicated by a value close to 1 and low suitability is indicated by a value close to 0.

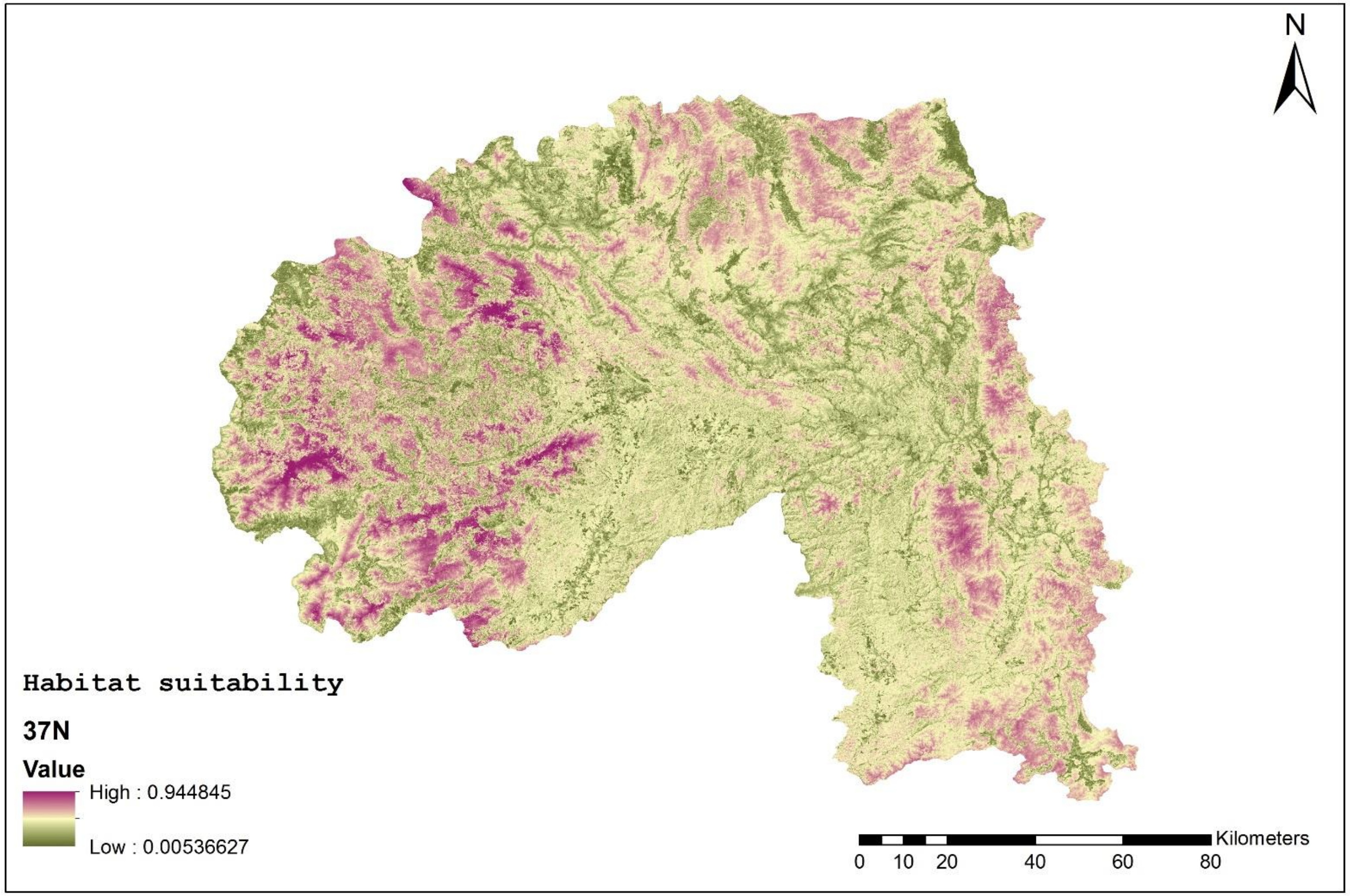

Figure 17. Habitat suitability map of *Dendrobium thyrsiflorum* in model 37N. High suitability is indicated by a value close to 1 and low suitability is indicated by a value close to 0.

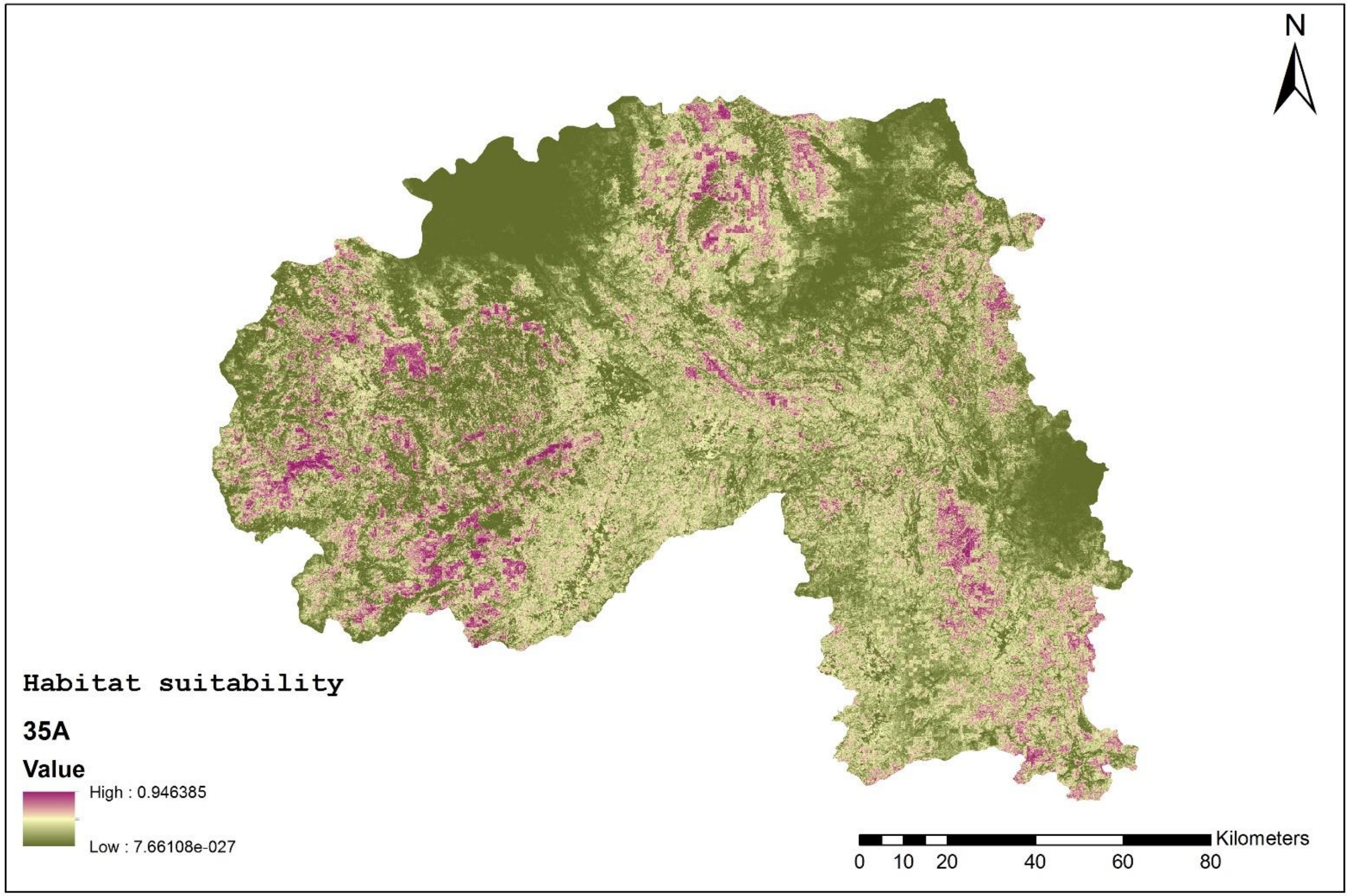

Figure 18. Habitat suitability map of *Luisia magniflora* in model 35A. High suitability is indicated by a value close to 1 and low suitability is indicated by a value close to 0.

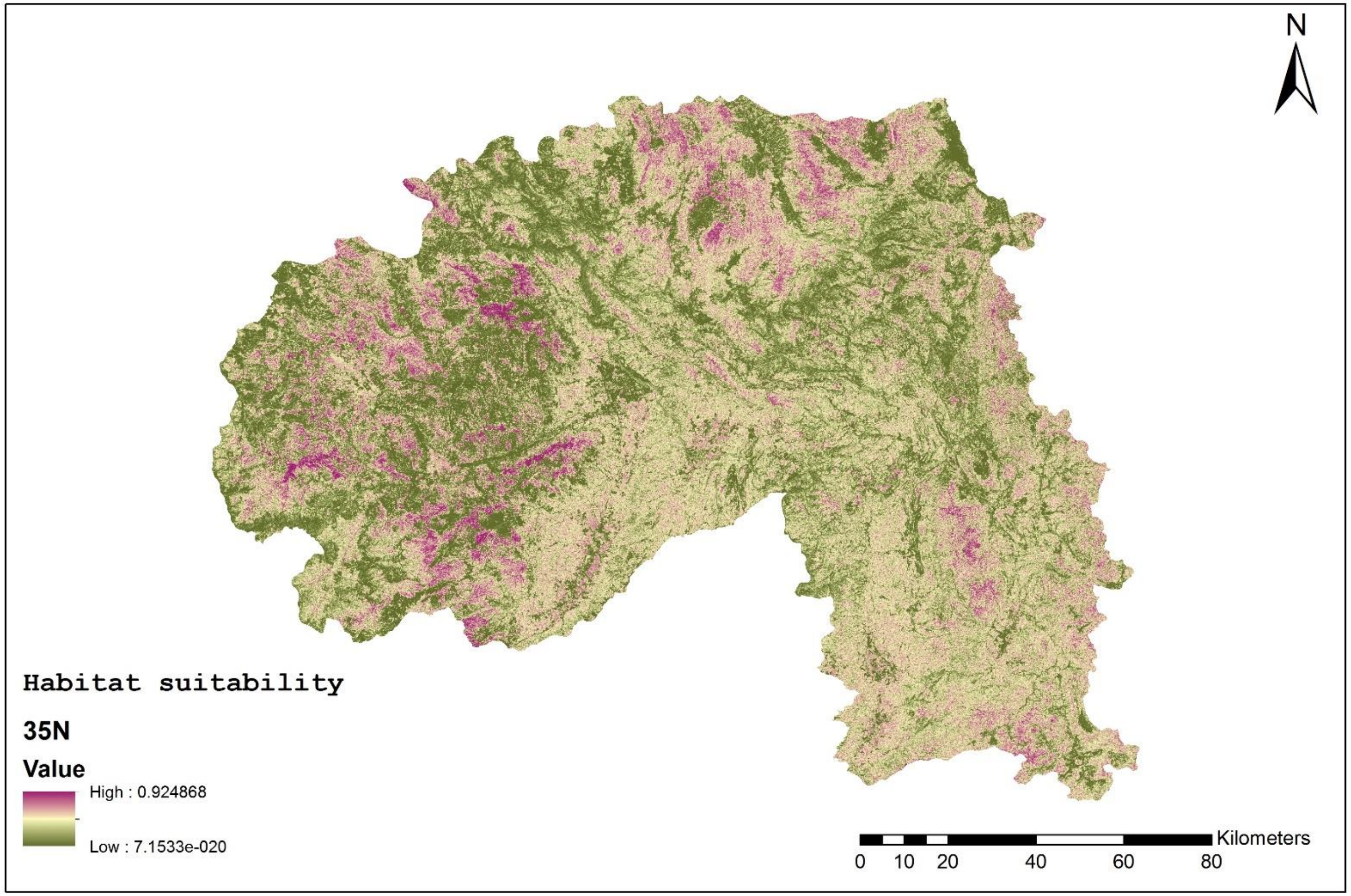

Figure 19. Habitat suitability map of *Luisia magniflora* in model 35N. High suitability is indicated by a value close to 1 and low suitability is indicated by a value close to 0.

4.4 Efficiency models

Table 14 displays the four most contributing factors in model 37A. These factors were used in creating the *Dendrobium thyrsiflorum* efficiency model with a random test percentage of 0. A training AUC value of 0.879 was obtained.

Table 14. Factors and their descriptions used in the *Dendrobium* efficiency model.

Layer name	Description
NIGHT	Lights at night
RDDIST	Distance to roads
DEM	Digital elevation model
tct_wet	Tasselled cap transformation for wetness

Figure 20 gives the ROC curve for the *Dendrobium* efficiency model. It is fair to say that the model is better than a random model at all points as the training data never falls below the random prediction.

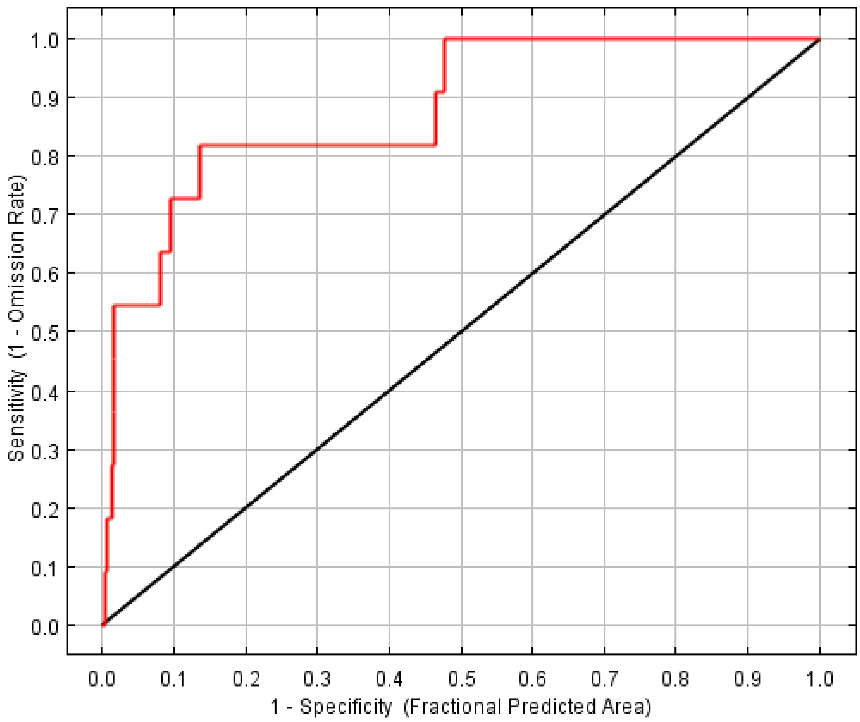

Figure 20. ROC curve for the *Dendrobium* efficiency model. Sensitivity vs 1-specificity. Training data is shown in red and the random prediction (AUC = 0.5) is shown in black.

The jackknife test illustrated in figure 21 found the factor that contributed most to the model alone was the lights at night (NIGHT), the proxy for human habitation density, it subsequently had the most useful data. When the distance to roads (RDDIST) factor was removed, the model suffered the most, suggesting that this environmental variable has the greatest amount of information that none of the other variables possess. The gain contributed by all factors working together is greater than any of the individual factors, showing that the distribution is affected by a combination of variables rather than a single one. These results correspond with the results from model 37A of which this model was based on.

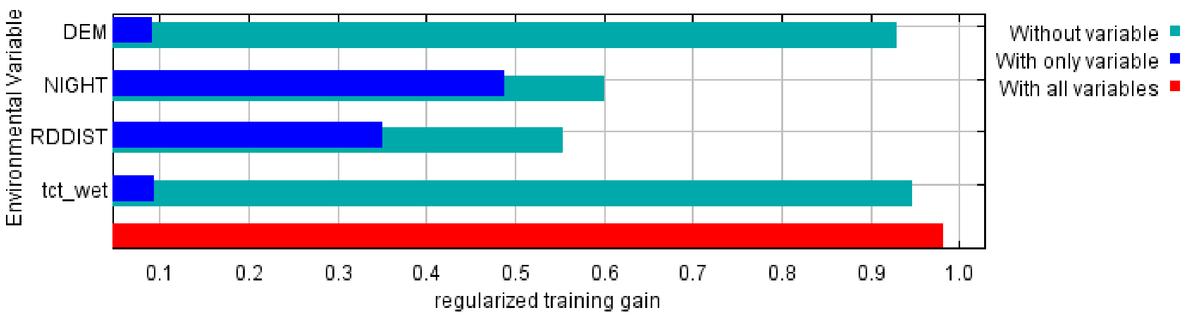

Figure 21. Jackknife of the regularised training gain for the *Dendrobium* efficiency model.

Table 15 displays the four most contributing factors in model 37A. These factors were used in creating the *Luisia magniflora* efficiency model with a random test percentage of 0. A training AUC value of 0.896 was obtained.

Table 15. Factors and their descriptions used in the *Luisia* efficiency model.

Layer name	Description
RDDIST	Distance to roads
tct_wet	Tasselled cap transformation for wetness
NIGHT	Lights at night
SLP	Slope

Figure 22 gives the ROC curve for model for the *Luisia* efficiency model. It is fair to say that the model is better than a random model at all points as the training data never falls below the random prediction.

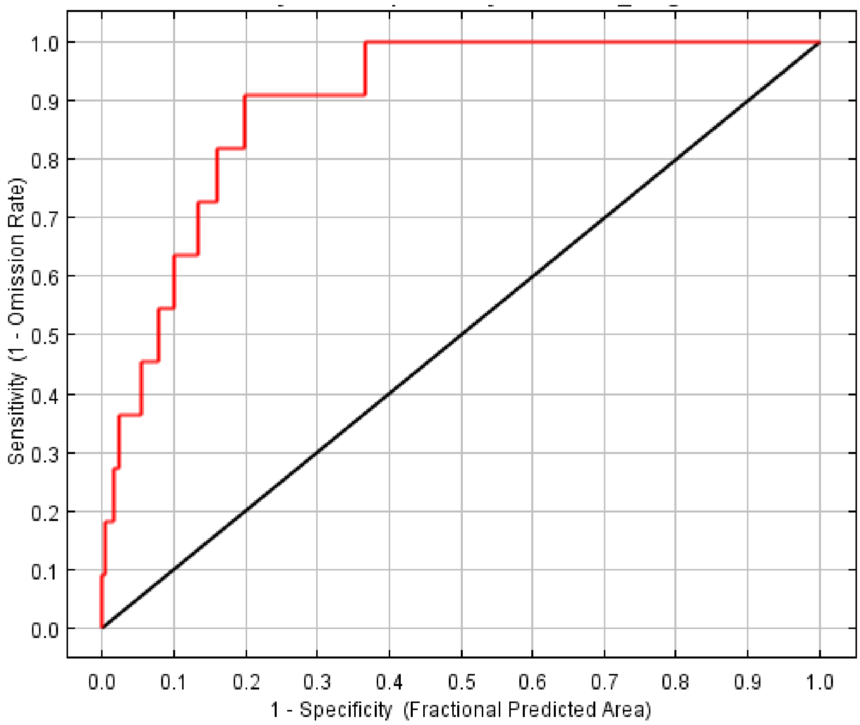

Figure 22. ROC curve for the *Luisia* efficiency model. Sensitivity vs 1-specificity. Training data is shown in red and the random prediction (AUC = 0.5) is shown in black.

The jackknife test illustrated in figure 23 found that the factor that contributed most to the model by itself was the tasseled cap transformation for wetness (tct_wet), it therefore had the most useful information. When the distance to roads (RDDIST) factor was removed, the model suffered the most, suggesting that this environmental variable has the greatest amount of information that none of the other variables possess. The gain contributed by all factors working together is greater than any of the individual factors, showing that the distribution is affected by a combination of variables rather than a single one. These results correspond with the results of model 35A of which this model was based on.

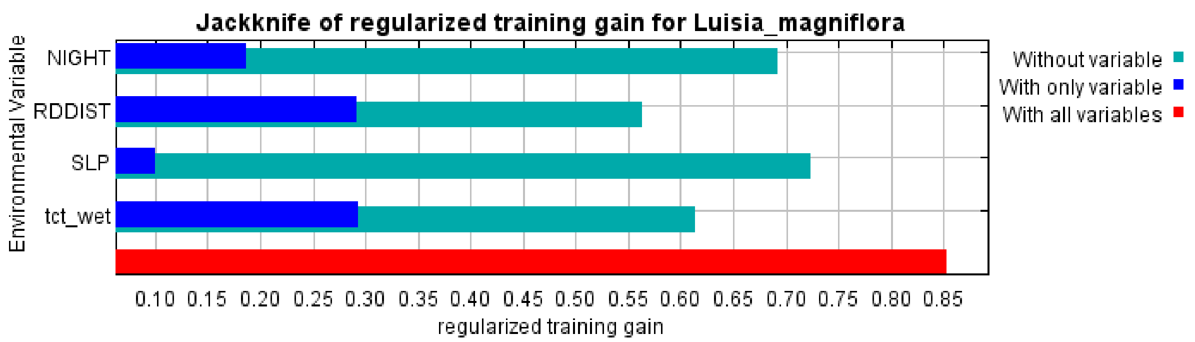

Figure 23. Jackknife of the regularised training gain for the *Luisia* efficiency model.

5.0 Discussion

This study has created habitat suitability maps, with MAXENT, of two orchid species with and without factors directly influenced by humans. An investigation into the relationship between the number of factors used and the information related in the subsequent model was carried out and a jackknife test of the environmental variables used in the final models was completed.

Although former studies have used MAXENT to determine the most influential factors in the distribution of animals and plants, other studies have looked at the orchids of Xishuangbanna (Liu *et al.* 2015), and researchers have looked at human disturbance in Xishuangbanna; however I have not found studies in which habitat suitability models are used in Xishuangbanna. I have also found no studies using MAXENT to directly investigate the impact of humans on a flora or fauna. The information regarding the two species in this study, *Luisia magniflora* and *Dendrobium thyrsiflorum,* is sparse and there is little information about what environmental variables influence their growth and distribution.

In the following sections I will discuss the environmental variables used in each model, the habitat suitability maps, the assumptions of these models, and future improvements and steps.

5.1 Predictability of the models

Since MAXENT only makes use of an environmental variable when there is evidence that it contributes to the distribution of a species, it stands to reason that there is a correlation between the number of environmental variables used in each model and the training AUC of the model. While there was only a correlation between the number of environmental variables and the test AUC in the *Dendrobium* models, I would expect that with a greater number of models, the correlation would prove to be significant for the *Luisia* too. The selection of the final models matches these findings since the models used 18 and 19 of the possible 20 environmental variables.

5.2 Presence of humans

The most informative environmental variables used in model 37A, in which all factors were modeled for *Dendrobium thyrsiflorum,* were the lights at night (49% contribution) and distance to roads (39.1%) factors. This shows that there is a strong influence of human presence and human access on *Dendrobium thyrsiflorum.* Model 35A, in which all factors were modeled for *Luisia magniflora,* was also highly influenced by the distance to roads (36.6%) and lights at night (20.5%) as well as the tasseled cap transformation for wetness (23.3%). This suggests that *Luisia magniflora* is dependent on specific moisture levels but is also influenced by the presence of humans.

The fact that locations of settlements and roads influence the distribution of these two orchid species suggests that the populations of these species are being adversely influenced by human presence. Since the distance to roads factor is informative along with the lights at night, it is likely that the populations are being impacted by the collection of these species rather than the presence of rubber plantations, deforestation, or general human presence which would be influenced by altitude or a vegetation index, respectively. However it is possible that the mere presence of the roads is disturbing the orchid populations through pollution, edge effects, or fragmentation.

5.3 Natural variables

Models 35N and 37N, in which only the natural factors were modeled for *Luisia magniflora* and *Dendrobium thyrsiflorum* respectively, show the factors influencing the distribution of each species when there is little or no disturbance by humans. *Dendrobium thyrsiflorum* is most influenced by altitude (45.9% contribution), and the tasseled cap transformation for wetness (42.6%). This is corroborated by the literature which shows that this species, like many other orchids in Xishuangbanna, has a lower altitude limit of about 1000m and is sensitive to extremes of moisture (Liu *et al.* 2015). *Luisia magniflora* is most influenced by the tasseled cap transformation for wetness (51.5%), slope (23.2%) and altitude (16.8%). This too makes sense, as *Luisia* species generally have a moisture preference.

5.4 Efficiency models

The results of the efficiency models show that a high AUC value and model accuracy can be obtained with the use of a small number of carefully selected factors. While much research must be done in selecting the appropriate factors, this high efficiency method could make it possible to model many different areas in a short period of time and therefore create a summary of all the known habitats of a species in a single project. The technique however, can only be used on species that have known factors affecting their distribution because the selection of the wrong factors could create inaccurate and misleading models.

5.5 This is a snapshot

It should be noted that Xishuangbanna is a rapidly changing region and the research done here is only a snapshot in time rather than an evolving landscape. Although the altitude, slope, and aspect of the region are unlikely to change drastically in the near future, the presence of settlements and roads, and the amount of forest are likely to change rapidly. The rubber industry has only benefited the people of the region – those owning the rubber have made magnificent profits and those working for the rubber farmers are generally paid fair wages – although profits have recently dwindled due to the high supply in the sector (Sreekar & Huang 2015; Hu 2016). There are also plans for the further connection of Xishuangbanna to China's high-speed railway system and more of Southeast Asia. With greater economic opportunity and access to the region, it is fair to predict that there will be an expansion of Xishuangbanna's settlements and roads in the future.

5.6 Future improvements and work

This study has shed light on several things that require further investigation.

The following include several environmental variables which could be used in further MAXENT models:

Lidar – Lidar is a process of using aircraft or spacecraft borne lasers to create high resolution terrain maps of the land's surface. This information could be useful to future models but is only available to government agencies within China.

Distance to rivers – The location of water may have an effect on the distribution of orchids or the trees that epiphytic orchids reside in. Although this information was not available at the time of modelling, it may be possible to incorporate it into future models.

Using the cost-distance tool in ArcGIS it is possible to relate the distance from roads with vegetation indices. This could be used in future studies to look at the trade-off for harvesters between travelling through thick vegetation (and therefore greater energy expenditure) and the increased probability of coming across orchids to harvest.

Bioclimatic layers – Many habitat suitability models incorporate bioclimatic layers, such as annual precipitation and mean temperature. When creating the models in this study, the decision to not use these factors was made as layers with the appropriate resolution were not available, however future studies may be able to take advantage of greater resolution layers.

Higher resolution vegetation indices – The Sentinel-2 satellites from ESA's Copernicus program have started recording multi-spectral images of Earth at 10m resolution (Drusch *et al.* 2012). Taking advantage of this higher resolution and the better sensors of future satellites may increase the predictive ability of MAXENT studies.

While the number of presence points used in this study was adequate to create habitat suitability models, it would be expected that a greater number of presence points would improve the model. While spending field days to discover a greater number of individuals has a very low return on investment (one individual of the target species is found roughly every ten field days), the use of herbaria and citizen science projects like the Global Biodiversity and Information Facility (gbif.org) could increase the number of known locations with little investment in time (GBIF 2015). This data was not taken advantage of in this study since the area where an orchid may have been in historical records may have been converted to rubber plantation or built on in the time since the record was made. Not only would this information be inaccurate, it may damage the model by providing information to MAXENT that suggests that the species can survive in a certain habitat it cannot. This problem can be rectified by travelling to these locations to confirm if the species is still present or at least there is the chance that it is present. While this method does not eliminate all field days, it reduces them to certain highly targeted areas. The Chinese National Herbarium and the Xishuangbanna Tropical Botanical Garden herbarium both have specimens of the target species that are available for researchers to study but citizen science programs in China are not yet popular.

Machine learning and computer vision – of the more speculative options, machine learning may drastically improve the ability of MAXENT as well as create new algorithms that are more able to quickly and accurately create habitat suitability models. Machine learning may also be incorporated in identifying different vegetation types from satellite images as well as the identification of species during field surveys and market surveys. This technique has already been demonstrated by Kumar *et al.* (2012) and it will not be long until computer vision methods are more available to the public.

The creation of these models has prepared all the needed materials to create more MAXENT models on the other orchid species within Xishuangbanna. While the preparation needed to create habitat suitability models takes some time, running the

MAXENT with presence data of other species takes a relatively small amount of time. Future studies may create habitat suitability models for all known species in Xishuangbanna of which data is available.

5.6 Why does this matter, and what does this teach us?

This study has been able to identify that humans are having a noticeable effect on the distributions of both *Dendrobium thyrsiflorum* and *Luisia magniflora*. Most likely, this effect is coming from wild harvesting of these species, as there is a link between the distance to roads and the distribution of each species. Additionally, this study has identified the most important factors in determining the distribution of these species without the presence of humans; *Dendrobium thyrsiflorum* is adapted to certain altitudes (most notably above 1000m) and is adapted to certain moisture levels, *Luisia magniflora* has a preference for certain moisture levels as well as certain slope gradients.

This study has also shown that it is possible to create informative habitat suitability models with very few environmental variables as long as the selection of factors is done carefully and with accurate information. Informative and rapidly fabricated habitat suitability models are essential to the effective management of species and ecosystems. With increasing access to environmental data as well as tools like MAXENT, the public and researchers have greater ability to create habitat suitability models of species they are interested in.

6.0 Conclusions

All three of the initial aims of this study have been met. This study has identified which factors determine the distribution of the two orchid species, it has shown that humans have an influence on their distribution, and it has shown that it is possible to create informative models with few highly selected factors.

There have been several methods recommended for mitigating the general wildlife trade. These include certification programmes (Treves & Jones 2010), public awareness (Zhang, Hua & Sun 2008; Williams *et al.* 2012), increased enforcement (Nijman & Shepherd 2012) and sustainable harvest (Milner-Gulland & Bennett 2003; Liu *et al.* 2014). Until recently, the approaches proposed for illegal trade mitigation have centred around enforcement, lately it has been realized that trade bans, in some cases can drive trade into the black market (Verissimo, Challender & Nijman 2012; Biggs *et al.* 2013) and increase overexploitation (Conrad 2012).

There are several of these options that may have an ability to reduce the disturbance caused by wild harvesting of these orchids. Liu *et al.* (2014) demonstrated a novel method of sustainably cultivating epiphytic orchids in the wild there is no evidence to suggest this would not be possible with the species in this study. Certification programmes in which the public are encouraged to check whether the orchids they are purchasing are sustainably harvested and public awareness programmes may also be successful. As the populations of these orchids are vastly reduced and they are present in orchid markets, a multifaceted approach in which many of these methods are implemented is encouraged.

Liu *et al.* (2015) have cautioned that although the rubber plantations of Xishuangbanna pose little risk to most of the orchids in the region (this study found little evidence that the two species studied here are influenced by rubber), climate change may allow rubber to be grown at higher altitudes in the near future. This would encroach on the areas available to these orchids (especially *Dendrobium thyrsiflorum*) and make conservation of these species challenging. Assisted migration may be useful in this case to move populations to areas not suitable for rubber plantations.

This study has demonstrated the fragility of orchid populations in Xishuangbanna and the robustness of habitat suitability models in their ability to model species distributions.

References

Araújo, M.B. & Peterson, A.T. (2012) Uses and misuses of bioclimatic envelope modeling. *Ecology*, **93**, 1527–1539.

Baig, M.H.A., Zhang, L., Shuai, T. & Tong, Q. (2014) Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance. *Remote Sensing Letters*, **5**, 423–431.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **222**, 1810–1819.

Bennie, J., Hill, M.O., Baxter, R. & Huntley, B. (2006) Influence of slope and aspect on long-term vegetation change in British chalk grasslands. *Journal of Ecology*, **94**, 355–368.

Biggs, D., Courchamp, F., Martin, R. & Possingham, H.P. (2013) Legal Trade of Africa's Rhino Horns. *Science*, **339**, 1038–1039.

Brewer, C.A. & Marlow, K.A. (1993) Color representation of aspect and slope simultaneously. *Autocarto conference* pp. 328–328. ASPRS American Society for Photogrammetry.

Cao, M. & Zhang, J. (1997) Tree species diversity of tropical forest vegetation in Xishuangbanna , SW China. *Biodiversity and Conservation*, **6**, 995–1006.

Chase, M.W., Cameron, K.M., Freudenstein, J.V., Pridgeon, A.M., Salazar, G., van den Berg, C. & Schuiteman, A. (2015) An updated classification of Orchidaceae. *Botanical Journal of the Linnean Society*, **177**, 151–174.

Chen, J.M. (1996) Evaluation of Vegetation Indices and a Modified Simple Ratio for Boreal Applications. *Canadian Journal of Remote Sensing*, **22**, 229–242.

Chen, Y.Y., Bao, Z.X., Qu, Y., Li, W. & Li, Z.Z. (2014) Genetic diversity and population structure of the medicinal orchid *Gastrodia elata* revealed by microsatellite analysis. *Biochemical Systematics and Ecology*, **54**, 182–189.

CITES. (1973) *Convention on International Trade in Endangered Species of Wild Fauna and Flora*. Washington, DC.

Conrad, K. (2004) Probability distributions and maximum entropy. *Entropy*, 1–27.

Conrad, K. (2012) Trade bans: a perfect storm for poaching? *Tropical Conservation Science*, **5**, 245–254.

Ding, G., Zhang, D., Ding, X., Zhou, Q., Zhang, W. & Li, X. (2008) Genetic variation and conservation of the endangered Chinese endemic herb *Dendrobium officinale* based on SRAP analysis. *Plant Systematics and Evolution*, **276**, 149–156.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P. & others. (2012) Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, **120**, 25–36.

eFloras. (2016a) *Dendrobium thyrsiflorum*. *Flora of China*, **25**, 368, 376.

eFloras. (2016b) *Luisia*. *Flora of China*, **25**, 8, 13, 426.

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, a., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, a., Li, J., Lohmann, L.G., Loiselle, B. a., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M.C., Peterson, a T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330–342.

ESRI. (2015) *ESRI ArcMap*. Environmental Systems Research Insti, Redlands, CA.

Ewen, J.G., Soorae, P.S. & Canessa, S. (2014) Reintroduction objectives, decisions and outcomes: global perspectives from the herpetofauna. *Animal Conservation*, **17**, 74–81.

GBIF. (2015) *2014 Science Review*. Global Biodiversity Information Facility, Copenhagen.

Goh, C.J. & Kavaljian, L.G. (1989) Orchid industry of Singapore. *Economic Botany*, **43**, 241–254.

Gong, P., Pu, R., Biging, G.S. & Larrieu, M.R. (2003) Estimation of forest leaf area index using vegetation indices derived from Hyperion hyperspectral data. *Geoscience and Remote Sensing, IEEE Transactions on*, **41**, 1355–1362.

Gravendeel, B., Smithson, A., Slik, F.J.W. & Schuiteman, A. (2004) Epiphytism and pollinator specialization: drivers for orchid diversity? *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 1523–1535.

Hastie, T., Tibshirani, R. & Friedman, J. (2009) The Elements of Statistical Learning. *Elements*, **1**, 337–387.

Hexagon Geospatial. (2014) *ERDAS IMAGINE 2015*. Hexagon Geospatial.

Hongmao, L., Zaifu, X., Youkai, X. & Jinxiu, W. (2002) Practice of conserving plant diversity through traditional beliefs: a case study in Xishuangbanna, southwest China. *Biodiversity & Conservation*, **11**, 705–713.

Hu, S. (2016) Personal communication.

Huang, B.-Q., Yang, X.-Q., Yu, F.-H., Luo, Y.-B. & Tai, Y.-D. (2008) Surprisingly high orchid diversity in travertine and forest areas in the Huanglong valley, China, and implications for conservation. *Biodiversity and Conservation*, **17**, 2773–2786.

Hughes, A.C. (2015) Introductory GIS & SDM Workshop.

Jaynes, E.T. (1963) *Information Theory and Statistical Mechanics*.

Joppa, L.N., Roberts, D.L. & Pimm, S.L. (2010) How many species of flowering plants are there? *Proceedings of the Royal Society of London B: Biological Sciences*.

Kumar, N., Belhumeur, P.N., Biswas, A., Jacobs, D.W., Kress, W.J., Lopez, I. & Soares, J.V.B. (2012) Leafsnap: A Computer Vision System for Automatic Plant Species Identification. *The 12th European Conference on Computer Vision (ECCV)*

Lahoz-Monfort, J.J. (2008) *Habitat Suitability Modelling for the Alaotran Gentle Lemur (Hapalemur Alaotrensis)*. Imerial College London, London.

Li, H., Ma, Y., Aide, T.M. & Liu, W. (2008) Past, present and future land-use in Xishuangbanna, China and the implications for carbon dynamics. *Forest Ecology and Management*, **255**, 16–24.

Liu, Q., Chen, J., Corlett, R.T., Fan, X., Yu, D., Yang, H. & Gao, J. (2015) Orchid conservation in the biodiversity hotspot of southwestern China: Orchid Conservation in Xishuangbanna. *Conservation Biology*, **29**, 1563–1572.

Liu, H., Luo, Y.-B., Heinen, J., Bhat, M. & Liu, Z.-J. (2014) Eat your orchid and have it too: A potentially new conservation formula for Chinese epiphytic medicinal orchids. *Biodiversity and Conservation*, **23**, 1215–1228.

Liu, H., Luo, Y. & Liu, H. (2010) Studies of Mycorrhizal Fungi of Chinese Orchids and Their Role in Orchid Conservation in China — A Review. *Botanical Review*, 241–262.

López-Pujol, J., Zhang, F.-M. & Ge, S. (2006) Plant Biodiversity in China: Richly Varied, Endangered, and in Need of Conservation. *Biodiversity and Conservation*, **15**, 3983–4026.

Lozier, J.D., Aniello, P. & Hickerson, M.J. (2009) Predicting the distribution of Sasquatch in western North America: anything goes with ecological niche modelling. *Journal of Biogeography*, **36**, 1623–1627.

Matheson, C.A. (2014) iNaturalist. *Reference Reviews*, **28**, 36–38.

McCormick, M.K. & Jacquemyn, H. (2014) What constrains the distribution of orchid populations? *New Phytologist*, **202**, 392–400.

Merow, C., Smith, M.J. & Silander, J.A. (2013) A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, **36**, 1058–1069.

Milner-Gulland, E.J. & Bennett, E.L. (2003) Wild meat: the bigger picture. *Trends in Ecology & Evolution*, **18**, 351–357.

NASA. (2013) *Landsat Data Continuity Mission Press Kit*. National Aeronautics and Space Administration (NASA).

NASA JPL. (2009) ASTER Global Digital Elevation Model.

Ng, T.B., Liu, J., Wong, J.H., Ye, X., Wing Sze, S.C., Tong, Y. & Zhang, K.Y. (2012) Review of research on Dendrobium, a prized folk medicine. *Applied Microbiology and Biotechnology*, **93**, 1795–1803.

Nijman, V. & Shepherd, C. (2012) The role of Lao PDR in the ivory trade. *TRAFFIC Bulletin*, **24**, 35–40.

Nouri, H., Beecham, S., Anderson, S. & Nagler, P. (2014) High Spatial Resolution WorldView-2 Imagery for Mapping NDVI and Its Relationship to Temporal Urban Landscape Evapotranspiration Factors. *Remote Sensing*, **6**, 580–602.

Orbital ATK. (2015) *Landsat 8 Fact Sheet: Continuing the Landsat Mission*. Orbital ATK.

Packwood, A. (2015) Personal communication.

Pearson, R.G., Raxworthy, C.J., Nakamura, M. & Townsend Peterson, a. (2007) Predicting species distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102–117.

Penfield, P. & Lloyd, S. (2003) Principle of Maximum Entropy Chapter 8. *6.050J Information and Entropy* MIT OpenCourseWare., Spring 2008 pp. 85–92. MIT OpenCourseWare, Massachusetts Institute of Technology.

Phillips, S.J. (2008) Transferability, Sample Selection Bias and Background Data in Presence-Only Modelling: A Response to Peterson et al. (2007). *Ecography*, **31**, 272–278.

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

Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J. & Ferrier, S. (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, **19**, 181–197.

Phillips, S., Dudík, M. & Schapire, R. (2004) A maximum entropy approach to species distribution modeling. *Proceedings of the twenty-first \ldots*, 655–662.

Pimm, S.L. & Raven, P. (2000) Biodiversity: Extinction by numbers. *Nature*, **403**, 843–845.

Porras-Alfaro, A. & Bayman, P. (2007) Mycorrhizal fungi of Vanilla: diversity, specificity and effects on seed germination and plant growth. *Mycologia*, **99**, 510–525.

Rasmussen, H.N. & Rasmussen, F.N. (2009) Orchid mycorrhiza: implications of a mycophagous life style. *Oikos*, **118**, 334–345.

Roujean, J.-L. & Breon, F.-M. (1995) Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment*, **51**, 375–384.

Sang, W., Ma, K. & Axmacher, J.C. (2011) Securing a Future for China's Wild Plant Resources. *BioScience*, **61**, 720–725.

Särkinen, T., Gonzáles, P. & Knapp, S. (2013) Distribution models and species discovery: the story of a new Solanum species from the Peruvian Andes. *PhytoKeys*, **20**, 1–20.

Shou-qing, Z. The Vulnerable and Endangered Plants of Xishuangbanna Prefecture, Yunnan Province, China. *Xishuangbanna's Endangered Flora*, 1–8.

Small, C., Pozzi, F. & Elvidge, C.D. (2005) Spatial analysis of global urban extent from DMSP-OLS night lights. *Remote Sensing of Environment*, **96**, 277–291.

Sreekar, R. & Huang, G. (2015) Personal communication.

Tian, H., Chen, L. & Xing, F. (2013) Species diversity and conservation of orchids in Nanling National Nature Reserve, Guangdong. *Biodiversity Science*, **21**, 224–231.

Treves, A. & Jones, S.M. (2010) Strategic tradeoffs for wildlife-friendly eco-labels. *Frontiers in Ecology and the Environment*, **8**, 491–498.

Verissimo, D., Challender, D. & Nijman, V. (2012) Wildlife trade in Asia: start with the consumer. *Asian Journal of Conservation Biology*, **1**, 49–50.

Version 4 DMSP-OLS Nighttime Lights Time Series. (2013)

Webber, B.L., Yates, C.J., Le Maitre, D.C., Scott, J.K., Kriticos, D.J., Ota, N., McNeill, A., Le Roux, J.J. & Midgley, G.F. (2011) Modelling horses for novel climate courses: insights from projecting potential distributions of native and alien Australian acacias with correlative and mechanistic models. *Diversity and Distributions*, **17**, 978–1000.

Williams, S.J. (2008) *The Identification and Conservation of Important Plant Areas : A Case Study from the Turks and Caicos Islands*. Imerial College London, London.

Williams, S.J., Jones, J.P.G. & Annewandter, R. (2014) Cultivation can increase harvesting pressure on overexploited plant populations. *Ecological Applications*, **24**, 2050–2062.

Williams, S.J., Jones, J.P.G., Clubbe, C. & Gibbons, J.M. (2012) Training programmes can change behaviour and encourage the cultivation of over-harvested plant species. *PLoS ONE*, **7**, 1–9.

Xing, X., Gai, X., Liu, Q., Hart, M.M. & Guo, S. (2014) Mycorrhizal fungal diversity and community composition in a lithophytic and epiphytic orchid. *Mycorrhiza*.

Xu, J., Grumbine, R.E. & Beckschäfer, P. (2014) Landscape transformation through the use of ecological and socioeconomic indicators in Xishuangbanna, Southwest China, Mekong Region. *Ecological Indicators*, **36**, 749–756.

Yuan, Y.H., Hou, B.W., Xu, H.J., Luo, J. & Ding, X.Y. (2011) Identification of the Geographic Origin of *Dendrobium thyrsiflorum* on Chinese Herbal Medicine Market Using Trinucleotide Microsatellite Markers. *Biological and Pharmaceutical Bulletin*, **34**, 1794–1800.

Zhang, Y.-B., But, P.P.-H., Wang, Z.-T. & Shaw, P.-C. (2005a) Current approaches for the authentication of medicinal Dendrobium species and its products. *Plant Genetic Resources: Characterization and Utilization*, **3**, 144–148.

Zhang, L., Hua, N. & Sun, S. (2008) Wildlife trade, consumption and conservation awareness in southwest China. *Biodiversity and Conservation*, **17**, 1493–1516.

Zhang, Z., Yan, Y., Tian, Y., Li, J., He, J. & Tang, Z. (2015) Distribution and conservation of orchid species richness in China. *Biological Conservation*, **181**, 64–72.

Zhang, G.-N., Zhong, L.-Y., Annie Bligh, S.W., Guo, Y.-L., Zhang, C.-F., Zhang, M., Wang, Z.-T. & Xu, L.-S. (2005b) Bi-bicyclic and bi-tricyclic compounds from Dendrobium thyrsiflorum. *Phytochemistry*, **66**, 1113–1120.

Zhou, D.Q. & Grumbine, R.E. (2011) National parks in China: Experiments with protecting nature and human livelihoods in Yunnan province, Peoples' Republic of China (PRC). *Biological Conservation*, **144**, 1314–1321.

Zhu, H., Cao, M. & Hu, H. (2006) Geological History, Flora, and Vegetation of Xishuangbanna, Southern Yunnan, China. *Biotropica*, **38**, 310–317.

Zhu, H., Wang, H. & Li, B. (1998) The Structure, Species Composition and Diversity of the Limestone Vegetation in Xishuangbanna, SW CHINA. *Gardens' Bulletin Singapore*, **50**, 5–30.

Appendices

Appendix 1. A list of all the individuals recorded and the market codes in a trial market survey of 9 markets in Xishuangbanna in January 2015. Market codes are MO – Mohan, ML – Mengla, MU – Menglun, RS1 – Roadside near Menglun, RS2 – Roadside near Jinghong, JH – Jinghong, NNS – Nannunshan, MH – Menghai, DL – Da luo.

Species	**Market**
Anoectochilus roxburghii	ML
Bulbophyllum affine	NNS
Bulbophyllum aurantiantum	NNS
Bulbophyllum orientale	RS1
Bulbophyllum orientale	NNS
Bulbophyllum repens	NNS
Cleisostoma williamsonii	NNS
Cymbidium bicolor	RS2
Cymbidium lowianum	NNS
Cymbidium mannii	RS1
Cymbidium sinense	RS2
Cymbidium sinense	NNS
Cymbisium aloifolium	NNS
Dendrobium aduncum	RS1
Dendrobium affinum	NNS
Dendrobium aphyllum?	RS1
Dendrobium brymerianum	ML
Dendrobium cariniferum	NNS
Dendrobium cariniferum	ML
Dendrobium chrysanthum	NNS
Dendrobium chrysotoxum	RS2
Dendrobium devonianum	NNS
Dendrobium gratiosissimum	ML
Dendrobium gratiosissimum	RS2
Dendrobium gratiosissimum	NNS
Dendrobium nobile	NNS
Dendrobium nobile	RS2
Dendrobium parsii??	NNS
Dendrobium thyrsiflorum	ML
Dendrobium thyrsiflorum	RS1
Dendrobium thyrsiflorum	RS2
Dendrobium thyrsiflorum	NNS
Dendrobium trigonopus	NNS
Eulophia spectabilis	NNS

Gastochilus obliqus	NNS
Geodorum densiflorum	NNS
Holcoglossum amesianum	NNS
Luisia magniflora	NNS
Luisia morsei	NNS
Oberonia ensifolium?	RS2
Phaius mishmensis	NNS
Pholidota articulata	NNS
Polystacha concreta	NNS
Rhynchostylis retuse	NNS
Triganopus	NNS
Vanda brunnea	RS2
Vanda brunnea	NNS
Vandasilla	NNS

Appendix 2. Each model and the corresponding number of factors and AUC values gained.

Model name	Number of factors	*Dendrobium* Training AUC	*Luisia* Training AUC	*Dendrobium* Test AUC	*Luisia* Test AUC
1	3	0.6728	0.75	0.8539	0.5051
2	4	0.7106	0.77	0.8249	0.5114
3	4	0.7234	0.7557	0.5793	0.5442
4	7	0.7365	0.77	0.606	0.5114
5	10	0.7365	0.7625	0.606	0.5359
6	13	0.7365	0.7905	0.606	0.5762
7	3	0.6982	0.7662	0.8216	0.5448
8	6	0.6982	0.7901	0.8216	0.5756
9	3	0.6982	0.7901	0.8216	0.5756
10	5	0.7474	0.8373	0.9275	0.44
11	5	0.7533	0.8237	0.7315	0.4254
12	8	0.7794	0.8373	0.7665	0.44
13	11	0.7794	0.8415	0.7665	0.4531
14	14	0.7794	0.8437	0.7665	0.4621
15	4	0.7567	0.8415	0.947	0.4531
16	7	0.7571	0.8437	0.9471	0.4621
17	10	0.7563	0.8437	0.9469	0.4621
18	3	0.6917	0.7899	0.8408	0.7346
19	4	0.6971	0.7907	0.8442	0.7294
20	2	0.6932	0.741	0.8324	0.6614
21	16	0.7939	0.9063	0.8665	0.5816

22	17	0.7822	0.8898	0.7569	0.6511
23	15	0.7803	0.8567	0.729	0.6074
24	17	0.7939	0.9063	0.8665	0.5816
25	18	0.8006	0.8999	0.8674	0.5656
26	16	0.7997	0.8732	0.842	0.5278
27	2	0.7323	0.6571	0.6845	0.8426
28	3	0.7322	0.7001	0.6843	0.9099
29	5	0.7949	0.8219	0.8799	0.8747
30	6	0.8007	0.8226	0.8812	0.877
31	4	0.8003	0.7941	0.8777	0.8571
32	18	0.8225	0.9095	0.8851	0.8021
33	19	0.8303	0.9054	0.8824	0.8003
34	17	0.8299	0.8813	0.8778	0.7621
35	19	0.8505	0.9179	0.9413	0.7323
36	20	0.8549	0.913	0.9358	0.725
37	18	0.8551	0.8926	0.9309	0.6902
38	6	0.7949	0.8227	0.8799	0.9107
39	7	0.8007	0.8223	0.8812	0.911
40	5	0.8003	0.8006	0.8777	0.8974
41	19	0.8225	0.9096	0.8851	0.841
42	20	0.8303	0.906	0.8824	0.841
43	18	0.8299	0.8855	0.8778	0.8053
44	20	0.8505	0.917	0.9413	0.7837
45	21	0.8549	0.9126	0.9358	0.7753
46	19	0.855	0.8951	0.9309	0.75
47	7	0.8364	0.8587	0.9512	0.6862
48	5	0.8363	0.8381	0.9497	0.6535
49	19	0.8505	0.9179	0.9413	0.7323
50	20	0.8549	0.913	0.9358	0.725
51	18	0.855	0.8926	0.9309	0.6902
52	19	0.8505	0.9179	0.9413	0.7323
53	20	0.8549	0.913	0.9358	0.725
54	18	0.855	0.8926	0.9309	0.6902
55	7	0.8355	0.8606	0.9536	0.7582
56	8	0.8364	0.8597	0.9512	0.7536
57	6	0.8363	0.8435	0.9497	0.7289
58	20	0.8505	0.917	0.9413	0.7837
59	18	0.8499	0.9126	0.9597	0.7753
60	19	0.855	0.8951	0.9309	0.75
61	20	0.8505	0.917	0.9413	0.7837
62	21	0.8549	0.9126	0.9358	0.7753
63	19	0.855	0.8951	0.9309	0.75

Appendix 3. Each model and the abbreviated factors used.

RAW stands for Landsat bands 4, 5, 6, TCT-BGW stands for tasseled cap transformations for brightness, greenness and wetness. TCT 123 stands for the tasseled cap transformations 4, 5, and 6.

Model name	Description
1	RAW
2	RAW, NDVI
3	NDVI, ALB, MSR, RNDVI
4	RAW, NDVI, ALB, MSR, RNDVI
5	RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW
6	RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123
7	TCT-BGW
8	TCT-BGW, TCT 123
9	RAW, TCT-BGW, TCT 123
10	RAW, NDVI, NIGHT
11	NDVI, ALB, MSR, RNDVI, NIGHT
12	RAW, NDVI, ALB, MSR, RNDVI, NIGHT
13	RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, NIGHT
14	RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
15	TCT-BGW, NIGHT
16	TCT-BGW, TCT 123, NIGHT
17	RAW, TCT-BGW, TCT 123, NIGHT
18	DEM, ASP, SLP
19	DEM, ASP, SLP, ASPSLP
20	DEM, ASPSLP
21	DEM, ASP, SLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123
22	DEM, ASP, SLP, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123
23	DEM, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123

24	DEM, ASP, SLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
25	DEM, ASP, SLP, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
26	DEM, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
27	RDDIST, RDDIR
28	RDDIST, RDDIR, RESIST
29	RDDIST, RDDIR, DEM, ASP, SLP
30	RDDIST, RDDIR, DEM, ASP, SLP, ASPSLP
31	RDDIST, RDDIR, DEM, ASPSLP
32	RDDIST, RDDIR, DEM, ASP, SLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123
33	RDDIST, RDDIR, DEM, ASP, SLP, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123
34	RDDIST, RDDIR, DEM, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123
35	RDDIST, RDDIR, DEM, ASP, SLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
36	RDDIST, RDDIR, DEM, ASP, SLP, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
37	RDDIST, RDDIR, DEM, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
38	RDDIST, RDDIR, RESIST, DEM, ASP, SLP
39	RDDIST, RDDIR, RESIST, DEM, ASP, SLP, ASPSLP
40	RDDIST, RDDIR, RESIST, DEM, ASPSLP
41	RDDIST, RDDIR, RESIST, DEM, ASP, SLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123
42	RDDIST, RDDIR, RESIST, DEM, ASP, SLP, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123
43	RDDIST, RDDIR, RESIST, DEM, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123
44	RDDIST, RDDIR, RESIST, DEM, ASP, SLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
45	RDDIST, RDDIR, RESIST, DEM, ASP, SLP, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
46	RDDIST, RDDIR, RESIST, DEM, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
47	RDDIST, RDDIR, DEM, ASP, SLP, ASPSLP, NIGHT
48	RDDIST, RDDIR, DEM, ASPSLP, NIGHT
49	RDDIST, RDDIR, DEM, ASP, SLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
50	RDDIST, RDDIR, DEM, ASP, SLP, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
51	RDDIST, RDDIR, DEM, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
52	RDDIST, RDDIR, DEM, ASP, SLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT

53	RDDIST, RDDIR, DEM, ASP, SLP, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT, NIGHT
54	RDDIST, RDDIR, DEM, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT, NIGHT
55	RDDIST, RDDIR, RESIST, DEM, ASP, SLP, NIGHT
56	RDDIST, RDDIR, RESIST, DEM, ASP, SLP, ASPSLP, NIGHT
57	RDDIST, RDDIR, RESIST, DEM, ASPSLP, NIGHT
58	RDDIST, RDDIR, RESIST, DEM, ASP, SLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
59	RDDIST, RDDIR, RESIST, DEM, ASP, SLP, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
60	RDDIST, RDDIR, RESIST, DEM, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
61	RDDIST, RDDIR, RESIST, DEM, ASP, SLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
62	RDDIST, RDDIR, RESIST, DEM, ASP, SLP, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT
63	RDDIST, RDDIR, RESIST, DEM, ASPSLP, RAW, NDVI, ALB, MSR, RNDVI, TCT-BGW, TCT 123, NIGHT