



HAL
open science

Transferability of species distribution models for the detection of an invasive alien bryophyte using imaging spectroscopy data

Sandra Skowronek, Ruben van de Kerchove, Bjorn Rombouts, Raf Aerts, Michael Ewald, Jens Warrie, Felix Schiefer, Carol Garzon-Lopez, Tarek Hattab, Olivier Honnay, et al.

► To cite this version:

Sandra Skowronek, Ruben van de Kerchove, Bjorn Rombouts, Raf Aerts, Michael Ewald, et al.. Transferability of species distribution models for the detection of an invasive alien bryophyte using imaging spectroscopy data. *International Journal of Applied Earth Observation and Geoinformation*, 2018, 68, pp.61–72. 10.1016/j.jag.2018.02.001 . hal-02357330

HAL Id: hal-02357330

<https://hal.science/hal-02357330>

Submitted on 29 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Manuscript Number: JAG-D-17-00704R1

Title: Transferability of species distribution models for the detection of an invasive alien bryophyte using imaging spectroscopy data

Article Type: Research Paper

Keywords: *Campylopus introflexus*; heath star moss; hyperspectral; Maxent; dune ecosystem; model transfer

Corresponding Author: Ms. Sandra Skowronek,

Corresponding Author's Institution:

First Author: Sandra Skowronek

Order of Authors: Sandra Skowronek; Ruben Van De Kerchove; Bjorn Rombouts; Raf Aerts; Michael Ewald; Jens Warrie; Felix Schiefer; Carol Garzon-Lopez; Tarek Hattab; Olivier Honnay; Jonathan Lenoir; Duccio Rocchini; Sebastian Schmidtlein; Ben Somers; Hannes Feilhauer

Abstract: Remote sensing is a promising tool for detecting invasive alien plant species. Mapping and monitoring those species requires accurate detection. So far, most studies relied on models that are locally calibrated and validated against available field data. Consequently, detecting invasive alien species at new study areas requires the acquisition of additional field data which can be expensive and time-consuming. Model transfer might thus provide a viable alternative. Here, we mapped the distribution of the invasive alien bryophyte *Campylopus introflexus* to i) assess the feasibility of spatially transferring locally calibrated models for species detection between four different heathland areas in Germany and Belgium and ii) test the potential of combining calibration data from different sites in one species distribution model (SDM).

In a first step, four different SDMs were locally calibrated and validated by combining field data and airborne imaging spectroscopy data with a spatial resolution ranging from 1.8 m to 4 m and a spectral resolution of about 10 nm (244 bands). A one-class classifier, Maxent, which is based on the comparison of probability densities, was used to generate all SDMs. In a second step, each model was transferred to the three other study areas and the performance of the models for predicting *C. introflexus* occurrences was assessed. Finally, models combining calibration data from three study areas were built and tested on the remaining fourth site. In this step, different combinations of Maxent modelling parameters were tested.

For the local models, the area under the curve for a test dataset (test AUC) was between 0.57-0.78, while the test AUC for the single transfer models ranged between 0.45-0.89. For the combined models the test AUC was between 0.54-0.9. The success of transferring models calibrated in one site to another site highly depended on the respective study site; the combined models provided higher test AUC values than the locally calibrated models for three out of four study sites. Furthermore, we also demonstrated the importance of optimizing the Maxent modelling

parameters. Overall, our results indicate the potential of a combined model to map *C. introflexus* without the need for new calibration data.

Opposed Reviewers:

Dear Sir or Madam,

We revised our manuscript entitled “Transferability of species distribution models for the detection of an invasive alien bryophyte using imaging spectroscopy data” according to the reviewers suggestions.

We hope that our manuscript meets the requirements of the International Journal of Applied Earth Observation and Geoinformation and look forward to hearing from you.

Yours sincerely,

Sandra Skowronek
(on behalf of all coauthors)

Reviewer 1

This manuscript presents a study on the transferability of models that predict the occurrence probability of an alien moss species in four different sites in Belgium and Germany. The authors present a rare dataset of having four hyperspectral datasets for four the test sites and similar ground truthing campaigns. The aim of the study was to test whether models created for one site would be yield better or similar results for another site, which was partly the case. I believe that this study is a good example for what is possible if multiple datasets are available, but they also did not really convince me that it makes sense to train a model at a very different site and then apply this model somewhere else (called simple transfer models). More interesting are the combined optimized results presented which show that in most of the cases the results were better than the original model, but also only marginally. Nevertheless, this study is an important step towards the idea of applying a well trained model to multiple sites in order to minimize costs and errors.

The manuscript is in general well written, the english was improved a native speaker (I guess? see the acknowledgements) and the content is sound. Hence I vote for minor revisions.

Reviewers comments	Answers	Changes in the manuscript
<p>In table 1, why do you use less plots for calibration then for validation? At least for Sylt and Averbode you have a Cal/Val ratio of 1/3. Usually people apply an 80/20 or 60/40 ratio for the Cal/Val data. Ah - I just saw that you the Val-Data do consist of p and a plots...okay. but still the ratio is not standard. Please explain.</p>	<p>We tried to apply a sampling scheme that is as efficient as possible. In a former publication we could show that reducing the number of calibration plots from 57 presence plots to about 22 presence plots for the Sylt study area did not significantly change the results. A similar finding was made in another study with spectral data. A relatively low number of calibration plots was thus found to be sufficient, if it covers the variety present in the study area. For validation plots, on the other side, the more the better. The more we have, the easier it is to understand how good the model really is and where it fails. We thus sampled as many validation plots as we could within a reasonable time frame of one or two weeks of field work (1 person). Also, for my dissertation I compared about 20 previous studies on mapping invasive species using hyperspectral remote sensing data, and did not really see any consistent ratio of cal/val plots that all studies used. Some used less cal than val plots, some the same amount, and some less val than cal plots.</p>	<p>Added sentence: <i>While a relatively low number of calibration plots was found to be sufficient (see Skowronek et al. 2017), we used as many plots as we could gather within a reasonable timeframe for validation.</i></p>
<p>Another question related to table 1. You have sampled "absence" plots, however Maxent generates absence data itself, right? These are</p>	<p>Yes, Maxent can generate background data itself, but this is not the same as absence data. Maxent generates random points, which are only used to calibrate the model. But to validate the model, we need real</p>	<p>Added sentence: <i>In all four study areas, we collected presence data to calibrate the model and presence/absence data to validate the prediction.</i></p>

<p>called "background", which you also mention in L168 (3000 random points). Why do you need your absence plots if the algorithm will not use them?</p>	<p>absences. We added a sentence to make this clearer.</p>	
<p>L192-194 "assessed the amount of spectral variance in the different study sites by calculating the standard deviation across the whole spectrum for the background and calibration datasets" why did you do this? please explain to the reader, it might not be obvious to everybody. I also do not understand how you did this technically. Does this mean you calculated the sd in a window of 3x3? was the variability in the spectral range measured or spatially for each band? what for? Is little variation good or not so good for you aim?</p>	<p>The spectral variance was calculated to have an estimate of the amount of variation/heterogeneity in both the background and calibration samples as this is known to have an effect on model transferability. It was obtained by calculating for every band the standard deviation over all background and calibration points, respectively. Please see paragraph 3 of section 4.1 for more information on the effect of spectral variation on model transferability.</p>	<p>We added "for each band and" on line 203 to clarify how this spectral variation was calculated</p>
<p>One more question on the background points. In line 204 you state that you use the default (10.000 points) but later on L212 and 213 you talk about 9.000 and 3.000 points. I find this pretty confusing. Could you try to make this part more clear?</p>	<p>We used the default settings, so the 10,000 background points for Steps I and II. For step III however, we are combining the data and thus also the background points. We tried using the full set of $3 \times 10,000 = 30,000$ background points, but this has led to excessive computing time, so we reduced it by using a random subset of 3,000 points from each dataset, so that we have $3 \times 3,000 = 9,000$ points. Indeed one sentence in the manuscript was wrong, stating that we always used 3,000 points. We corrected this sentence. We also moved the relevant sentence up in the paragraph to make it less confusing and added another sentence.</p>	<p>We added the following explanations:</p> <p>"The background sample for each study site consisted of a large number of random points located within the biotope types of each study area"</p> <p>"For each of these models (192 in total) we used a selection of 9,000 background points obtained by combining a randomly selected subset of 3,000 background points per site. The lower number of background points was used due to limits in computing time."</p>
<p>L 215: "For Step I, we also derived the confusion matrices using kappa</p>	<p>Yes, something got mixed up in that sentence. We did use kappa as a threshold to derive the presence-</p>	<p>"For Step I, we also derived the presence / absence map from the probability maps</p>

<p>statistic as threshold to derive presence-absence maps from the probability maps." Here is also something wrong. You did not derive the confusion matrix using kappa, but of course the info correctly classified P/As. Did you optimize kappa for creating the probability maps?</p>	<p>absence map (there are multiple thresholds in the "evaluate" function of the dismo package and kappa worked best). The sentence was revised.</p>	<p>using kappa as threshold. We then also derived the confusion matrices and compared the overall accuracies (OA)."</p>
<p>I also do not get the meaning of the spectral variability plots in Figure 4 or better to say the relevance. for the calibration points, these are then locations at which the moss was present? Is the spectral variability</p>	<p>The spectral variability plots were calculated to assess the range of conditions which appeared in the different study areas as it is known that this can have an effect on the performance when transferring models. Please see paragraph 3 of 4.1 for an interpretation on the effect of this spectral variability.</p>	
<p>In Line 423 you mention that choosing an independent validation dataset is important. Independent in which sense? Validation data is always not the the same data as the calibration data, so the word "independent" is, to me at last, confusing. Do you mean spatially independent? please explain.</p>	<p>With independent we mean that a different sampling scheme is applied. While it is necessary to use preferential sampling for the calibration dataset in order to sample all possible variations of the species occurrence, for the validation data, a random sampling approach should be the goal. In many studies, however, this is not the case, that's why we point it out.</p>	<p>Changed to "separate independent"</p>
<p>L217 OAC -> OA (I never saw OAC before anywhere) please correct throughout the manuscript</p>	<p>Yes, indeed.</p>	<p>Done</p>

<p>Fig.3 and all other maps. Your legend looks strange and they all differ between the maps, making them effectively not really comparable. I suggest having a unified legend for all maps and turn the 1.3 e-12 simply into a zero. You should also better state in the caption that predictions are in percentage of occurrence probability of c.introflexus</p>	<p>We agree that a common legend will make the comparison easier - we had first stretched it to make the differences within one prediction more visible, but indeed a unified legend is better for comparison. We modified the maps accordingly and revised the figure caption.</p>	<p>Done, see new Figure 4 and Maps in Supplement</p>
<p>L243: remove blank space after "details)"</p>	<p>Ok</p>	<p>Done</p>
<p>L277: "calibration AUC values (between 0.94 and 0.96)" maybe you should state that these values can be found in the appendix?</p>	<p>This seems to be a misunderstanding, as there is no figure in the appendix that contains the calibration AUC values.</p>	
<p>L290: write four instead of 4</p>	<p>Ok</p>	<p>Done</p>
<p>L291: were combined instead of was combined?</p>	<p>Ok</p>	<p>Done</p>
<p>L374: if these are lower...lower then what? please explain</p>	<p>There is no fixed value we can give here as it depends on the target species as well as the other vegetation present in the pixel.</p>	<p>Changed "lower" to "low"</p>
<p>L396: please write out lq and lqhp - it is not clear here what these mean.</p>	<p>Ok</p>	<p>Done, "mainly lq (linear and quadratic) and lqhp (linear, quadratic, hinge and product)"</p>
<p>L418: One should always prevent/test for sampling bias. <- talking of sampling bias - this was not mentioned in the discussion at all. As you state here that this is an important issue, shouldn't you include the sampling bias in your discussion as well (in 4.1 and 4.2)? and also decide on whether to use prevent or test, prevent/test looks awkward.</p>	<p>Indeed. Prevent for sampling bias makes more sense, so we deleted "/test for". The sampling bias (spatial sorting bias, Hijmans et al. 2012) was calculated and the values are given in chapter 3. We also added a sentence to the discussion in chapter 4.4. We think it fits better here than in chapters 4.1 and 4.2, as it is a factor contributing to uncertainty, even if it was found to be very low for this study.</p>	<p>Changed to "mitigate sampling bias". Added sentence: And while there was no sampling bias (spatial sorting bias, see chapter 2.3) for Sylt and Liereman, there was a relatively small bias for Kalmthout and Averbode.</p>

Reviewer 2

This well-written manuscript deserves publication in the JAG journal. The objective was to evaluate the transferability of Maxent classification models for detecting one invasive species using hyperspectral APEX data. The manuscript adds to the previous publications by the authors in other journals related to the topic.

In terms of revision, some important methodological aspects related to atmospheric correction and airborne data acquisition are missing, as detailed in my comments. The advantages of the Maxent classification over other conventional classification techniques require also clarification. In order to support the discussion of results, it is important to add reflectance spectra of the invasive species and of the background to inspect for spectral differences across sites. Finally, please, insert a Conclusions section.

I added my comments sequentially with page numbering. However, it was just at the end of reading that I found out solutions and responses in the supplementary material for [comments 6 and 11](#). Therefore, I suggest that the authors migrate some reflectance spectra (species and background) from the supplementary material (Last two figures; insert unit for wavelength) into Results, and discuss the spectra in the new proposed section (see comments 6 on the Results section).

<p>1. Abstract: It should be continuous (without paragraphs).</p>	<p>Ok</p>	<p>Done</p>
<p>2. Lines 141 to 148: In the discussion of Table 1, please, highlight the gaps between the flight campaigns and fieldwork activities (up to two years for the Kalmthout site) as well as the differences in spatial resolution between the campaigns. Add a line with the time of image acquisition (GMT) since significant differences in solar zenith angle can affect the reflectance across sites of the invasive species and of the background. Justify in the text why you think that such problems in the experimental design do not influence on your investigation. Are there any differences in phenological stages for the invasive species considering</p>	<p>Ok, we added the time of the image acquisition in table 1 and highlighted the differences in spatial resolution and gaps between field work and flight campaigns in the accompanying paragraph.</p> <p>We did not observe a significant phenological difference between the dates when the image data was acquired. <i>C. introflexus</i> is a bryophyte that does not usually undergo major phenological changes between July and September, when our image data was acquired. We only observed changes later in the year after major rainfalls.</p> <p>Nevertheless, the time gap between image acquisition and field campaign could lead to a slight over or underestimation, especially on low cover plots. However, we expect the changes from one year to the next to be relatively small at the current stage of invasion.</p>	<p>Added text The spatial resolution was highest for Sylt and lowest for Kalmthout. While for Sylt and Averbode, the calibration data was collected less than one month before or after the flight campaigns took place, the calibration data for Liereman was collected about one year after the flight, and the data for Kalmthout only two years after the flight.</p> <p>Sentence added to section 4.4: We did not observe a significant phenological difference between the dates when the image data was acquired.</p> <p>and</p> <p>And the time gap between image acquisition and field campaigns of about one year for Liereman and two years for Kalmthout may have led to a</p>

<p>the dates of image acquisition? Please, clarify in the text.</p>	<p>We thus estimate that neither the time gap between field campaign and image acquisition nor the timing of the image acquisition have a negative impact on our analysis. As for the effect of the spatial differences, please see discussion in section 4.1.</p>	<p>slight over or underestimation of the abundance of <i>C. introflexus</i> on single plots.</p>
<p>3. Line 143: In a study of imaging spectroscopy, it is important to mention the spectral resolution (bandwidth) of the APEX in the VNIR and SWIR spectral intervals. There was just a general mention about that in Abstract.</p>	<p>We agree and modified the descriptive paragraph accordingly.</p>	<p>Airborne imaging spectroscopy data, acquired by the Airborne Prism EXperiment (APEX) spectrometer was used within this study. APEX is an airborne imaging spectrometer which collects information between 380nm and 2500nm with a Full Width at Half Maximum (FWHM) ranging from 3 nm to 12 nm (after spectral binning) in the visible and near-infrared spectral region, and from 9 nm to 12 nm in the SWIR region. Apex data were acquired</p>
<p>4. Line 144: Please, add a few lines to clarify how the atmospheric correction was performed over the APEX data. For instance, what was the APEX band used for water vapor determination (940 nm or 1140 nm)? What was the selected atmosphere and aerosol models? How was the visibility determined? Is the standard APEX processing based on MODTRAN4? In short, provide more details on this important methodological step.</p>	<p>We added the following description:</p>	<p>The data were geometrically and atmospherically corrected using the standard processing applied to APEX (Sterckx <i>et al.</i>, 2016; Vreys <i>et al.</i>, 2016) at VITO's Central Data Processing Center. The processing chain is based on the MODTRAN4 software, in which the model atmosphere was set to "mid-latitude summer" and the employed aerosol type was "rural". The main atmospheric parameters (water vapor content and visibility) were derived from ground-based measurements using a Microtops sunphotometer and spectral ground control points, measured by means of an ASD spectrometer, were used as reference spectra. Where Microtops and/or ASD measurements were not available, all parameters were iteratively tuned to ensure a minimum spectral distortion in the water vapor absorption bands jointly with a high consistency between APEX spectra and reference spectra</p>

		from available spectral libraries. After atmospheric correction,
<p>5. Lines 159-165: Please, mention the advantages of Maxent compared to others conventional classifiers that can work also as a one-class classifier (e.g., SAM, MTMF etc.). What are the main differences between them to justify Maxent selection? For instance, the concept of background is also used in the MTMF.</p>	<p>A general advantage of Maxent is that it is well known to most ecologists, and easy to manipulate using the freely available standalone software. A specific advantage to SAM is that we only need to sample presences and absences of the target species, and do not need a lot of data for each endmember class present in the study area.</p> <p>As far as we know, MTMF is only readily available through ENVI, which is an expensive commercial software, which is usually not available to people working in nature conservation.</p> <p>In Skowronek et al. 2017 (Ecological Informatics) we also compared the performance of Maxent to SVM and BRT and found no major differences.</p>	<p>Sentences added: General advantages of Maxent are that it is relatively easy to use and freely available, either through R or through its standalone software. Moreover, as a one-class classifier, it only requires presence data to be collected in the field for model calibration, which greatly reduces the amount of field work necessary. In Skowronek et al. 2017 we compared the performance of Maxent, Support Vector Machine and Boosted Regression Trees and found that all three classifiers allowed for the detection of the two target species with similar success rates.</p>
<p>6. Line 228: To strength the manuscript and facilitate comprehension of the results, I suggest that the authors add a new short section (e.g., 3.1. Spectral reflectance of the invasive species and background across sites) to show, for each site, average APEX reflectance spectra of the plots having the invasive species and average curves of the background. I think this important to highlight the major spectral features of the invasive species detected by the sensor; the eventual spectral differences across sites due to phenology and gaps in data acquisition; and the eventual differences across sites between the backgrounds. These factors should be considered in the Discussion section.</p>	<p>We agree that this would enhance readability and thus added a new section. 3.1 Spectral reflectance of the invasive species and background across sites and created a new figure 3 which is a combination of former figure 4 and one of the supplement figures.</p> <p>Furthermore, spectral differences across sites due to phenology and gaps in data acquisition are now discussed in the uncertainties section (4.4) in greater detail. Please see answers to comment 13.</p> <p>For differences between backgrounds we also added more information, please see answers to comment 11.</p> <p>Additionally we also remade the figure in the supplement by inserted units for the wavelength and adding a legend.</p>	<p>Text added to section 3.1</p> <p>Figure 3 shows the mean reflectance and the spectral variability for all four study sites for the calibration as well as the background data. It is important to point out that the calibration spectra are averages of all calibration plots, which may contain very high or very low amounts of the target species. For the background data, it is important to note that this data may eventually also contain a few single data point where the target species is present, as this data is randomly selected. Overall, Sylt had higher mean reflectance values in the VIS/NIR, both for the calibration and background points. Averbode had the highest mean reflectance in the SWIR. Kalmthout on the other hand had consistently lower reflectance values. The spectral variability within the calibration datasets was overall highest for Liereman. For Sylt and Averbode, we observed relatively high variability in the VIS/NIR and in the SWIR,</p>

		respectively (Fig. 3). For the background points, Sylt showed the highest variability, followed by Liereman in the NIR and by Averbode in the SWIR. The Kalmthout calibration and background spectra contained only little spectral variation.
7. Line 237: In Figure 3, please, insert the North arrow inside the figure. Add geographical coordinates.	Ok	Done, see Figure 4
8. Line 240: Please, define the abbreviations and acronyms when they first appear in the text. After that, just use the abbreviations to save space in the text (for instance, SWIR).	Ok	Done
9. Line 251: In Figure 4, what represents the three shaded portions of the figure? Interestingly, the bands with the largest values of importance are coincident with spectral intervals close to strong water vapor absorptions. Maybe you can use the shade to show spectral regions excluded from the analysis due to atmospheric absorption and noise.	The shaded part is just a graphical feature to make the figure more readable. We changed it according to your suggestion.	Done, see Figure 5
10. Line 289: Figure 7 was misplaced in the text just after Figure 4. Please, correct.	Oh yes indeed, will be corrected.	Figures were placed in correct order.
11. Line 301: In the true color composites of Figure 6 and in the description between lines 169 and 170, it seems that the background composition (e.g., sand dunes, grasslands) for Maxent classification is not the same across the sites. If it is correct, please, add a	Indeed the background composition is not identical for all four study sites - would be strange if it was. We added a paragraph to the methods section on the most frequently occurring background vegetation. However, as there is much more than one vegetation type in each background, we do not want to add a line to Table 1, as this would be too much of a simplification of the real situation to reduce each area to one predominant	Regarding the background vegetation, the following paragraph was added: While the most abundant vegetation types on Sylt include <i>Empetrum nigrum</i> dominated heathland making up about 2/3 of the study area, other important vegetation types include grey dunes vegetation, <i>Erica-tetralix</i>

<p>new line in Table 1 to clarify the predominant background for the Maxent classification on each site. Are the authors modelling the same invasive species over very different reflective backgrounds or are they mixed on each scene? How the reflectance of the predominant background affects classification results? Please, clarify these aspects in the text.</p>	<p>background type (especially for the Belgian study sites, there are always several vegetation types making up nearly similar amounts of the background area, which cannot be easily summed up in one line). Also different classification schemes were used for the vegetation mapping in the different study sites, which further complicates the comparison - anyone interested in the details needs to consult the original biotope maps and classification schemes (references are given). How the different background affects the results is being discussed in section 4.1.</p>	<p>and <i>Ammophila arenaria</i> dominated areas. For the Belgian study sites, the most abundant biotope types sites include vegetation types dominated by <i>Caluna vulgaris</i>, <i>Molinia caerulea</i> and <i>Erica tetralix</i>.</p> <p>One Figure was moved from the supplementary material to the new section 3.1 and explanations were added, see answer to comment 6.</p>
<p>12. Line 301: In Figure 6, there is space to accommodate the legend at the upper left side of the figure without obliterating the results.</p>	<p>Ok</p>	<p>Done</p>
<p>13. Line 425, discussion of uncertainties: In addition to the listed uncertainties, clarify if the differences in spatial resolution between the campaigns affected the results. Do the same for the time of image acquisition (it should mentioned in Table 1; GMT) and for possible differences in phenological stages of the species across sites.</p>	<p>Spatial resolution: this point is discussed in section 4.1, but we added an additional sentence to section 4.4 (uncertainties) pointing to that section.</p> <p>Difference for time of image acquisition: We added a line in Table 1 with the time in GMT and added a sentence to section 4.4</p> <p>Possible differences in phenological stages: The invasive bryophyte <i>C. introflexus</i> did not undergo any major phenological change across time, neither was there a pronounced difference across sites. We added a sentence to section 4.4.</p>	<p>added sentences: We did not observe a significant phenological difference between the dates when the image data was acquired or between the different study sites. All imagery was acquired around noon local time (see table 1). We thus estimate that the timing of the image acquisition did not have any major impact on the results. A factor that did affect the results in a significant way was the different spatial resolution (see section 4.1 for details).</p>
<p>14. Line 458: Define EnMAP, if not done before.</p>	<p>Ok</p>	<p>Done, "EnMAP (Environmental Mapping and Analysis Program)"</p>
<p>15. Please, it is very important to add a Conclusions section.</p>	<p>We added a conclusions section.</p>	<p>Added section: 5 Conclusions In this study we successfully transferred species distribution models for <i>Campylopus introflexus</i> which were calibrated at different sites using airborne imaging spectroscopy as explanatory variables. Our results demonstrate that model</p>

		<p>transfer success was determined by a combination of i) the spectral heterogeneity of the calibration dataset and how adequately it represents the spectral heterogeneity of the target dataset, ii) the spatial resolution of the calibration dataset as well as the iii) parametrization and complexity of the used model. As more remote sensing datasets become available, those techniques can improve model results or be used to avoid additional time-consuming field work. This is especially relevant for a time- and cost-efficient repetitive monitoring of invasive plant species, as it is impossible to frequently map invasive species over large scales using traditional field mapping techniques. However, we do need this type of information to be able to assess the spread of invasive species and manage them accordingly. This study therefore explores challenges related to model transfer and gives practical recommendations regarding data collection, data analysis and evaluation of the results.</p>
<p>16. References: Please, revise them for missing information. Some conference papers could be replaced by journal papers from the same authors. This is the case of the study by Müllerová et al. (2016) from the ISPRS conference, which was published in the JAG journal: Müllerová et al. (2013). Remote sensing as a tool for monitoring plant invasions: Testing the effects..... Int. J. Applied Earth Observation and Geoinformation, 25: 55-65.</p> <p>There are other references related to the topic in the</p>	<p>Thanks for those suggestions! The references were revised for missing information.</p> <p>We integrated Robinson et al. 2016 and Fernandes et al. 2014. Müllerová et al. 2013 does not support our statement, we decided to cite Müllerová et al. 2017 instead as it fits best.</p>	<p>Sentence added to the discussion: However, a similar transferability approach could be applied to multispectral satellite data such as worldview-2 or 3, which are readily available for larger areas, and have proven to be useful for mapping certain invasive plant species (e.g. Robinson et al. 2016, Fernandes et al. 2014).</p>

JAG journal that you can eventually consider in the literature review or discussion of results:

- Robinson et al. (2016). Testing the discrimination and detection limits of WorldView-2 imagery on a challenging invasive plant target. *Int. J. Applied Earth Observation and Geoinformation*, 44: 23-30.
- Fernandes et al. (2014). Optimal attributes for the object based detection of giant reed in riparian habitats: A comparative study between Airborne High Spatial Resolution and WorldView-2 imagery. *Int. J. Applied Earth Observation and Geoinformation*, 32: 79-91.
- Dorigo et al. (2012). Mapping invasive *Fallopia japonica* by combined spectral, spatial, and temporal analysis of digital orthophotos. *Int. J. Applied Earth Observation and Geoinformation*, 19: 185-195.

Please, verify if these references are adequate or just ignore them, if the case.

Highlights

- An invasive alien bryophyte was mapped on four sites using hyperspectral data.
- Transferred species distribution models sometimes outperformed local models.
- High potential of combining field data to create a more general model.
- Optimizing model parameters is very important for a successful transfer.

Transferability of species distribution models for the detection of an invasive alien bryophyte using imaging spectroscopy data

1 Sandra Skowronek¹, Ruben Van De Kerchove², Bjorn Rombouts³, Raf Aerts^{3,4}, Michael Ewald⁵,
2 Jens Warrie⁴, Felix Schiefer¹, Carol Garzon-Lopez⁶, Tarek Hattab⁶, Olivier Honnay⁴, Jonathan
3 Lenoir⁶, Duccio Rocchini⁷, Sebastian Schmidtlein⁵, Ben Somers³, Hannes Feilhauer¹

4 ¹Institute of Geography, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany*

5 ²Flemish Institute for Technological Research (VITO), Mol, Belgium

6 ³Division of Forest, Nature and Landscape, Department of Earth and Environmental Sciences, KU Leuven,
7 Leuven, Belgium

8 ⁴Ecology, Evolution and Biodiversity Conservation Section, Department of Biology, KU Leuven, Leuven,
9 Belgium

10 ⁵Institute of Geography and Geoecology, Karlsruhe Institute of Technology, Karlsruhe, Germany

11 ⁶Unité de Recherche « Ecologie et dynamique des systèmes anthropisés » (EDYSAN, FRE 3498, CNRS-UPJV),
12 Université de Picardie Jules Verne, Amiens, France

13 ⁷Department of Biodiversity and Molecular Ecology, Fondazione Edmund Mach, San Michele all'Adige, Italy

14

15 * **Correspondence:**

16 Sandra Skowronek

17 Sandra.skowronek@fau.de

18 **Keywords:** *Campylopus introflexus*, heath star moss, hyperspectral, Maxent, dune ecosystem,
19 model transfer

20 **Abstract**

21 Remote sensing is a promising tool for detecting invasive alien plant species. Mapping and monitoring those
22 species requires accurate detection. So far, most studies relied on models that are locally calibrated and
23 validated against available field data. Consequently, detecting invasive alien species at new study areas
24 requires the acquisition of additional field data which can be expensive and time-consuming. Model transfer
25 might thus provide a viable alternative. Here, we mapped the distribution of the invasive alien bryophyte
26 *Campylopus introflexus* to i) assess the feasibility of spatially transferring locally calibrated models for species
27 detection between four different heathland areas in Germany and Belgium and ii) test the potential of
28 combining calibration data from different sites in one species distribution model (SDM). In a first step, four
29 different SDMs were locally calibrated and validated by combining field data and airborne imaging
30 spectroscopy data with a spatial resolution ranging from 1.8 m to 4 m and a spectral resolution of about 10
31 nm (244 bands). A one-class classifier, Maxent, which is based on the comparison of probability densities,
32 was used to generate all SDMs. In a second step, each model was transferred to the three other study areas
33 and the performance of the models for predicting *C. introflexus* occurrences was assessed. Finally, models
34 combining calibration data from three study areas were built and tested on the remaining fourth site. In this
35 step, different combinations of Maxent modelling parameters were tested. For the local models, the area
36 under the curve for a test dataset (test AUC) was between 0.57-0.78, while the test AUC for the single

37 transfer models ranged between 0.45-0.89. For the combined models the test AUC was between 0.54-0.9.
38 The success of transferring models calibrated in one site to another site highly depended on the respective
39 study site; the combined models provided higher test AUC values than the locally calibrated models for three
40 out of four study sites. Furthermore, we also demonstrated the importance of optimizing the Maxent
41 modelling parameters. Overall, our results indicate the potential of a combined model to map *C. introflexus*
42 without the need for new calibration data.

43 **Declarations of interest:** none

44 **Funding**

45 This study is part of the project DIARS (Detection of Invasive plant species and Assessment of their impact on
46 ecosystem properties through Remote Sensing), which is funded by the ERA-Net BiodivERsA. The respective
47 national funders are the Agence Nationale de la Recherche [grant number ANR-13-EBID-0006], the Belgian
48 Federal Science Policy Office [grant numbers SR/67/315 and BR/132/A1/DIARS-BE] and the German Research
49 Foundation [grant numbers FE 1331/3-1 and SCHM 2153/9-1].

50 **Acknowledgments**

51 We would like to thank the nature conservation authority of Northern Friesland and the private land owners
52 for granting us permission to conduct research in the protected dune heathlands on the island of Sylt. For the
53 Belgian study sites, we thank Natuurpunt and the Research Institute for Nature and Forest (INBO) for giving
54 us access to the different study sites as well as providing us with existing data sets. Many thanks to Phillip
55 Brodrick for proofreading the manuscript.

56

57 **1 Introduction**

58 Remote sensing is a promising tool for the detection and monitoring of invasive alien plant species (Bradley,
59 2013). Invasive alien plants can be identified from different remote sensing platforms like unmanned aerial
60 vehicles (UAVs) (e.g. Michez et al., 2016; Müllerová et al., 2017), airborne platforms (e.g. Cheng, 2007; Mirik
61 et al., 2013; Skowronek et al., 2017a, 2017b) or from satellites (e.g. Proctor et al., 2012; Somers and Asner,
62 2013). In particular, imaging spectroscopy data hold a high potential due to their high spectral resolution,
63 which allows differentiating characteristic species from the surrounding vegetation (He et al., 2011; Huang
64 and Asner, 2009).

65 The large majority of studies on mapping the distribution of invasive alien plant species have relied on
66 models that are calibrated (trained) and validated (tested) using field data specific to a particular location
67 (referred to hereafter as site-specific models). The spatial transfer of species distribution models might be a
68 useful tool for mapping the distribution of invasive alien species in the following two situations: when limited
69 resources are available to carry out field work and remote sensing data are available for a larger area and
70 when the detection of recently invaded sites is of interest, but manual search of the area to calibrate a site-
71 specific model is not feasible. The transferability of species distribution models has been investigated in
72 several recent studies which mainly evaluated the performance of different algorithms (Duque-Lazo et al.,
73 2016; Heikkinen et al., 2012; Wenger and Olden, 2012), or focused on the tuning of model settings (e.g.
74 Moreno-Amat et al., 2015; Muscarella et al., 2014). While most of these studies relied on climatic,
75 topographic, soil, or similar data as predictor variables, few studies have examined the success of model
76 transfer using spectral data (with the exception of Tuanmu et al., 2011, for example). However, He et al.
77 (2015) highlighted the potential of airborne hyperspectral remote sensing data in species distribution
78 modelling due to its high spectral and relatively high spatial resolution as well as a high spatial coverage.

79 One main challenge for model transferability is that individual models may be limited by site-specific
80 information, causing the model to be overfit to a certain location (Anderson and Gonzalez, 2011; Moreno-
81 Amat et al., 2015). Jiménez-Valverde et al. (2011) suggest combining data from several locations to calibrate
82 an overall species distribution model for invasive alien species to predict on a new area. One of the most
83 frequently used algorithms for species distribution modelling is Maxent (Merow et al., 2013). Two important
84 parameters govern the functionality of Maxent: the regularization multiplier (β), and the number of
85 considered feature classes to construct the model (fc) (Elith et al., 2011; Merow et al., 2013; Radosavljevic
86 and Anderson, 2014). To reduce over-fitting and to generate a simpler and potentially more transferable
87 model, we can increase β and limit fc . Elith et al. (2011) mention that Maxent is relatively stable when
88 dealing with correlated input variables compared with other methods (for example stepwise regression).
89 Consequently, there is less of a need for pre-selection of predictor variables when using Maxent. However,
90 the selection of model metaparameters is important for Maxent to perform optimally. Warren and Seifert
91 (2010) proposed to use information criteria for model selection in order to avoid selecting overly complex
92 models.

93 In this study, we evaluated the transferability of Maxent models based on airborne imaging spectroscopy for
94 detecting the invasive alien bryophyte *Campylopus introflexus*. This species was classified to be one of the
95 100 worst invaders in Europe (DAISIE, 2015). As a relatively small and inconspicuous species lacking
96 characteristic features like colourful flowers, it was chosen to show whether remote sensing is a useful tool

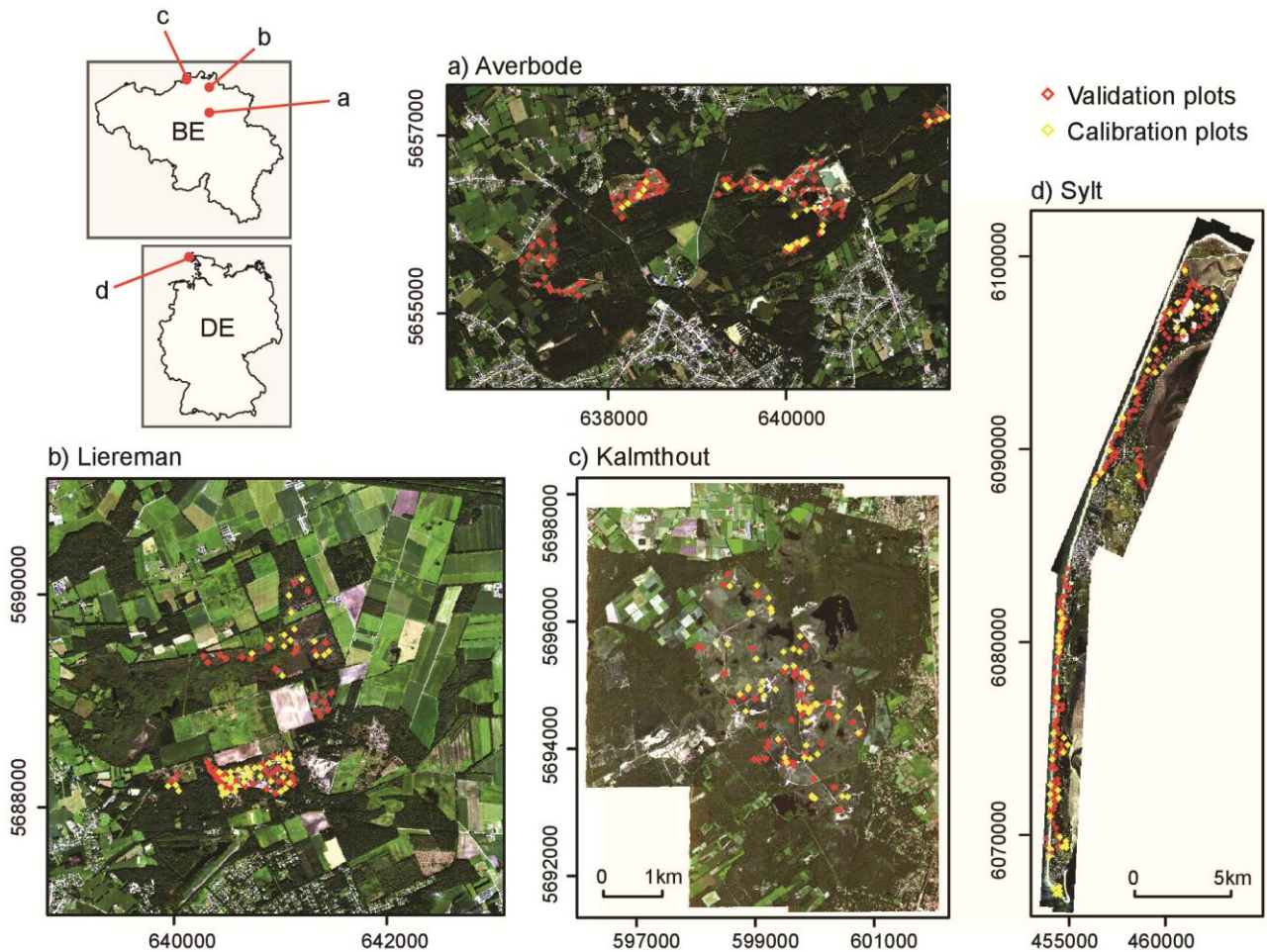
97 to detect such a species. Also, bryophytes constitute a largely understudied group of species among the
98 invasive alien plants (Essl et al., 2014; Mateo et al., 2015).

99 We use four different study sites located in Germany and Belgium where we collected independent
100 calibration and validation datasets. This study builds further on the work of (Skowronek et al., 2017b) which
101 used Maxent modelling (using default settings) to map the distribution of *C. introflexus* based on airborne
102 imaging spectroscopy on the island of Sylt, Germany. Our research questions are: (1) How well can we
103 transfer models from one site to another? (2) Does combining data from multiple study sites improve the
104 prediction? (3) How do parameter settings affect model performance?

105 **2 Materials and Methods**

106 **2.1 Study areas and study species**

107 We used three sites from Belgium in this study: the dune areas within Kalmthoutse Heide (Ka; 51°24'00" N,
108 4°26'00" E), Landschap de Liereman (Li; 51°20'00" N, 5°01'00" E), and Averbode Bos & Heide (Av; 51°02'30"
109 N, 4°58'00" E). A fourth site, the dune areas of the island of Sylt (Sy; 54°55'00" N, 8°20'00" E), was located in
110 north-western Germany. All study sites have a temperate climate. The study sites in Belgium are located 30-
111 60 km from one another, and have a distance of about 450-500 km to the island of Sylt. All four study sites
112 are shown in Figure 1.



113

114 Figure 1: Study areas (a) Averbode Bos & Heide (Av); (b) Landschap de Liereman (Li); (c) Kalmthoutse Heide
115 (Ka) and (d) Sylt (Sy). A true colour composite derived from the APEX data is used as background

116 Within each study site, we limited our area of interest, using the available biotope maps (Instituut voor
117 Natuurbehoud, 2016; LEGUAN, 2012; Natuurpunt, 2012) to identify areas where our target species
118 *C. introflexus* might be present – mainly dune areas and a few grassland areas. The dunes on the island of
119 Sylt are mainly coastal dunes, Kalmthout consists of inland dunes, and the majority of Averbode and
120 Liereman was recently converted into heathland by cutting down planted pine forests. **While the most**

121 abundant vegetation types on Sylt include *Empetrum nigrum* dominated heathland making up about ⅓ of the
122 study area, other important vegetation types include grey dunes vegetation, *Erica-tetralix* and *Ammophila*
123 *arenaria* dominated areas. For the Belgian study sites, the most abundant biotope types sites include
124 vegetation types dominated by *Caluna vulgaris*, *Molinia caerulea* and *Erica tetralix*. Sylt covers an area of
125 24.2 km² and the Kalmthout study site covers 8.0 km². The two other sites, Liereman and Averbode, are
126 significantly smaller and cover 1.4 km² and 1.2 km², respectively.

127 All study sites show high degrees of invasion by the heath star moss, *C. introflexus*. First introduced to Europe
128 in 1941 (Richards, 1963), *C. introflexus* is known to mainly invade coastal and inland dunes and reduce the
129 diversity of the native dune communities and potentially change succession rates (Biermann and Daniels,
130 1997; Ketner-Oostra and Sýkora, 2004). *Campylopus introflexus* prefers acidic soils and benefits from
131 nitrogen deposition. A promising management approach is to cover *C. introflexus* with sand through the re-
132 activation of dunes (Boxel et al., 1997; Ketner-Oostra and Sykora, 2000), but to date, almost no attempts
133 have been made to manage *C. introflexus* occurrences within our study areas.

134

135 **2.2 Data acquisition**

136 Field and remote sensing data were acquired between 2013 and 2015 (Table 1). In each of the study areas, a
137 stratified sampling approach was used to lay out a set of 3 m x 3 m calibration (presence) plots, while a
138 random sampling approach was used for laying out validation plots (Fig. 1, Table 1). **In all four study areas,**
139 **we collected presence data to calibrate the model and presence/absence data to validate the prediction.**
140 **While a relatively low number of calibration plots was found to be sufficient (see Skowronek et al. 2017), we**
141 **used as many plots as we could gather within a reasonable timeframe for validation.** For all plots, the cover
142 of *C. introflexus* was recorded by dividing the plot in four equal parts and visually estimating and summing up
143 the cover of *C. introflexus* on each of the subplots. For Liereman and Kalmthout, a differential GPS (Trimble
144 GeoExplorer 6000) was used to determine the plot position and a differential correction was applied after
145 data collection, while for Kalmthout and Averbode, no differential correction could be performed, as the
146 device (Ashtech mobile mapper 10) did not allow for this feature. All positions are averages of at least 100
147 measurements.

148

149 **Airborne imaging spectroscopy data**, acquired by the Airborne Prism EXperiment (APEX) spectrometer were
150 used within this study. APEX is an airborne imaging spectrometer which collects information between
151 380nm and 2500nm with a **Full Width at Half Maximum (FWHM) ranging from 3 nm to 12 nm (after spectral**
152 **binning) in the visible and near-infrared spectral region, and from 9 nm to 12 nm in the SWIR region. Apex**
153 **data were acquired** by the Flemish Institute of Technology (VITO, Mol, Belgium) with **different** spatial
154 resolution ranging between 1.8 m x 1.8 m and 4 m x 4 m, depending on the study site (Table 1). **The spatial**
155 **resolution was highest for Sylt and lowest for Kalmthout. While for Sylt and Averbode, the calibration data**
156 **was collected less than one month before or after the flights campaigns took place, the calibration data for**
157 **Liereman was collected about one year after the flight, and the data for Kalmthout only two years after the**
158 **flight.**

159 The data were geometrically and atmospherically corrected using the standard processing applied to APEX
 160 (Sterckx et al., 2016; Vreys et al., 2016) at VITO's Central Data Processing Center. The processing chain is
 161 based on the MODTRAN4 software (Berk et al., 1999) in which the model atmosphere was set to "mid-
 162 latitude summer" and the employed aerosol type was "rural". The main atmospheric parameters (water
 163 vapor content and visibility) were derived from ground-based measurements using a Microtops
 164 sunphotometer and spectral ground control points, measured by means of an ASD spectrometer, were used
 165 as reference spectra. Where Microtops and/or ASD measurements were not available, all parameters were
 166 iteratively tuned to ensure a minimum spectral distortion in the water vapor absorption bands jointly with a
 167 high consistency between APEX spectra and reference spectra from available spectral libraries. After
 168 atmospheric correction, bands from both ends of the spectra and bands disturbed by water absorption were
 169 removed (bands between 1320-1447 nm and 1762-1988 nm: selected based on visual interpretation, i.e.
 170 noisy profile). Thus, a total of 244 spectral bands (between 426 nm and 2425 nm) were used in the
 171 subsequent analyses.

172 Table 1: Characteristics of the field data and the remote sensing data for each study site, p – presence plots,
 173 a – absence plot

174

Data	Sylt	Averbode	Liereman	Kalmthout
Flight dates	Jul-14	Sep-14	Sep-14	Jul-13
Fieldwork dates	Jul/Aug-14	Aug-14 & May-15	Sep-15	Aug-15
Number of calibration plots (presence plots)	57	27	49	50
Number of validation plots (presence and absence plots)	150 (48 p, 102 a)	93 (66 p, 27 a)	51 (28 p, 23 a)	50 (35 p, 15 abs)
GPS device	Trimble/ Mobile mapper	Ashtech Mobile mapper	Trimble, post- processed	Trimble, post- processed
Flight time (GMT)	11:21-12:13	12:48-13:12	12:21-12:39	11:05-11:18
Pixel size APEX data	1.8 m x 1.8 m	2.8 m x 2.8 m	2.8 m x 2.8 m	4 m x 4 m
Plot size	3 m x 3 m			

175

176 2.3 Data analysis

177 All species distribution models were built with Maxent (Phillips et al., 2004), a one-class classifier, which
178 differentiates the target species from a background sample based on the comparison of probability densities.
179 Maxent makes an estimate of the ratio between the conditional density of the predictors at the presence
180 sites and the unconditional density of the predictors across the study area, where the distance between
181 those densities is minimized. The logistic output of the model represents an estimate of the probability that
182 the species is present in a certain location. For detailed information on the model, see Phillips et al. (2006)
183 and Elith et al. (2011). **General advantages of Maxent are that it is relatively easy to use and freely available,**
184 **either through R or through its standalone software. Moreover, as a one-class classifier, it only requires**
185 **presence data to be collected in the field for model calibration which greatly reduces the amount of field**
186 **work necessary. In Skowronek et al. (2017a) we compared the performance of Maxent, Support Vector**
187 **Machine and Boosted Regression trees and found that all three classifiers allowed for the detection of the**
188 **two target species with similar success rates.**

189 Thus, for calibrating **the Maxent models** we used presence-only data (calibration dataset **collected in the**
190 **field**) and a random background sample with the 244 spectral bands serving as predictor variables. The
191 background sample for each study site consisted of **a large number of** random points located within the
192 biotope types of each study area, where the target species was potentially present (mostly dune areas and
193 natural grasslands). To delineate this area, we used existing biotope maps (INBO, 2016; LEGUAN, 2012;
194 Natuurpunt, 2012). To evaluate model performance, we used the independent validation dataset, containing
195 both presence and absence plots. The number of calibration and validation plots for each study site is given
196 in Table 1. The value of each calibration and validation plot is a weighted mean of the pixel values located
197 within the boundaries of each 3 m x 3 m field plot.

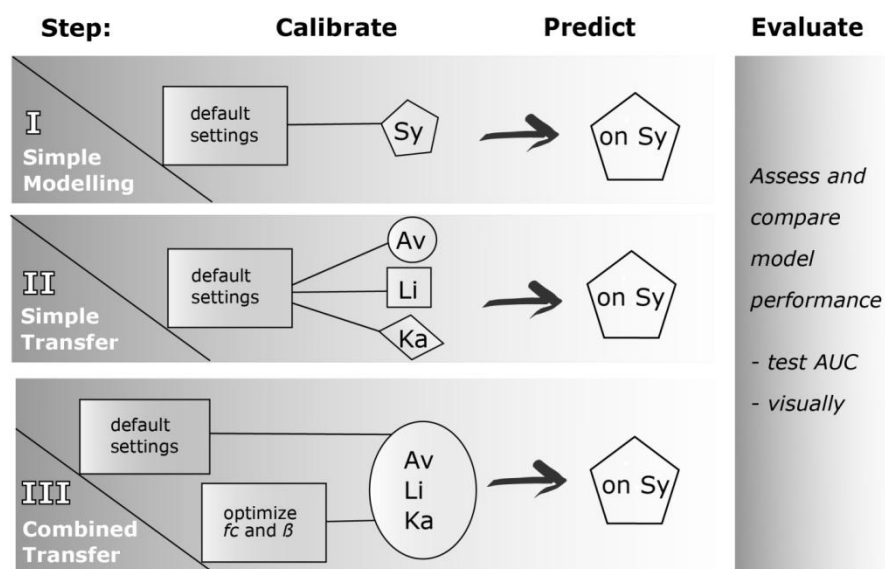
198 Within Maxent, there are two important modelling parameters. The first parameter is the regularization
199 multiplier (β), which may reduce over-fitting as it ensures that the empirical constraints are not being fit too
200 rigorously and by penalizing the model proportionally to the coefficients magnitude (Merow *et al.* 2013). The
201 other parameter is the feature class (fc), of which Maxent currently has six: linear (l), product (p), quadratic
202 (q), hinge (h), threshold (t) and categorical. For more information on the feature classes please see Phillips *et al.*
203 *et al.* (2006) and Elith *et al.* (2011). When using the default settings, β is 1 and the number of allowed feature
204 classes (fc) depends on the number of calibration plots.

205 Prior to starting the analysis, we tested the calibration and validation data for spatial sorting bias. It is
206 defined as the “difference between the geographic distance from testing-presence to training-presence sites
207 and the geographic distance from testing-absence (or testing-background) to training-presence sites”
208 (Hijmans, 2012). This spatial sorting bias can have a large impact on model performance (Hijmans, 2012;
209 Syfert *et al.*, 2013). Consequently, we followed Hijmans and Elith (2015) by calculating an indicator for spatial
210 sorting bias. If the indicator is 1, it means there is no bias, whereas an indicator of 0 means that a strong bias
211 exists.

212 Next, three different types of models were constructed, as outlined in Figure 2. In Step I, we calibrated and
213 tested a separate model for each study site using the calibration and validation datasets for that particular
214 site (simple modelling, Fig. 2). In this step, we also compared the relative importance of the different bands
215 (predictor variables) in the resulting model and assessed the amount of spectral variance in the different
216 study sites by calculating the standard deviation **for each band and** across the whole spectrum for the

217 background and calibration datasets. Subsequently, in Step II, for each study site we predicted the
 218 distribution of our target species using the models of the three other study areas generated in Step I,
 219 respectively. We evaluated the predictions by comparing them with the independent validation data sets
 220 (simple transfer, Fig. 2). This resulted in a total of 12 different validations, three for each study site, as each
 221 Step I model was applied on the three other areas. Finally, in Step III, we combined the calibration data and
 222 the background points of three different study sites and used these to build a single global model, which was
 223 then projected on the remaining fourth study site (combined transfer, Fig. 2), for each combination of sites.

224



225

226 Figure 2: Workflow for each study area using one study site (Sylt) as an example

227 We made use of the default settings for Maxent ($\beta=1$, fc =default, 10,000 background points) for Step I and II
 228 (simple modelling and simple transfer), whereas in Step III (combined transfer) we also tested the effect of
 229 varying the model parameters fc and β , as our results using the default settings indicated a highly complex
 230 and possibly overfit model (see section 3.3). We tested β values between 0.5 and 4 at 0.5 intervals, as values
 231 above the default have been found to produce better results (Radosavljevic and Anderson, 2014; Warren et
 232 al., 2014) as well as different combination of the feature classes linear (l), quadratic (q), hinge (h), product
 233 (p), and threshold (t), the model being restricted to the following feature classes: lq; lqp; h; qh; qhp; qhpt.
 234 For each of these models (192 in total) we used a selection of 9,000 background points obtained by
 235 combining a randomly selected subset of 3,000 background points per site. The lower number of background
 236 points was used due to limits in computing time. The Akaike information criterion (AIC) was used to select
 237 the best model (Warren and Seifert, 2010).

238 To evaluate model performance, we calculated the area under the curve for the independent validation data
 239 (test AUC) for all models. Additionally, for Step I, we also derived the presence / absence map from the
 240 probability maps using kappa as threshold. We then also derived the confusion matrices using the
 241 independent validation dataset and compared the overall accuracies (OA), sensitivity and specificity. For
 242 Steps II and III, we calculated a transferability index Tr_{AUC} (Heikkinen et al., 2012) from the obtained test

243 AUCs, which is a simple ratio between the test AUC for the transferred model (from Step II or III) and the test
244 AUC for the non-transferred simple model (from Step I) for each study site.

$$245 \quad Tr_{AUC} = \frac{testAUC_{Site1 \rightarrow Site2}}{testAUC_{Site2 \rightarrow Site2}} \quad (1)$$

246 When Tr_{AUC} is >1 , the transferred model performs better than the original model for that site, when it is <1 ,
247 the transferred model shows lower performance. Moreover, we compared the resulting probability maps
248 visually in order to evaluate the model performance.

249 All analysis were carried out using R Statistical Software 3.3.1 (R Development Core Team, 2016), QGIS 2.16
250 (QGIS Development Team, 2016) and pktools (Kempeneers, 2016). We mainly used the r-packages dismo
251 (Hijmans et al., 2016), raster (Hijmans, 2016) and rgdal (Bivand et al., 2016).

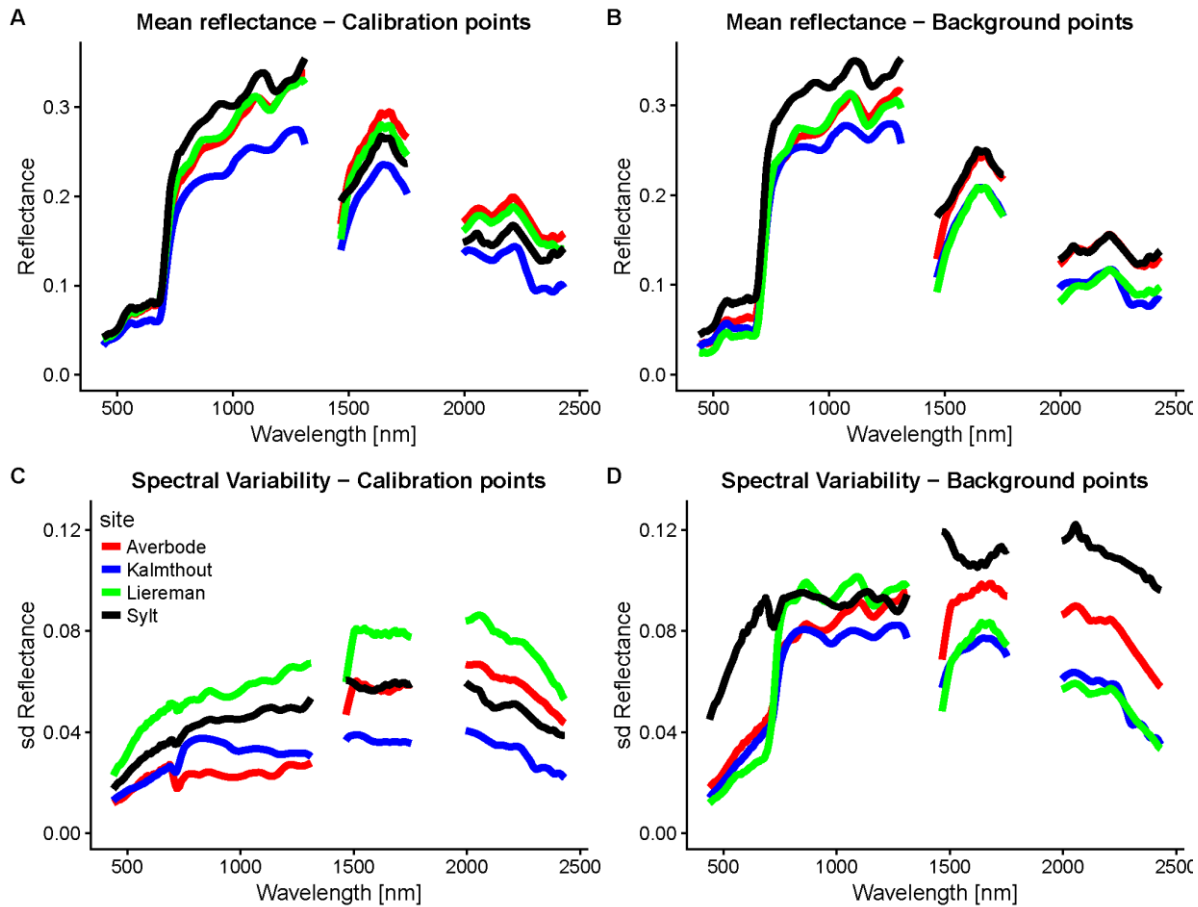
252 **3 Results**

253 **3.1 Spectral reflectance of the calibration and background plots across sites**

254
255 Figure 3 shows the mean reflectance and the spectral variability for all four study sites for the calibration
256 plots as well as the background plots. It is important to point out that the calibration spectra are averages of
257 all calibration plots, which may contain very high or very low amounts of the target species. For the
258 background data, it is important to note that this data may eventually also contain a few single data point
259 where the target species is present, as this data is randomly selected.

260 Overall, Sylt had higher mean reflectance values in the VIS/NIR, both for the calibration and background
261 points. Averbode showed the highest mean reflectance in the SWIR while Kalmthout on the other hand had
262 consistently lower reflectance values. The spectral variability within the calibration datasets was overall
263 highest for Liereman. For Sylt and Averbode, we observed relatively high variability in the VIS/NIR and in the
264 SWIR, respectively (Fig. 3). For the background points, Sylt showed the highest variability, followed by
265 Liereman in the NIR and by Averbode in the SWIR. The Kalmthout calibration and background spectra
266 contained only little spectral variation.

267

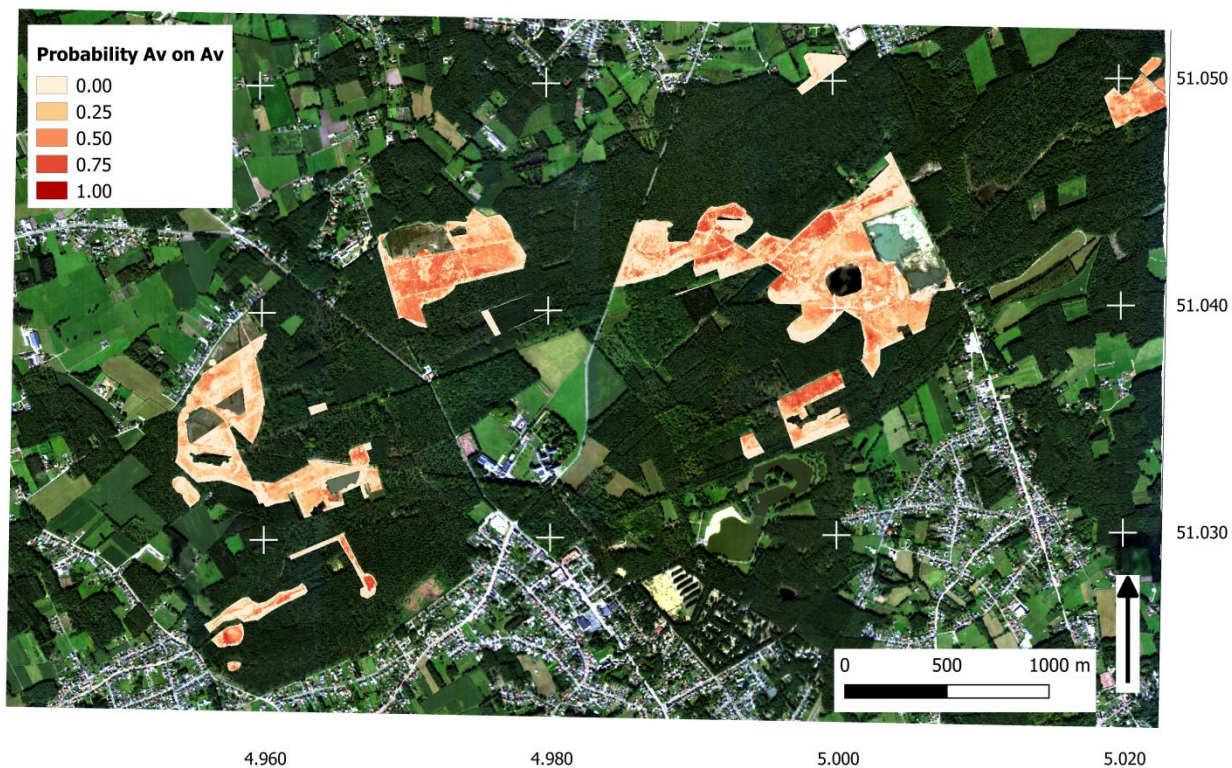


268

269 Figure 3: Mean reflectance and spectral variability of the calibration and the background data (measured as
270 the standard deviation per band for all plots) in a spectral range of 380 – 2500 nm.

271 3.2 Simple modelling and band importance

272 Site-specific models for mapping *C. introflexus* (Step I) resulted in OA values between 0.59 and 0.82 and test
273 AUC values between 0.57 and 0.85 (Table 2), with the range of AUC values 0.6-0.7, 0.7-0.8, 0.8-0.9 and 0.9-1
274 meaning poor, fair, good, and excellent model accuracy, respectively. Note that AUC values below 0.5 means
275 predictions are opposite to expectations. OA values were highest for the larger study sites, Sylt and
276 Kalmthout, and lower for Liereman and Averbode (Table 2). The value indicating spatial sorting bias was 1 for
277 Sylt and Liereman, meaning that there was no spatial sorting bias, and 0.88 for Kalmthout and 0.87 for
278 Averbode, indicating a relatively small bias. Calibration AUC values were between 0.87 and 0.93. An example
279 of this simple modelling for Averbode can be found in Fig. 4.

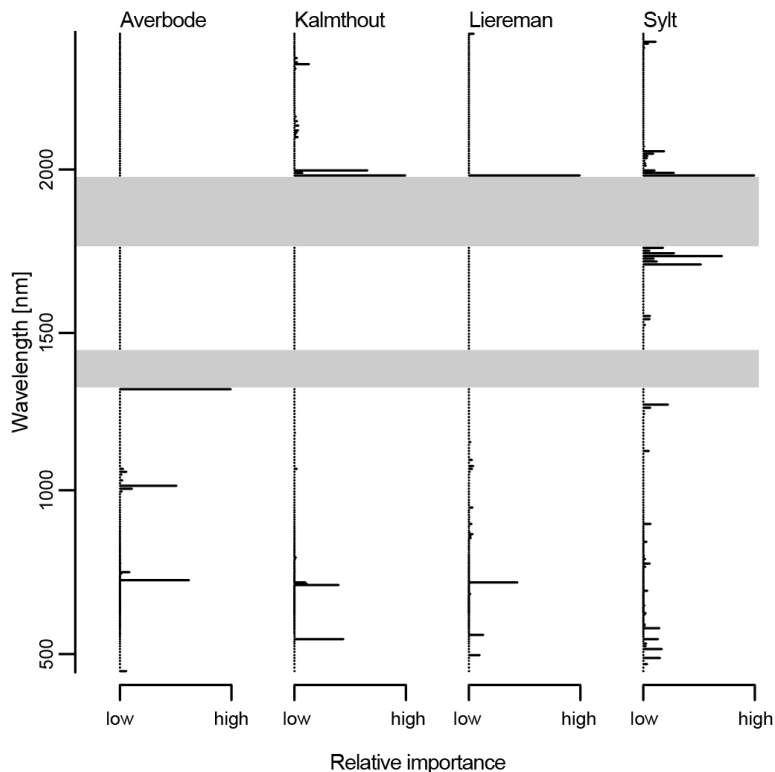


280

281 Figure 4: Predictions of the simple modelling (Step I) for Averbode showing the occurrence probability of
282 *Campylopus introflexus*; see Supplement 1 for all predictions for Steps I, II and III.

283 For three study sites (Liereman, Sylt, Kalmthout), the most important spectral band for modelling the
284 distribution of *C. introflexus* was located in the short wave infrared (SWIR, between 1500 and 2500 nm) at
285 1988 nm (Fig. 5). Plots with high covers of *C. introflexus* have higher reflectance values in the SWIR,
286 indicating a lower water content of those plots compared to the surrounding vegetation (see Skowronek et

287 al. 2017b for details). For Averbode, most important bands were located in the near infrared (NIR, between
 288 700 and 1400 nm). High variable importance was also indicated for a few bands in the visible (VIS, between
 289 400 and 700 nm), but the importance of this region was generally lower.



290

291 **Figure 5:** Relative band importance for the simple modelling (Step I) for each study site. **Gray shaded areas**
 292 **indicate bands that were removed from the data set prior to the analyses.**

293

294 **3.3 Simple transfer**

295 When evaluating the model calibrated on one study site and applied on the validation dataset of a different
 296 site (Step II, simple transfer), test AUC ranged between 0.45 and 0.85 (Table 2). For a total of six transfers,
 297 the resulting transferability index Tr_{AUC} was larger than one, indicating that the transfer model was more
 298 successful than the original model, while it was below one for a total of five transfers, indicating a less
 299 successful transfer (Fig. 6). The models calibrated for Sylt and Liereman showed slightly higher test AUC
 300 values when transferred to most other study sites, while results for the Kalmthout model were mixed.
 301 Transferring the Averbode model to the other study areas always resulted in lower test AUCs.

302 The visual evaluation confirmed that models with a Tr_{AUC} around or above one displayed similar patterns and
 303 that maximum probabilities were within the same range as the respective original model for each area. The
 304 predictions of models calibrated for Sylt and Liereman were generally very similar compared to the
 305 predictions of the original model (Fig. 7 and Supplement 1). On the other hand, models with a lower Tr_{AUC}

306 tended to have different patterns and lower maximum probabilities. Predictions resulting from the Averbode
 307 model showed the least similar pattern when transferred. In general, all transferred models showed
 308 smoother transitions than the original predictions. The full predictions for all study areas are provided in
 309 Supplement 1.

310 Table 2: Test AUC values for all 3 steps. For Step III – optimized, the test AUC values correspond to the test
 311 AUC values of those models with the lowest AIC value. For Step I, OA as well as sensitivity and specificity are
 312 displayed.

Step	Model	Applied on			
		Averbode	Kalmthout	Liereman	Sylt
Step I (single-site model)	AUC	0.61	0.85	0.57	0.78
	OA	0.63	0.82	0.59	0.76
	Sensitivity	0.62	0.86	0.50	0.69
	Specificity	0.67	0.73	0.70	0.79
Step II (simple transfer)	AUC Averbode	-	0.79	0.45	0.77
	AUC Kalmthout	0.58	-	0.57	0.82
	AUC Liereman	0.67	0.89	-	0.80
	AUC Sylt	0.72	0.84	0.62	-
Step III (transfer of multi-site models)	AUC Default	0.65	0.78	0.56	0.71
	AUC Optimized	0.70	0.90	0.54	0.83

313

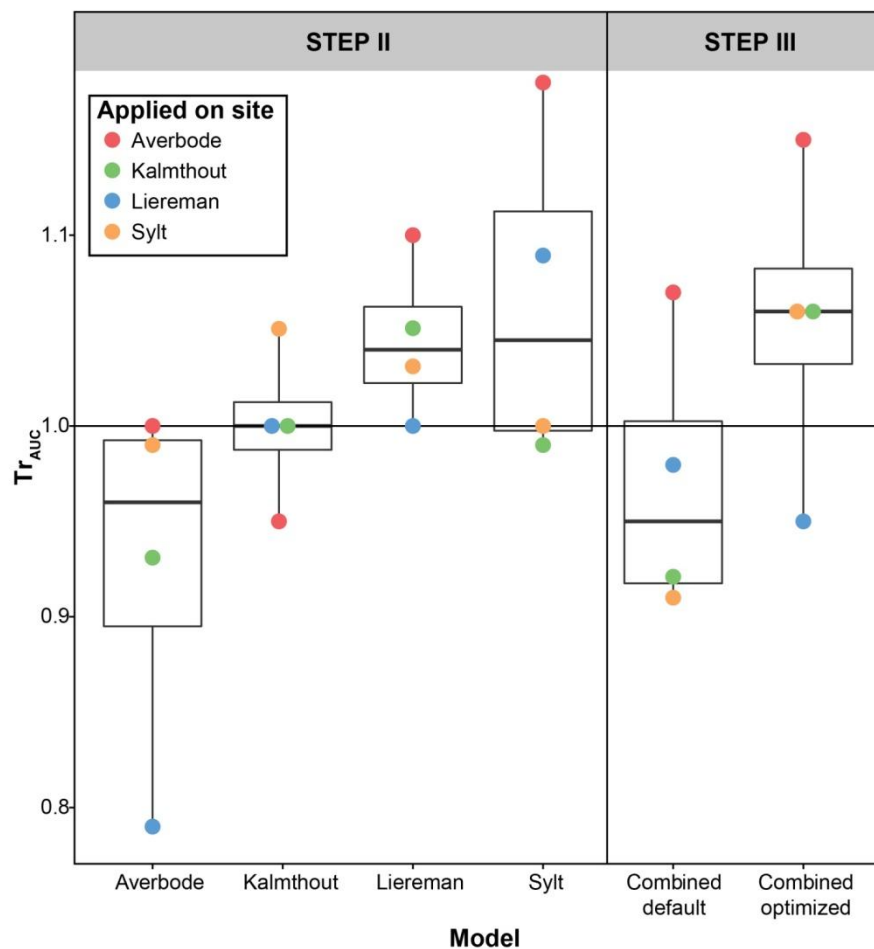
314 **3.4 Combined transfer**

315 For models based on calibration data from three different study sites (Step III, combined transfer), using the
 316 default settings resulted in very high calibration AUC values (between 0.94 and 0.96), while test AUC values
 317 were between 0.56 and 0.78 (Table 2). As those models were calibrated with a higher total number of
 318 presence plots, more feature classes were allowed for by the Maxent default settings (compared to the
 319 models in Step I and II). Tr_{AUC} values ranged between 0.91 and 1.07 (Fig.6).

320 Varying β and fc , we observed the tendencies demonstrated in Figure 8. We found that β values above the
 321 default of 1 mostly resulted in higher test AUC values, while the trends in fc were less obvious. Based on the
 322 AIC, we selected the combined model with optimized parameter settings for each of the study sites, which

323 resulted in test AUC values ranging between 0.54 and 0.9. The resulting Tr_{AUC} ranged between 0.95 and 1.15,
324 and thus had a higher median than any of the simple transfer or than the combined transfer using the
325 default settings, as shown in [Figure 7](#).

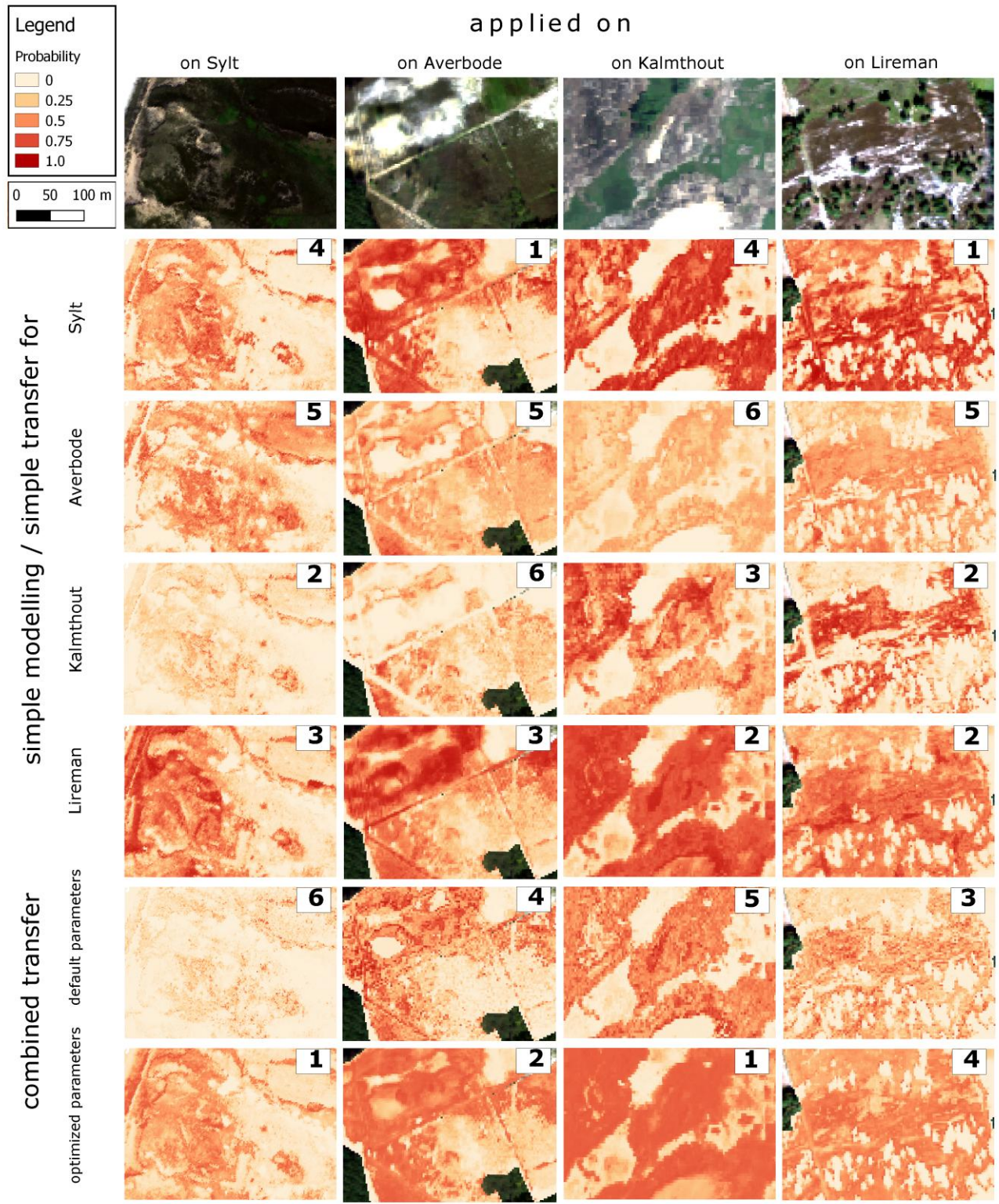
326 A visual evaluation of the resulting probability maps (see complete predictions in Supplement 1 and subsets
327 in [Figure 7](#)) showed that the combined models with an optimized parameter setting tended to show
328 smoother, more gradual transitions than the combined model with the default parameter settings.
329 Generally, they also predicted larger areas with presence of *C. introflexus*. Especially for Sylt, the optimized
330 prediction very much resembled the original prediction generated in Step I.



331

332 [Figure 6](#): The transferability index (Tr_{AUC}) for the four different models for each study site applied on the
333 respective three other study sites (STEP II) as well as the combined model (STEP III) using the default settings
334 (combined default), and the combined model using the optimized parameter settings (combined optimized)
335 applied on the respective study site that was not included in calibrating the model. The transferability index
336 is a ratio between the test AUC values of the transferred model and the local model (see section 2.3 for
337 details). A value of $Tr_{AUC} > 1$ indicates a better model than the model from Step I (simple modelling).

338



339

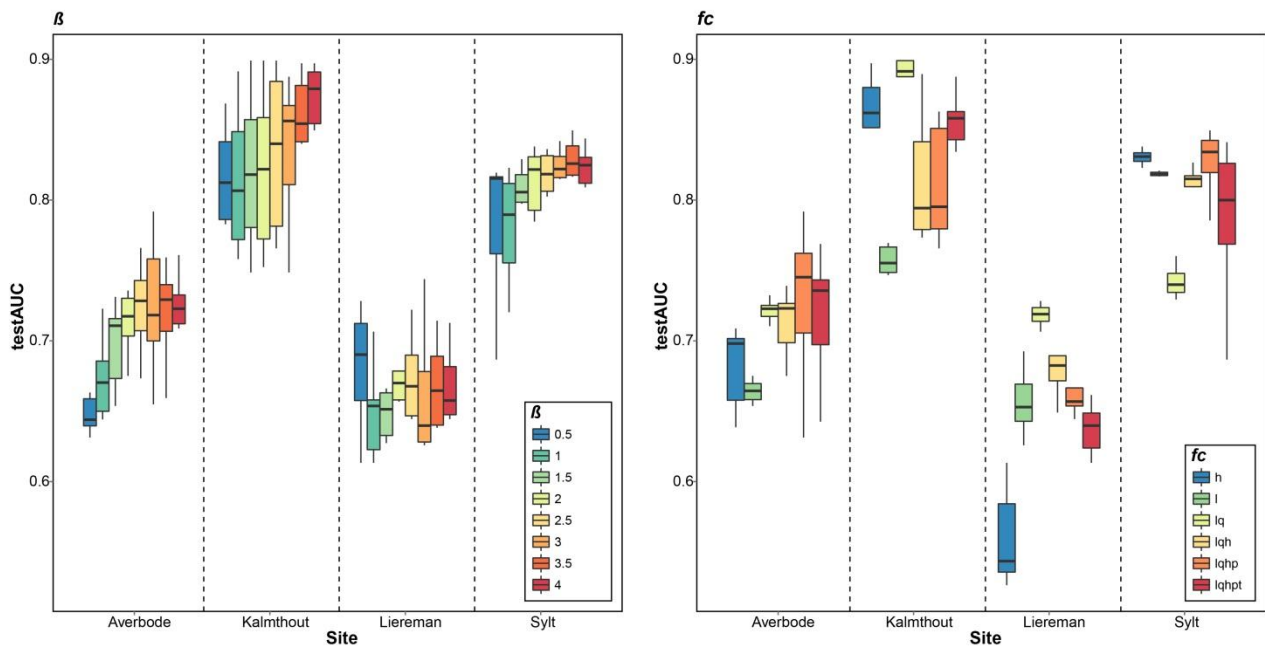
340 **Figure 7: Model results for the four different study areas using three different approaches (Step I, II and III).**

341 The number in the right corner of each subset indicates the rank according to test AUC values, 1 being the
 342 highest and 6 the lowest rank. The test AUC indicates how well the model performs, while the probability

343 shown in the maps indicates how much *C. introflexus* is present within each subset according to the different
344 model predictions.

345 **3.5 Comparison of simple modelling, simple transfer and combined transfer**

346 The simple transfer of models calibrated in one study site and validated on one another (Step II, simple
347 transfer) showed that for all study areas at least one of the transferred models outperformed the original
348 models (Step I, simple modelling). For the simple transfer, visual interpretations confirmed that especially
349 the Sylt and Liereman models showed good performances when transferred to other sites; the Averbode
350 model showed very low performances at all other study sites, and Kalmthout models performed better than
351 the local model for one area. The combined transfer models with optimized parameter settings (Step III,
352 combined transfer) outperformed the large majority of the simple transfer models as well as the combined
353 transfer using the default settings for three study areas (Fig. 6). Moreover, three optimized combined models
354 all had a transferability index >1, indicating that they performed similar or better than the original model
355 calibrated in the same area.



356
357 **Figure 8:** Effect of changing the Maxent parameter β and feature class (fc) on the observed test AUC values
358 for the combined models (Step III) created to map the invasive bryophyte *C. introflexus* in a dune habitat in
359 **four** different study sites **were** combined to generate the model and testing
360 model performance on the respective fourth site.

361 **4 Discussion**

362 **4.1 How successful is the model transfer and how do the characteristics of the input data affect model**
363 **performances?**

364 Projecting species distribution models to new areas – testing their transferability in space – is an important
365 topic in species distribution modelling (Heikkinen et al., 2012; Randin et al., 2006), which could be used for a
366 time- and cost-efficient large scale mapping of invasive alien plant species. Several factors can influence the
367 transferability of models using imaging spectroscopy data as predictor variables. First, model transferability is
368 likely to be affected by the amount of spectral variation in the model, which depends on the complexity and
369 heterogeneity of the vegetation in the respective target site. Several studies have discussed how the choice
370 of the background influences the prediction for Maxent (Merow et al., 2013). As a result, model
371 transferability greatly depends on the area selected for background sampling, the embedded heterogeneity
372 in the spectral signals of the co-occurring vegetation, as well as the phenological stage of the vegetation. The
373 latter plays a role as the reflectance signal of vegetation is largely determined by biochemical and biophysical
374 properties of the canopy. As these properties are subject to change with the phenological development of
375 the vegetation over the course of the year, spectral differences between the target species and the
376 background vegetation vary. This is especially true for some invasive alien plant species, where the
377 phenology differs substantially from that of the surrounding vegetation (Bradley, 2013). Transfer of a model
378 to a new site should thus consider the phenological stages of the vegetation at the time of data acquisition.
379 This could partly explain the generally lower test AUC values for Averbode and Liereman site, where the
380 remote sensing data was collected Mid-September, while the remote sensing data for Sylt and Kalmthout
381 was collected Mid-July.

382 Andrew and Ustin (2008) showed that the detectability of invasive alien species is highly dependent on the
383 specific environment of the study site. Hence, it is important to note that Maxent models and other
384 modelling techniques used in species distribution modelling are statistical or correlative-based models that
385 can only be transferred within the range of the calibration data (cf. interpolation). Predicting to areas outside
386 of the range of the calibration data (cf. extrapolation), on the other hand, will potentially lead to a number of
387 issues, which require a rigorous assessment (Elith and Leathwick, 2009; Jiménez-Valverde et al., 2011). In this
388 study, we found that at least one of the transferred models (simple transfer) outperformed the local model
389 (simple modelling). Furthermore, the more generalized model (combined transfer with optimized parameter
390 settings) outperformed most of the simple transfers. These findings may seem rather surprising at first
391 glance, as most previous studies on the potential model transferability indicate that models have a weaker
392 performance when they are applied to a new area (e.g. Barbosa et al., 2009; Heikkinen et al., 2012; Randin et
393 al., 2006).

394 However, as shown in [Figure 3](#) and the additional figure in Supplement 2, the range of conditions covered by
395 the models with good transferability for the simple transfer are larger than the range of conditions available
396 in the new area where those models are transferred and thus perform better than the original models of the
397 focal area. Hence, most of the successful transfers are typical cases of interpolations and thus consistent with
398 our findings. On the other hand, if the presence plots do not adequately represent the variability of the
399 spectral signal of the target species, a clear distinction might be difficult. This could partly explain the poor
400 performance of the Averbode model when applied to other sites (Step II) and its completely different use of
401 spectral bands: Averbode shows the most monotonic vegetation and a more sparse vegetation cover than
402 the other study sites. The transfer of the Averbode model is thus a good example of extrapolating beyond
403 the range of conditions for which it was calibrated. The relatively higher performance of the Sylt and
404 Liereman models, however, could be explained by the higher spectral variability which was embedded in the

405 calibration and background dataset used to calibrate the Maxent model (Fig. 3). It might also explain the
406 success of the combined model, which automatically covers a more comprehensive set of conditions than
407 any single model, thus increasing the probability of model interpolation at the expense of model
408 extrapolation.

409 Another point is that the spectral and spatial resolution as well as the quality of the remote sensing data
410 could influence the results (He et al., 2011). If these are low, the signal of the target species might be less
411 pronounced. While the spectral resolution was similar for all study sites, the spatial resolution varied. The
412 relatively lower performance of the Kalmthout model and the higher performance of the Sylt model could
413 also be explained by a relatively low/high spatial resolution: 4 m x 4 m and 1.8 m x 1.8 m respectively.

414 Those findings suggest that for the simple transfer, models based on remote sensing data with a higher
415 spatial resolution, which were calibrated in spectrally more heterogeneous areas and which correctly
416 identified the spectral band areas that are important for the species, are likely to perform well when
417 transferred to new areas. On the other hand, one has to be careful when transferring datasets that contain
418 less spectral heterogeneity, and have a lower spatial resolution, as these may not correctly identify the
419 reflectance signal that represents the target species. It also suggests that combining data from different
420 study sites may improve the overall model performance and limit the cases of model extrapolation and thus
421 should be considered if data from multiple sites are available.

422 **4.2 How do different model parameters affect models' performance?**

423 Our results show that the parameter settings for Maxent highly affect the model performance in a combined
424 modelling approach, and that models with an optimized parameter combination (based on minimizing the
425 AIC values) outperform models using the default settings (except for Liereman). Generally, using a β -value
426 higher than the default and varying fc produced models with a high transferability. We found that the choice
427 of fc was a very important factor in determining the model performance. The same effect was shown by
428 Moreno-Amat et al. (2015), while Syfert et al. (2013) concluded that the variation of fc only has a minor
429 effect on the model performance. Using only linear features did not produce the best results, as also shown
430 by Anderson and Gonzalez (2011), who compared models using only linear features with models using linear
431 and quadratic features. Based on the AIC values, a restriction to hinge features produced the best models,
432 while the highest test AUC values were found for different feature classes; mainly lq (linear and quadratic)
433 and $lqhp$ (linear, quadratic, hinge and product) depending on the model, as shown in Figure 6. Other studies
434 found that using less feature classes generally produces simpler models (Merow et al., 2013), which do not
435 necessarily perform less well. For example, Elith et al. (2011) found similar performance for using only hinge
436 features compared to using all possible feature classes.

437 Our finding that larger β -values mostly lead to a higher model performance agrees with existing literature.
438 Most authors recommend using β -values between 1 and 5 (Merow et al., 2013; Moreno-Amat et al., 2015;
439 Radosavljevic and Anderson, 2014). Warren et al. (2014), however, used a range of 0 to 15, and stated that a
440 wide range of different β -values was used for the optimal models selected using AIC. Shcheglovitova and
441 Anderson (2013) found that for small sample sizes, it is best to couple complex features (allow for more
442 feature classes) with higher regularization (higher β). The findings highlight the importance of understanding
443 the critical role of parameter tuning and model selection, which can drastically alter the resulting predictions.

444 **4.3 Recommendations**

445 Summarizing our findings in section 4.1 and 4.2 and the results of previous studies, we recommend
446 implementing the following strategies for transferring species distribution models which are based on
447 imaging spectroscopy data:

448 Concerning the input data:

449 i) The calibration data should adequately represent the spectral heterogeneity of the target species
450 and the surrounding vegetation. If available, data from different sites should be combined.

451 ii) Transfer should be made using data with the same or a slightly higher spatial resolution than the
452 target data set, and should be collected within the timeframe when the vegetation is in a similar
453 phenological stage.

454 iii) One should always **mitigate** sampling bias.

455 Concerning analysis:

456 iv) The effect of variable selection/reduction of the dimensionality of the input data should be tested.

457 v) Model parameters should be optimized (for Maxent by varying β and testing different fc).

458 Concerning output evaluation:

459 vi) The evaluation of the prediction with a **separate** independent validation dataset should always be
460 accompanied by a careful and sceptical visual examination by (local) experts.

461 **4.4 Uncertainties, future research needs and potential applications**

462 The impact of reducing the number of predictor variables was not investigated in this study, as Maxent has
463 shown to be less affected by collinearity issues than some other classifiers (Elith et al., 2011). However,
464 Warren et al. (2014) found that the variable selection had a larger effect than changing the regularization
465 parameter and recent remote sensing studies suggest that reducing the number of input variables by using
466 spectral indices (Tuanmu et al., 2011) or using reflectance-derived information on plant traits instead of
467 reflectance spectra is likely to improve model performance (Feilhauer et al., 2017). As those approaches
468 require complex additional processing steps, which are not in line with the scope of this study, which aims at
469 a simple, reproducible approach, we did not test the effect within this study. However, we acknowledge that
470 this question should be addressed in future research.

471 For all areas, there was a time lag of a few weeks to several months between the remote sensing acquisition
472 and the fieldwork campaign, which may cause a slight under or overestimation in the species cover. **This is
473 especially true for Liereman and Kalmthout. However, we did not observe a significant phenological
474 difference between the dates when the image data was acquired or between the different study sites. All
475 imagery was acquired around noon local time (see table 1). We thus estimate that the timing of the image
476 acquisition did not have any major impact on the results. A factor that did affect the results in a significant
477 way was the different spatial resolution (see section 4.1 for details).**

478

479 Additional uncertainties may occur due to the different GPS devices that were used. While we used devices
480 with differential correction for data collection in Liereman and Kalmthout, the devices used in Averbode and
481 Kalmthout did not have this option, which may led to larger position uncertainties on those study sites in
482 addition to the position uncertainties in the remote sensing data. Furthermore, our validation datasets,
483 particularly for Kalmthout and Liereman, were relatively small. We chose to still work with these datasets as
484 having to deal with a small amount of occurrences represents a real-world scenario for the (early) detection
485 of invasive alien plant species, where informed decisions have to be made with a limited amount of data.
486 However, collecting larger field datasets for validation might further enhance our understanding of the
487 model performances and transferability success (Bean et al., 2012). Another important factor influencing the
488 model are the soil reflectances, as some of the plots contain quite a high amount of bare soil. We did not
489 separately assess the influence of the soil reflectance due to a lack of adequate data, but doing so could
490 enhance the understanding of the different model performances. **Finally, while there was no sampling bias**
491 **(spatial sorting bias, see chapter 2.3) for Sylt and Liereman, there was a relatively small bias for Kalmthout**
492 **and Averbode.**

493 A simple transfer approach can be useful in the context of an early detection of invasive alien plant species.
494 In case remote sensing data with a similar resolution is available for an area, applying a model that was
495 formerly created for another dataset with similar vegetation composition might enable us to detect recently
496 invaded spots without having to manually search the whole area first in order to find enough spots to
497 calibrate a model for that area. For widely distributed species, such a model transfer might give us a good
498 first overview of the general distribution patterns and may guide following research or management
499 activities.

500 While we currently may not have very many situations where multiple imaging spectroscopy datasets are
501 available for the same study species in different regions to build a combined model, this might change in the
502 near future with the launch of hyperspectral satellite missions, such as **EnMAP (Environmental Mapping and**
503 **Analysis Program)**, where imaging spectroscopy data with a 30 m x 30 m resolution will be available
504 worldwide. While this spatial resolution is certainly too coarse for mapping *C. introflexus*, it might be
505 interesting for mapping larger species or vegetation types. **However, a similar transferability approach could**
506 **be applied to multispectral satellite data such as WorldView-2 or 3, which are readily available for larger**
507 **areas, and have proven to be useful for mapping certain invasive plant species (e.g. Fernandes et al., 2014;**
508 **Robinson et al., 2016).**

509 For a large scale mapping of *C. introflexus*, more research should be conducted on the usefulness **of such**
510 multispectral satellite data that might provide the necessary spatial and spectral resolution at lower costs
511 than the airborne hyperspectral data used in this study. For a cost-efficient mapping of *C. introflexus* at
512 smaller scales, the feasibility of mapping the species using multispectral data collected with unmanned aerial
513 vehicles (UAV) should be tested. Furthermore, a similar transferability approach could be applied for a large
514 remote sensing dataset where field data is only available within a few smaller subsets of the area. Our study
515 indicates a good model transferability using imaging spectroscopy data, but more research is necessary to
516 test model transferability for different species, different biotope types and different available spectral data
517 types.

518 **5 Conclusion**

519 In this study we successfully transferred species distribution models for *Campylopus introflexus* which were
520 calibrated at different sites using airborne imaging spectroscopy as explanatory variables. Our results
521 demonstrate that model transfer success was determined by a combination of i) the spectral heterogeneity
522 of the calibration dataset and how adequately it represents the spectral heterogeneity of the target dataset,
523 ii) the spatial resolution of the calibration dataset as well as the iii) parametrization and complexity of the
524 used model. As more remote sensing datasets become available, those techniques can improve model
525 results or be used to avoid additional time-consuming field work. This is especially relevant for a time- and
526 cost-efficient repetitive monitoring of invasive plant species, as it is impossible to frequently map invasive
527 species over large scales using traditional field mapping techniques. However, we do need this type of
528 information to be able to assess the spread of invasive species and manage them accordingly. This study
529 therefore explores challenges related to model transfer and gives practical recommendations regarding data
530 collection, data analysis and evaluation of the results.

531

532 6 References

- 533 Anderson, R.P., Gonzalez, I., 2011. Species-specific tuning increases robustness to sampling bias in
534 models of species distributions: An implementation with Maxent. *Ecol. Modell.* 222, 2796–
535 2811. <https://doi.org/10.1016/j.ecolmodel.2011.04.011>
- 536 Andrew, M., Ustin, S., 2008. The role of environmental context in mapping invasive plants with
537 hyperspectral image data. *Remote Sens. Environ.* 112, 4301–4317.
538 <https://doi.org/10.1016/j.rse.2008.07.016>
- 539 Barbosa, A.M., Real, R., Mario Vargas, J., 2009. Transferability of environmental favourability
540 models in geographic space: The case of the Iberian desman (*Galemys pyrenaicus*) in Portugal
541 and Spain. *Ecol. Modell.* 220, 747–754. <https://doi.org/10.1016/j.ecolmodel.2008.12.004>
- 542 Bean, W.T., Stafford, R., Brashares, J.S., 2012. The effects of small sample size and sample bias on
543 threshold selection and accuracy assessment of species distribution models. *Ecography (Cop.)*.
544 35, 250–258. <https://doi.org/10.1111/j.1600-0587.2011.06545.x>
- 545 Berk, A., Anderson, G.P., Bernstein, L.S., Acharya, P.K., Dothe, H., Matthew, M.W., Adler-Golden,
546 S.M., Chetwynd, Jr., J.H., Richtsmeier, S.C., Pukall, B., Allred, C.L., Jeong, L.S., Hoke, M.L.,
547 1999. MODTRAN4 radiative transfer modeling for atmospheric correction, in: Larar, A.M.
548 (Ed.), . p. 348. <https://doi.org/10.1117/12.366388>
- 549 Biermann, R., Daniels, F.J.A., 1997. Changes in a lichen-rich dry sand grassland vegetation with
550 special reference to lichen synusiae and *Campylopus introflexus*. *Phytocoenologia* 27, 257–273.
551 <https://doi.org/10.1127/phyto/27/1997/257>
- 552 Bivand, R., Keitt, T., Rowlingson, B., 2016. rgdal: Bindings for the Geospatial Data Abstraction
553 Library [WWW Document]. URL <https://cran.r-project.org/package=rgdal>
- 554 Boxel, J.H., Jungerius, P.D., Kieffer, N., Hampele, N., 1997. Ecological effects of reactivation of
555 artificially stabilized blowouts in coastal dunes. *J. Coast. Conserv.* 3, 57–62.
556 <https://doi.org/10.1007/BF02908179>
- 557 Bradley, B. a., 2013. Remote detection of invasive plants: a review of spectral, textural and
558 phenological approaches. *Biol. Invasions* 16, 1411–1425. <https://doi.org/10.1007/s10530-013-0578-9>
- 560 Cheng, Y.-B., 2007. Mapping an invasive species, kudzu (*Pueraria montana*), using hyperspectral
561 imagery in western Georgia. *J. Appl. Remote Sens.* 1, 13514. <https://doi.org/10.1117/1.2749266>
- 562 DAISIE, 2015. DAISIE - 100 of the Worst [WWW Document]. URL [http://www.europe-](http://www.europe-alien.org/speciesTheWorst.do)
563 [alien.org/speciesTheWorst.do](http://www.europe-alien.org/speciesTheWorst.do) (accessed 6.11.15).
- 564 Duque-Lazo, J., van Gils, H., Groen, T.A., Navarro-Cerrillo, R.M., 2016. Transferability of species
565 distribution models: The case of *Phytophthora cinnamomi* in Southwest Spain and Southwest
566 Australia. *Ecol. Modell.* 320, 62–70. <https://doi.org/10.1016/j.ecolmodel.2015.09.019>
- 567 Elith, J., Leathwick, J.R., 2009. Species Distribution Models: Ecological Explanation and Prediction
568 Across Space and Time. *Annu. Rev. Ecol. Evol. Syst.* 40, 677–697.
569 <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- 570 Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation
571 of MaxEnt for ecologists. *Divers. Distrib.* 17, 43–57. [https://doi.org/10.1111/j.1472-](https://doi.org/10.1111/j.1472-4642.2010.00725.x)
572 [4642.2010.00725.x](https://doi.org/10.1111/j.1472-4642.2010.00725.x)
- 573 Essl, F., Steinbauer, K., Dullinger, S., Mang, T., Moser, D., 2014. Little, but increasing evidence of
574 impacts by alien bryophytes. *Biol. Invasions* 16, 1175–1184. [https://doi.org/10.1007/s10530-](https://doi.org/10.1007/s10530-013-0572-2)
575 [013-0572-2](https://doi.org/10.1007/s10530-013-0572-2)
- 576 Feilhauer, H., Somers, B., van der Linden, S., 2017. Optical trait indicators for remote sensing of
577 plant species composition: Predictive power and seasonal variability. *Ecol. Indic.* 73, 825–833.

578 <https://doi.org/10.1016/j.ecolind.2016.11.003>

579 Fernandes, M.R., Aguiar, F.C., Silva, J.M.N., Ferreira, M.T., Pereira, J.M.C., 2014. Optimal
580 attributes for the object based detection of giant reed in riparian habitats: A comparative study
581 between airborne high spatial resolution and WorldView-2 imagery. *Int. J. Appl. Earth Obs. Geoinf.* 32, 79–91. <https://doi.org/10.1016/j.jag.2014.03.026>

582 He, K.S., Bradley, B.A., Cord, A.F., Rocchini, D., Tuanmu, M.-N., Schmidtlein, S., Turner, W.,
583 Wegmann, M., Pettorelli, N., 2015. Will remote sensing shape the next generation of species
584 distribution models? *Remote Sens. Ecol. Conserv.* 1, 4–18. <https://doi.org/10.1002/rse2.7>

585 He, K.S., Rocchini, D., Neteler, M., Nagendra, H., 2011. Benefits of hyperspectral remote sensing for
586 tracking plant invasions. *Divers. Distrib.* 17, 381–392. <https://doi.org/10.1111/j.1472-4642.2011.00761.x>

587 Heikkinen, R.K., Marmion, M., Luoto, M., 2012. Does the interpolation accuracy of species
588 distribution models come at the expense of transferability? *Ecography (Cop.)*. 35, 276–288.
589 <https://doi.org/10.1111/j.1600-0587.2011.06999.x>

590 Hijmans, R.J., 2016. raster: Geographic Data Analysis and Modeling [WWW Document]. URL
591 <https://cran.r-project.org/package=raster>

592 Hijmans, R.J., 2012. Cross-validation of species distribution models: removing spatial sorting bias
593 and calibration with a null model. *Ecology* 93, 679–688. <https://doi.org/10.1890/11-0826.1>

594 Hijmans, R.J., Elith, J., 2015. Species distribution modeling with R. R CRAN Proj. 79 pp.

595 Hijmans, R.J., Phillips, S., Leathwick, J., Elith, J., 2016. dismo: Species Distribution Modeling
596 [WWW Document]. URL <https://cran.r-project.org/package=dismo>

597 Huang, C., Asner, G.P., 2009. Applications of Remote Sensing to Alien Invasive Plant Studies.
598 *Sensors* 9, 4869–4889. <https://doi.org/10.3390/s90604869>

599 Instituut voor Natuurbehoud, 2016. BIOLOGISCHE WAARDERINGSKAART [WWW Document].
600 URL [http://www.geopunt.be/download?container=bwk2&title=Biologische waarderingskaart -
601 Natura 2000 Habitatkaart](http://www.geopunt.be/download?container=bwk2&title=Biologische%20waarderingskaart%20Natura%202000%20Habitatkaart) (accessed 2.2.18).

602 Jiménez-Valverde, A., Peterson, A.T., Soberón, J., Overton, J.M., Aragón, P., Lobo, J.M., 2011. Use
603 of niche models in invasive species risk assessments. *Biol. Invasions* 13, 2785–2797.
604 <https://doi.org/10.1007/s10530-011-9963-4>

605 Kempeneers, P., 2016. pktools 2.6.7 Processing Kernel for geospatial data [WWW Document]. URL
606 <http://pktools.nongnu.org> (accessed 2.2.18).

607 Ketner-Oostra, R., Sykora, K. V., 2000. Vegetation succession and lichen diversity on dry coastal
608 calcium-poor dunes and the impact of management experiments. *J. Coast. Conserv.* 6, 191–206.
609 <https://doi.org/10.1007/BF02913815>

610 Ketner-Oostra, R., Sýkora, K. V., 2004. Decline of lichen-diversity in calcium-poor coastal dune
611 vegetation since the 1970s, related to grass and moss encroachment. *Phytocoenologia* 34, 521–
612 549. <https://doi.org/10.1127/0340-269X/2004/0034-0521>

613 LEGUAN, 2012. Kartierung der Salzwiesen und Dünen an der Westküste von Schleswig-Holstein
614 2011-2012 – Biotopkartierung Sylt.

615 Mateo, R.G., Broennimann, O., Petitpierre, B., Muñoz, J., van Rooy, J., Laenen, B., Guisan, A.,
616 Vanderpoorten, A., 2015. What is the potential of spread in invasive bryophytes? *Ecography*
617 (Cop.). 38, 480–487. <https://doi.org/10.1111/ecog.01014>

618 Merow, C., Smith, M.J., Silander, J.A., 2013. A practical guide to MaxEnt for modeling species'
619 distributions: What it does, and why inputs and settings matter. *Ecography (Cop.)*. 36, 1058–
620 1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>

621 Michez, A., Piégay, H., Jonathan, L., Claessens, H., Lejeune, P., 2016. Mapping of riparian invasive
622 species with supervised classification of Unmanned Aerial System (UAS) imagery. *Int. J. Appl. Earth Obs. Geoinf.* 44, 88–94. <https://doi.org/10.1016/j.jag.2015.06.014>

626 Mirik, M., Ansley, R.J., Steddom, K., Jones, D.C., Rush, C.M., Michels, G.J., Elliott, N.C., 2013.
627 Remote distinction of a noxious weed (Musk Thistle: *Carduus Nutans*) using airborne
628 hyperspectral imagery and the support vector machine classifier. *Remote Sens.* 5, 612–630.
629 <https://doi.org/10.3390/rs5020612>

630 Moreno-Amat, E., Mateo, R.G., Nieto-Lugilde, D., Morueta-Holme, N., Svenning, J.C., García-
631 Amorena, I., 2015. Impact of model complexity on cross-temporal transferability in Maxent
632 species distribution models: An assessment using paleobotanical data. *Ecol. Modell.* 312, 308–
633 317. <https://doi.org/10.1016/j.ecolmodel.2015.05.035>

634 Müllerová, J., Brůna, J., Bartaloš, T., Dvořák, P., Vítková, M., Pyšek, P., 2017. Timing Is Important:
635 Unmanned Aircraft vs. Satellite Imagery in Plant Invasion Monitoring. *Front. Plant Sci.* 8, 1–13.
636 <https://doi.org/10.3389/fpls.2017.00887>

637 Muscarella, R., Galante, P.J., Soley-Guardia, M., Boria, R.A., Kass, J.M., Uriarte, M., Anderson,
638 R.P., 2014. ENMeval: An R package for conducting spatially independent evaluations and
639 estimating optimal model complexity for <sc>Maxent</sc> ecological niche models.
640 *Methods Ecol. Evol., Remote Sensing and Digital Image Processing* 5, 1198–1205.
641 <https://doi.org/10.1111/2041-210X.12261>

642 Natuurpunt, 2012. 1st monitoringsrapport Averbode Bos & Heide. Bijlage 4.1.2.: Actuele
643 natuurtypes.

644 Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species
645 geographic distributions. *Ecol. Modell.* 190, 231–259.
646 <https://doi.org/10.1016/j.ecolmodel.2005.03.026>

647 Phillips, S.J., Dudík, M., Schapire, R.E., 2004. A maximum entropy approach to species distribution
648 modeling. *Proc. twenty-first Int. Conf. Mach. Learn.* 655–662.
649 <https://doi.org/10.1145/1015330.1015412>

650 Proctor, C., Robinson, V., He, Y., 2012. Multispectral detection of European frog-bit in the South
651 Nation River using Quickbird imagery. *Can. J. Remote Sens.* 38, 476–486.
652 <https://doi.org/10.5589/m12-040>

653 QGIS Development Team, 2016. QGIS Geographic Information System. Open Source Geospatial
654 Foundation Project [WWW Document]. URL <http://www.qgis.org/>

655 R Development Core Team, 2016. R: A language and environment for statistical computing.

656 Radosavljevic, A., Anderson, R.P., 2014. Making better Maxent models of species distributions:
657 Complexity, overfitting and evaluation. *J. Biogeogr.* 41, 629–643.
658 <https://doi.org/10.1111/jbi.12227>

659 Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M., Guisan, A., 2006. Are
660 niche-based species distribution models transferable in space? *J. Biogeogr.* 33, 1689–1703.
661 <https://doi.org/10.1111/j.1365-2699.2006.01466.x>

662 Richards, P.W., 1963. *Campylopus introflexus* (Hedw.) Brid. and *C. polytrichoides* De Not. in the
663 British Isles; a preliminary account. *Trans. Brit. Bryol. Soc.* 4, 404–417.

664 Robinson, T.P., Wardell-Johnson, G.W., Pracilio, G., Brown, C., Corner, R., van Klinken, R.D.,
665 2016. Testing the discrimination and detection limits of WorldView-2 imagery on a challenging
666 invasive plant target. *Int. J. Appl. Earth Obs. Geoinf.* 44, 23–30.
667 <https://doi.org/10.1016/j.jag.2015.07.004>

668 Shcheglovitova, M., Anderson, R.P., 2013. Estimating optimal complexity for ecological niche
669 models: A jackknife approach for species with small sample sizes. *Ecol. Modell.* 269, 9–17.
670 <https://doi.org/10.1016/j.ecolmodel.2013.08.011>

671 Skowronek, S., Asner, G.P., Feilhauer, H., 2017a. Performance of one-class classifiers for invasive
672 species mapping using airborne imaging spectroscopy. *Ecol. Inform.* 37, 66–76.
673 <https://doi.org/10.1016/j.ecoinf.2016.11.005>

674 Skowronek, S., Ewald, M., Isermann, M., Van De Kerchove, R., Lenoir, J., Aerts, R., Warrie, J.,
675 Hattab, T., Honnay, O., Schmidtlein, S., Rocchini, D., Somers, B., Feilhauer, H., 2017b.
676 Mapping an invasive bryophyte species using hyperspectral remote sensing data. *Biol. Invasions*
677 19, 239–254. <https://doi.org/10.1007/s10530-016-1276-1>

678 Somers, B., Asner, G.P., 2013. Invasive species mapping in hawaiian rainforests using multi-
679 temporal hyperion spaceborne imaging spectroscopy. *IEEE J. Sel. Top. Appl. Earth Obs.*
680 *Remote Sens.* 6, 351–359. <https://doi.org/10.1109/JSTARS.2012.2203796>

681 Sterckx, S., Vreys, K., Biesemans, J., Iordache, M.-D., Bertels, L., Meuleman, K., 2016.
682 Atmospheric correction of APEX hyperspectral data. *Misc. Geogr.* 20.
683 <https://doi.org/10.1515/mgrsd-2015-0022>

684 Syfert, M.M., Smith, M.J., Coomes, D.A., 2013. The Effects of Sampling Bias and Model
685 Complexity on the Predictive Performance of MaxEnt Species Distribution Models. *PLoS One*
686 8, e55158. <https://doi.org/10.1371/journal.pone.0055158>

687 Tuanmu, M.-N., Viña, A., Roloff, G.J., Liu, W., Ouyang, Z., Zhang, H., Liu, J., 2011. Temporal
688 transferability of wildlife habitat models: implications for habitat monitoring. *J. Biogeogr.* 38,
689 1510–1523. <https://doi.org/10.1111/j.1365-2699.2011.02479.x>

690 Vreys, K., Iordache, M.-D., Biesemans, J., Meuleman, K., 2016. Geometric correction of APEX
691 hyperspectral data. *Misc. Geogr.* 20, 11–15. <https://doi.org/10.1515/mgrsd-2016-0006>

692 Warren, D.L., Seifert, S.N., 2010. Ecological niche modeling in Maxent: the importance of model
693 complexity and the performance of model selection criteria. *Ecol. Appl.* 21, 335–342.
694 <https://doi.org/10.1890/10-1171.1>

695 Warren, D.L., Wright, A.N., Seifert, S.N., Shaffer, H.B., 2014. Incorporating model complexity and
696 spatial sampling bias into ecological niche models of climate change risks faced by 90
697 California vertebrate species of concern. *Divers. Distrib.* 20, 334–343.
698 <https://doi.org/10.1111/ddi.12160>

699 Wenger, S.J., Olden, J.D., 2012. Assessing transferability of ecological models: An underappreciated
700 aspect of statistical validation. *Methods Ecol. Evol.* 3, 260–267. <https://doi.org/10.1111/j.2041-210X.2011.00170.x>

701
702

Figure 1
[Click here to download high resolution image](#)

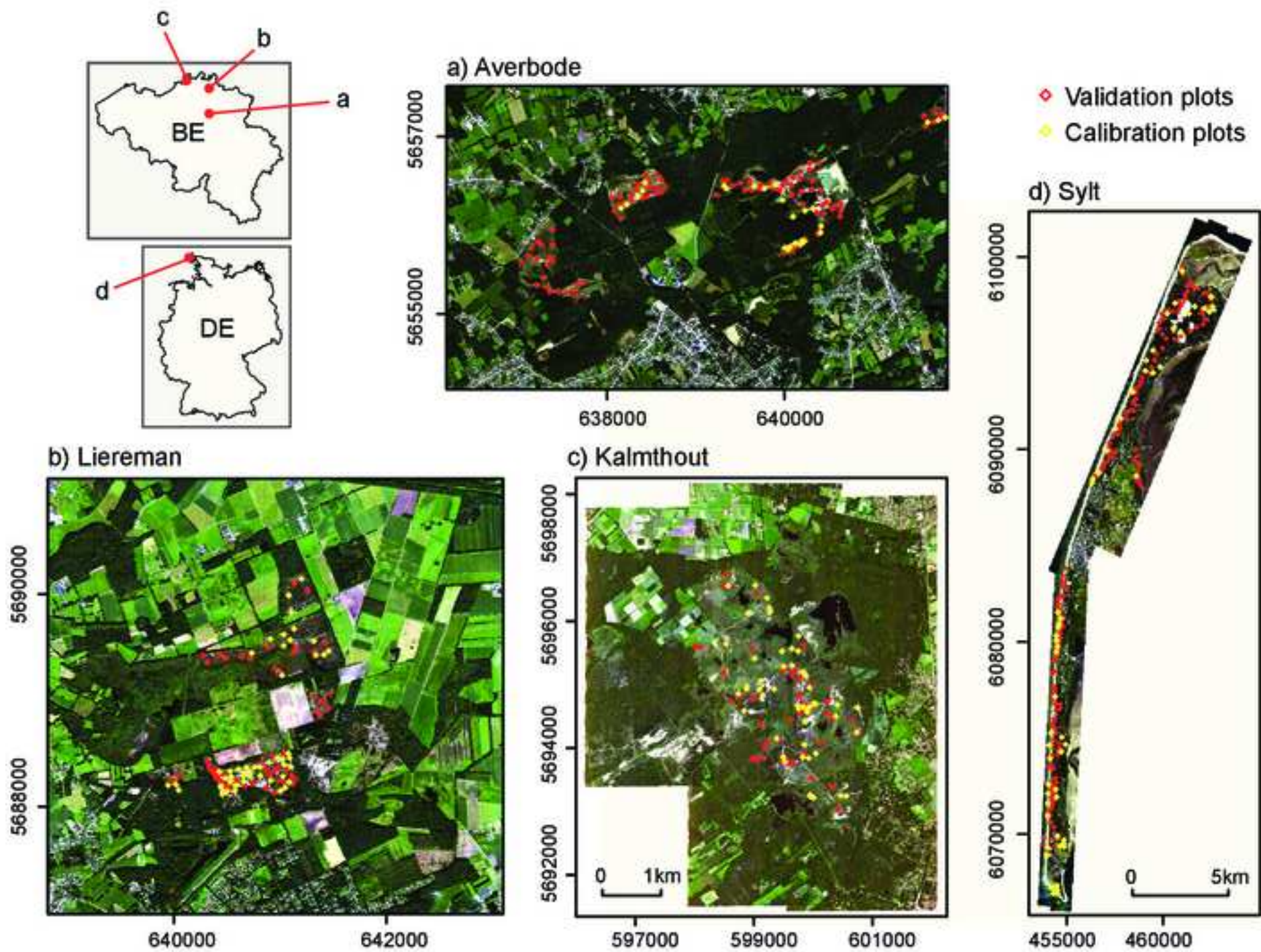


Figure 2

Step:

Calibrate

Predict

Evaluate

I
Simple
Modelling

default
settings

Sy



on Sy

II
Simple
Transfer

default
settings

Av

Li

Ka



on Sy

III
Combined
Transfer

default
settings

optimize
 fc and β

Av
Li
Ka



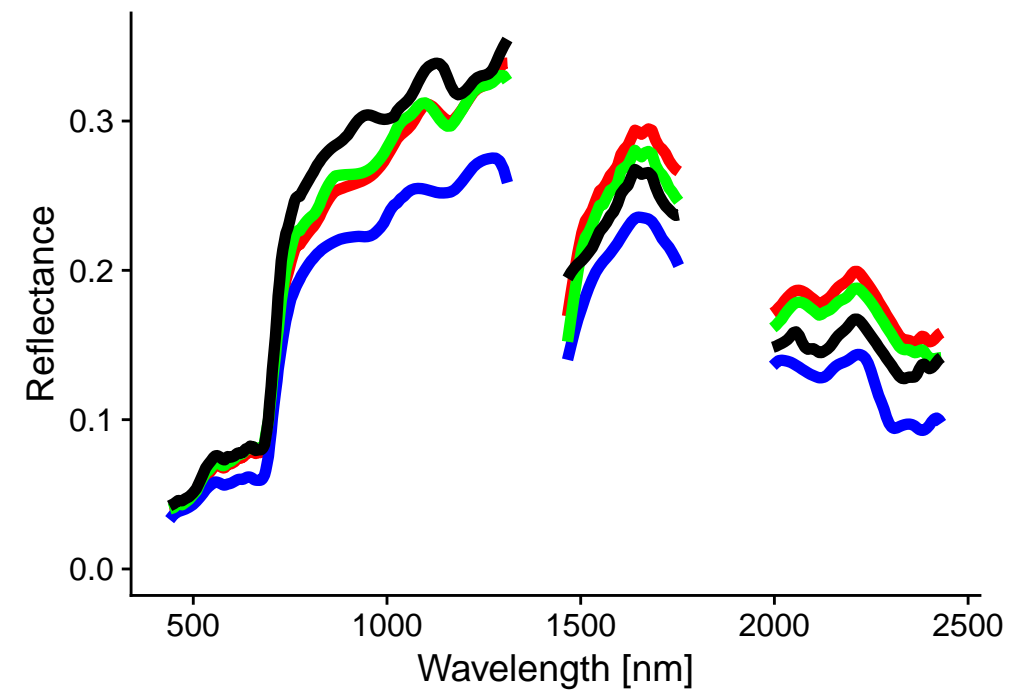
on Sy

*Assess and
compare
model
performance*

- test AUC
- visually

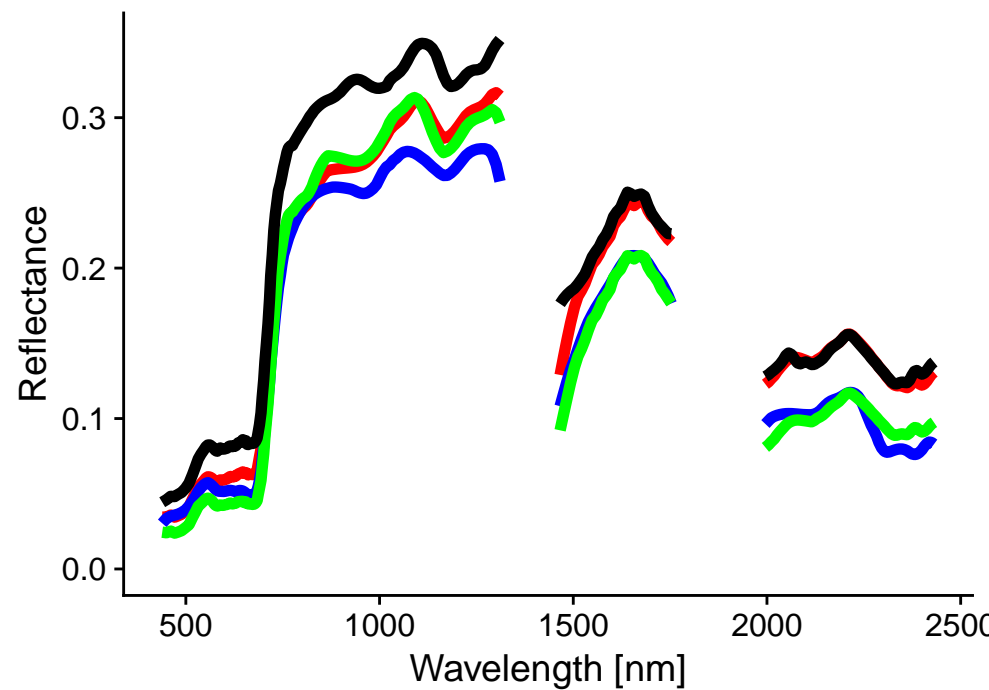
Figure 3

Mean reflectance – Calibration points



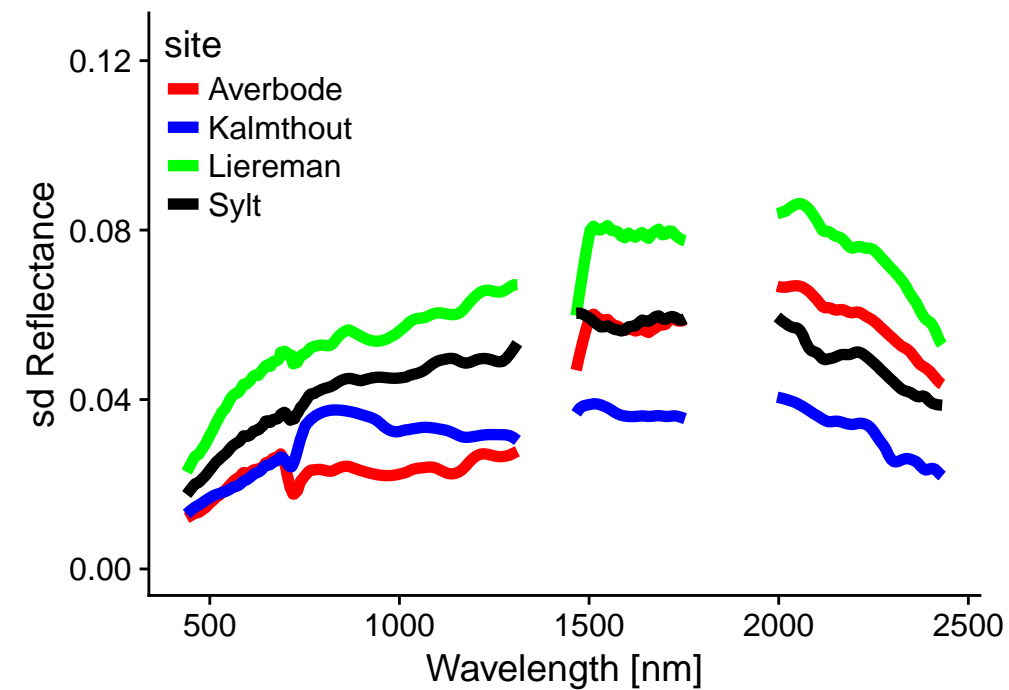
B

Mean reflectance – Background points



C

Spectral Variability – Calibration points



D

Spectral Variability – Background points

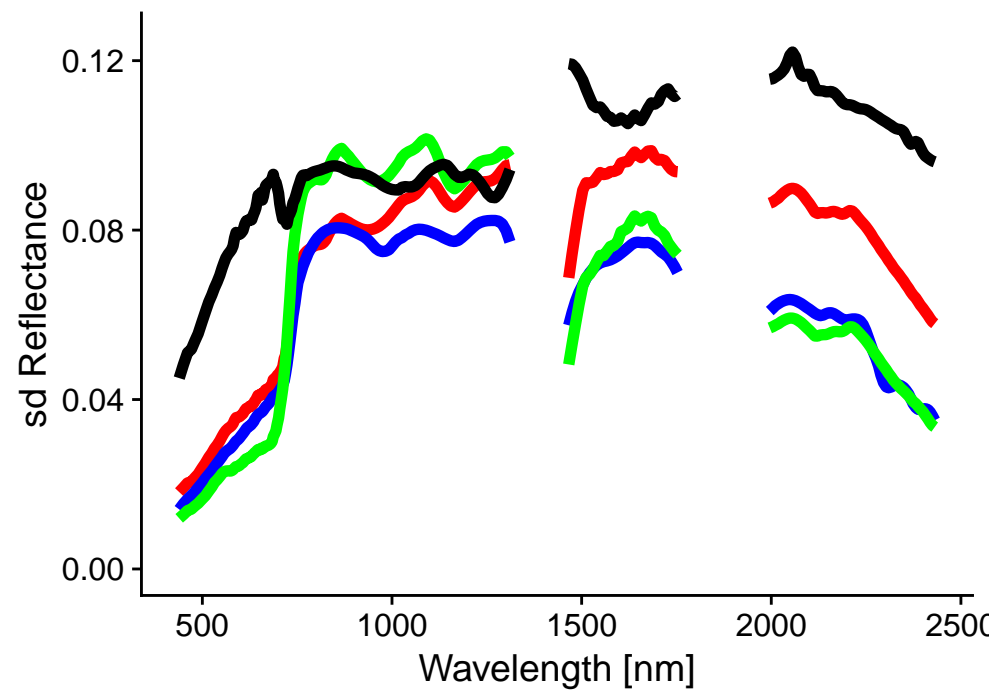


Figure 4
[Click here to download high resolution image](#)

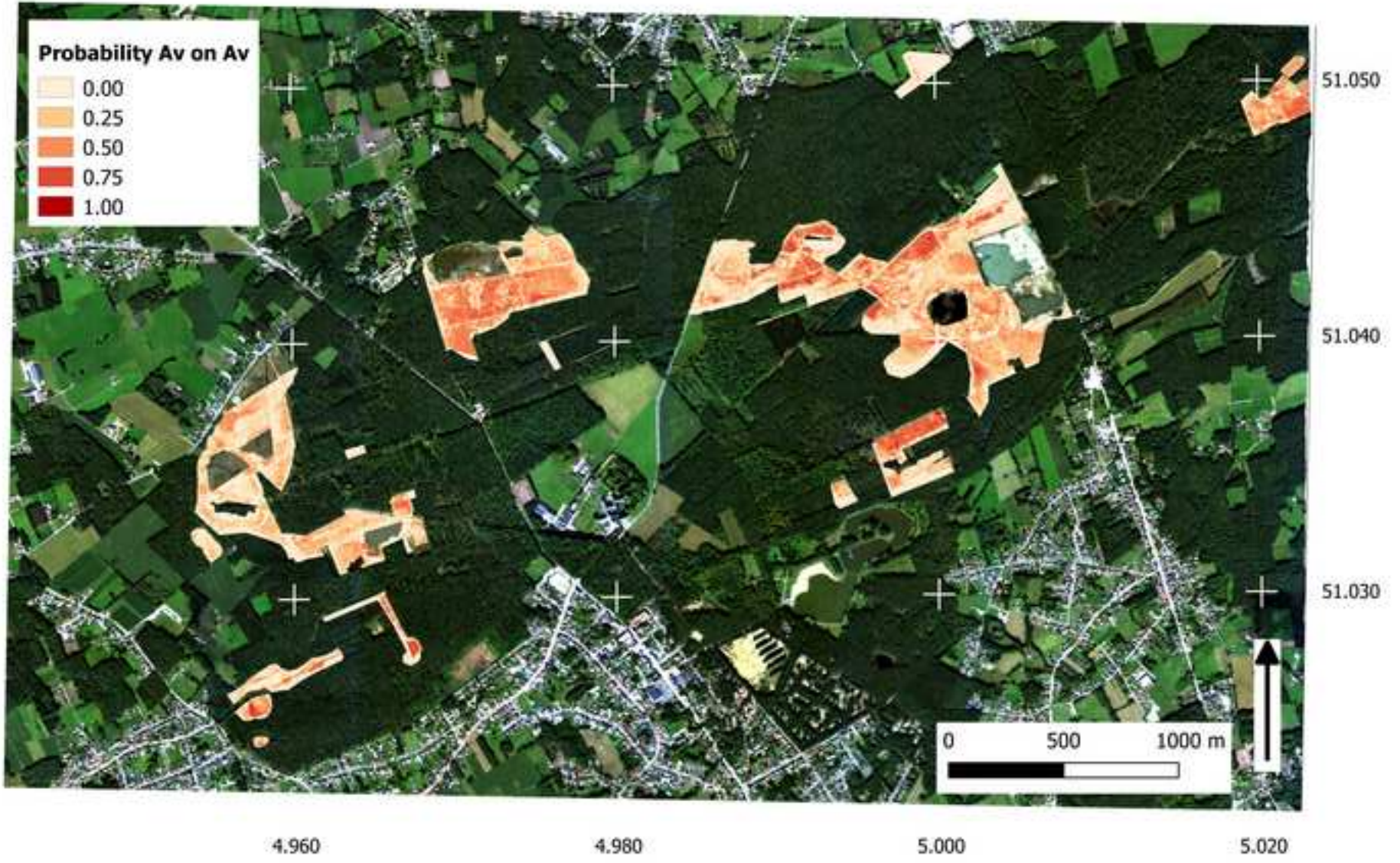


Figure 5

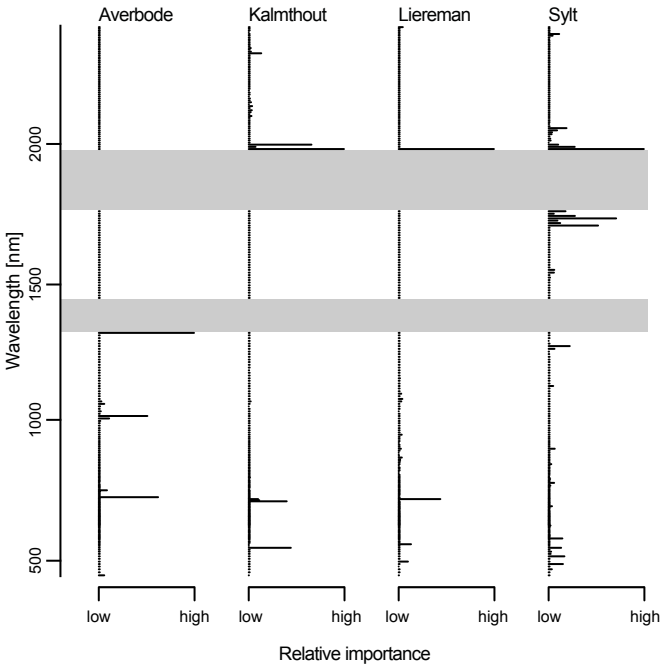


Figure 6

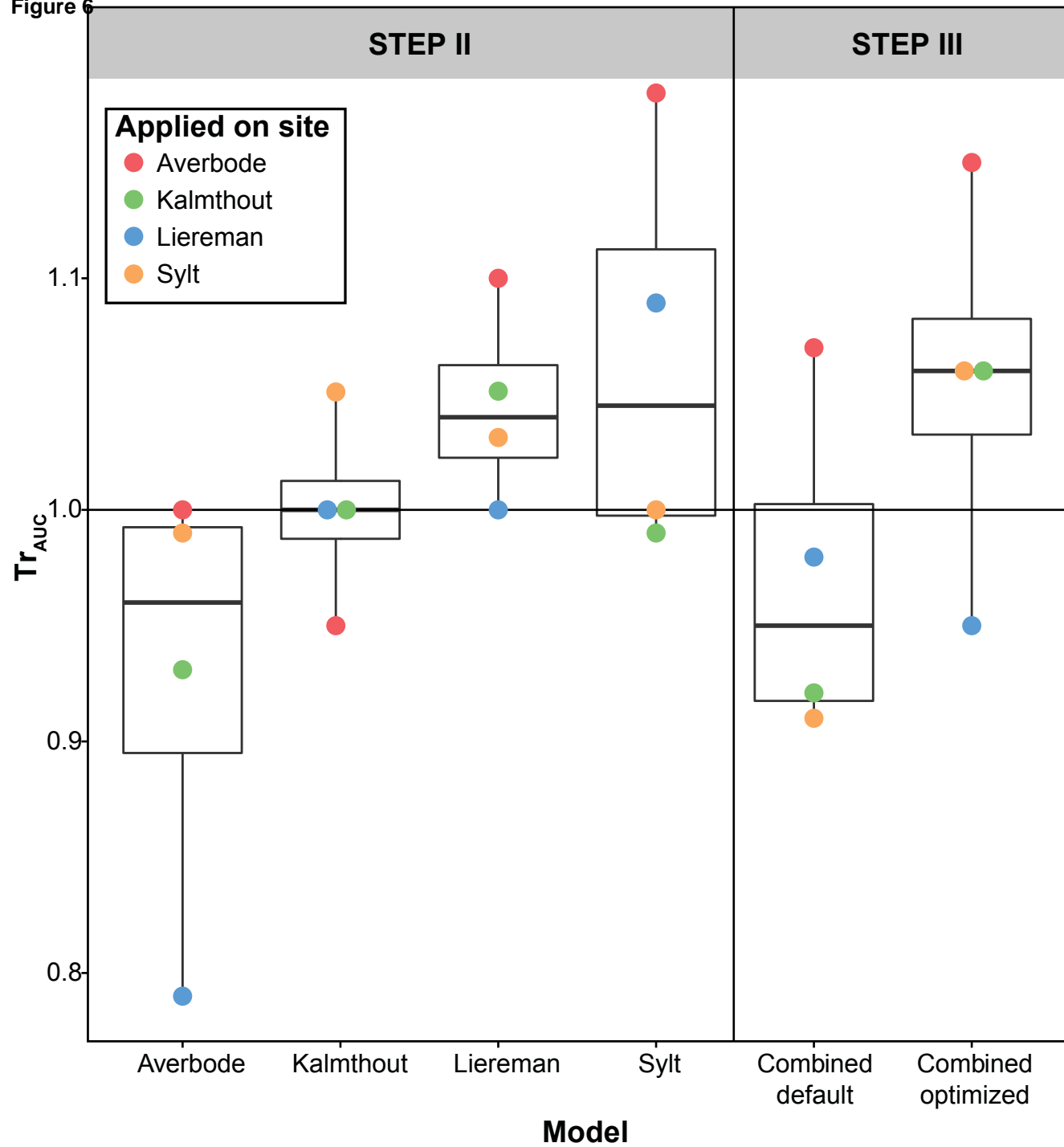


Figure 7
[Click here to download high resolution image](#)

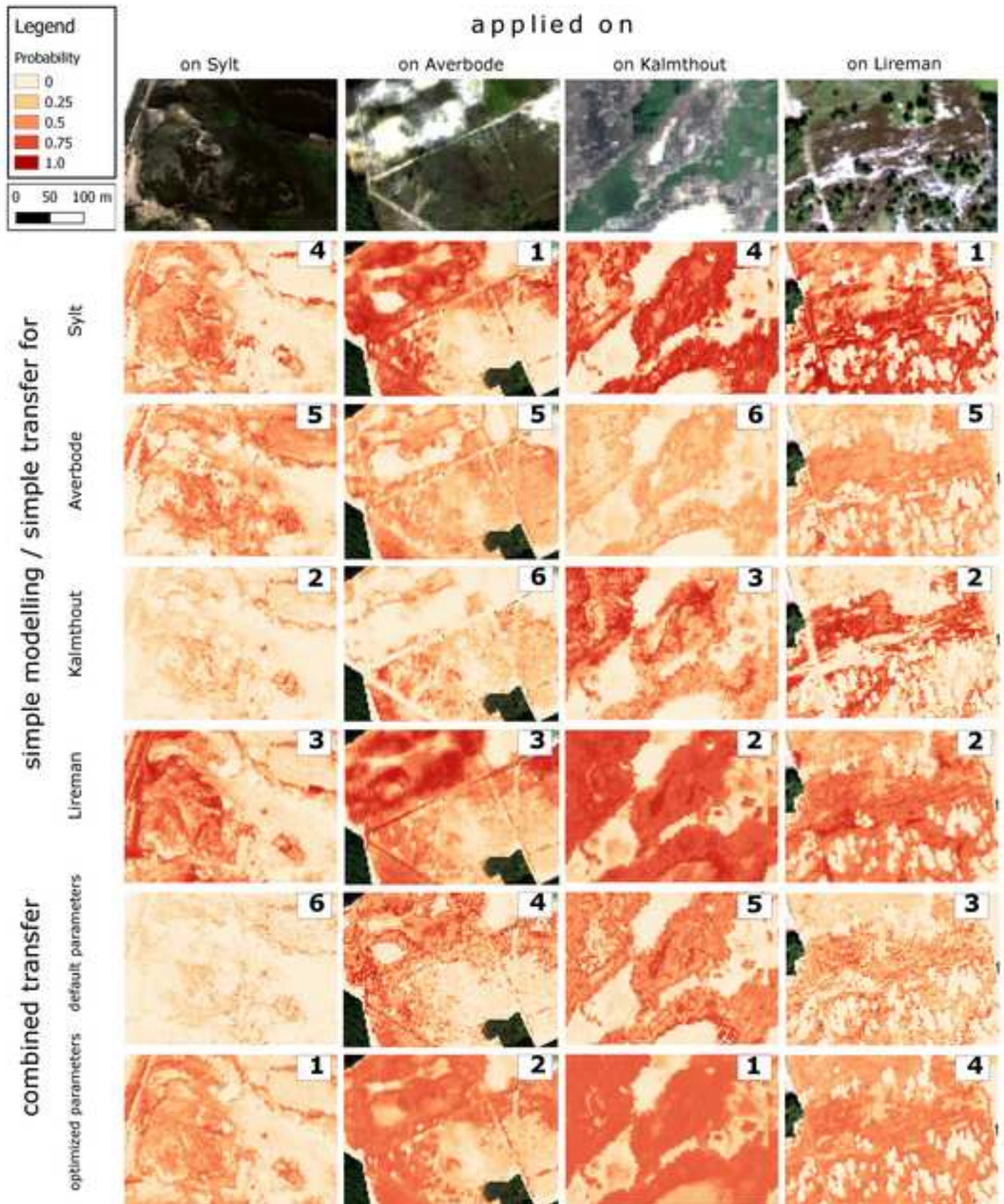
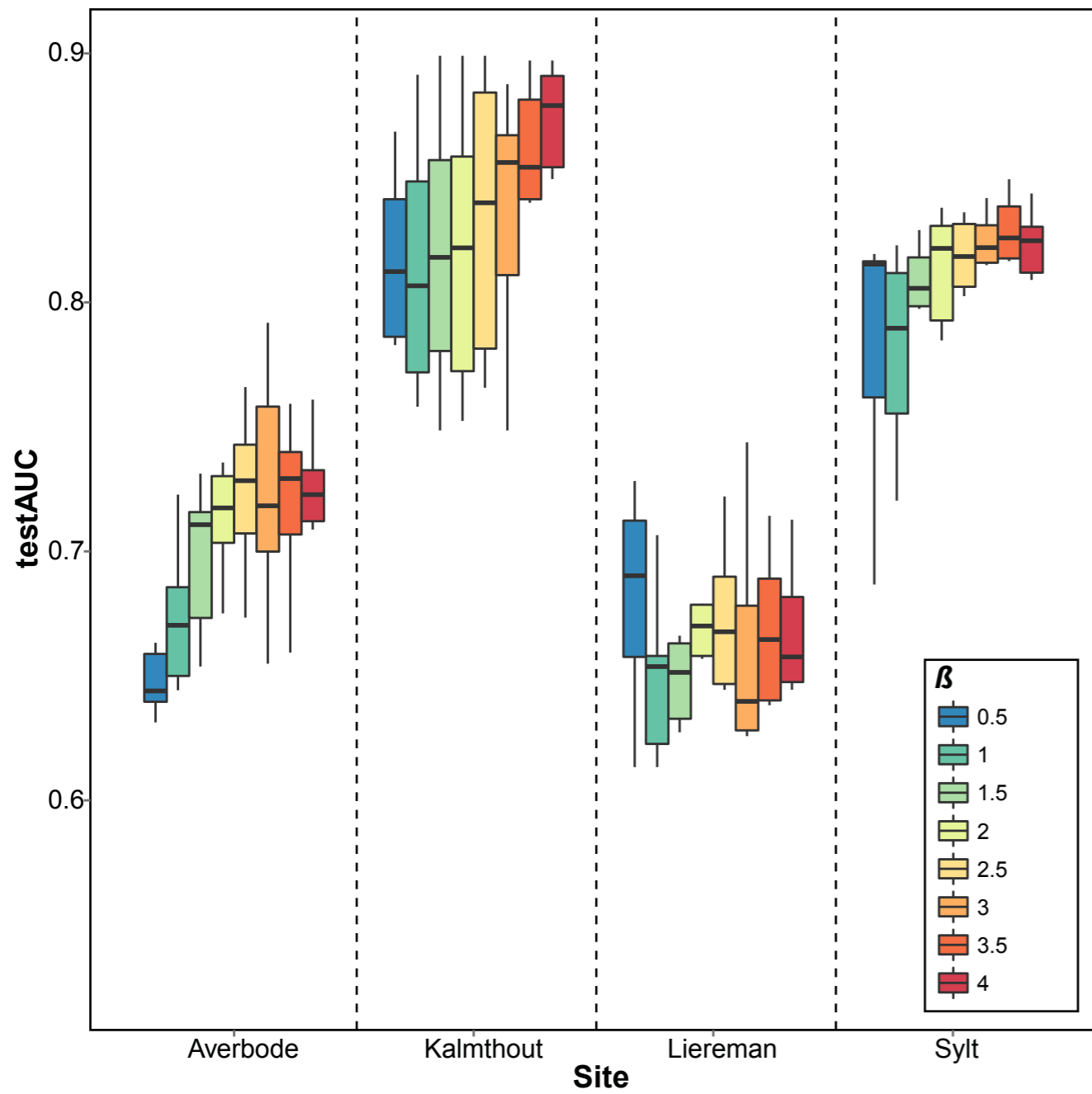
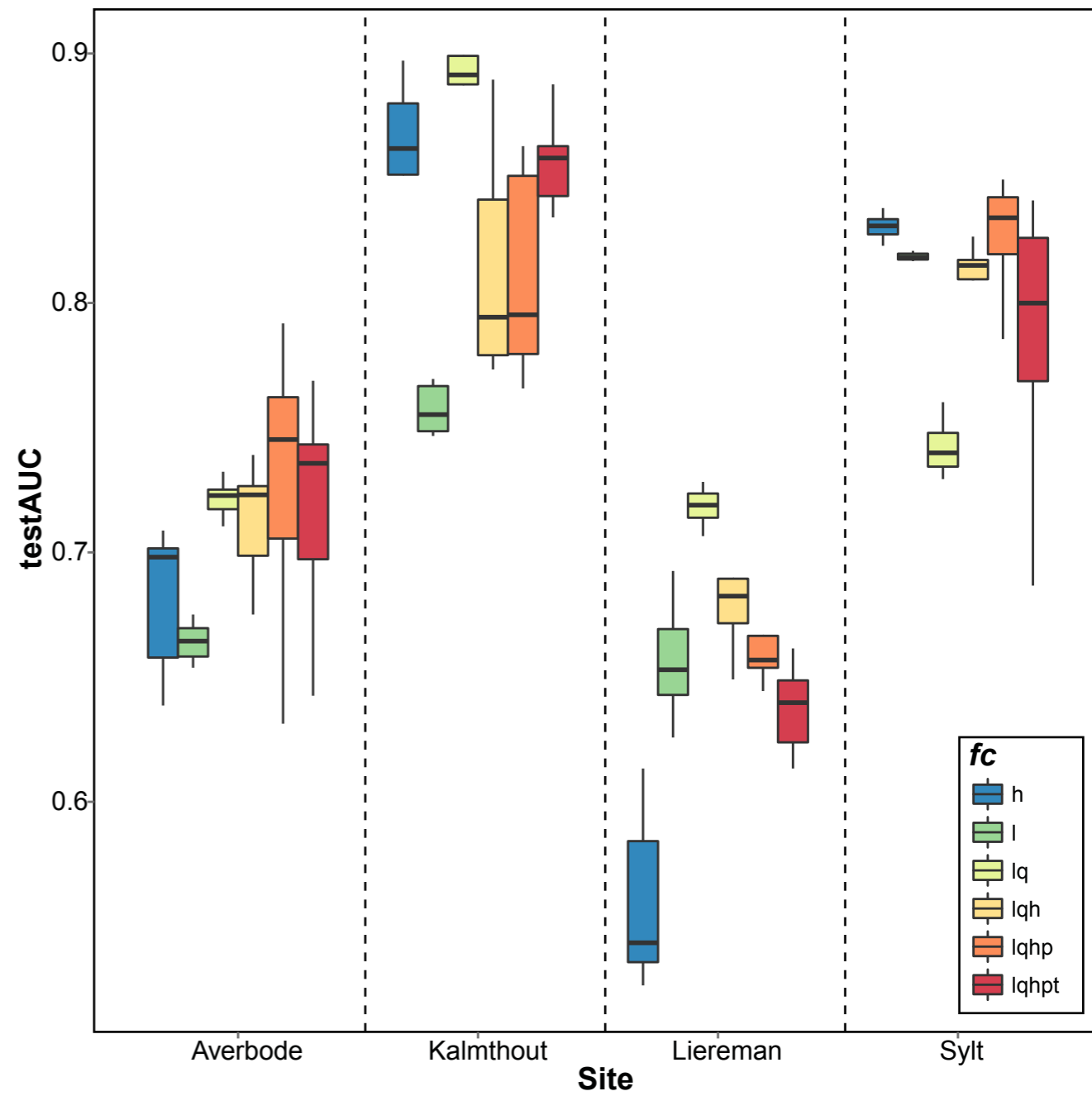


Figure 8

a) β b) fc 

Supplement 1

Occurrence probability of *Campylopus introflexus* according to the model predictions made in Step I (simple modelling), Step II (simple transfer) and Step III (combined transfer)

Averbode



0 500 1000 m





0 500 1000 m





0 500 1000 m







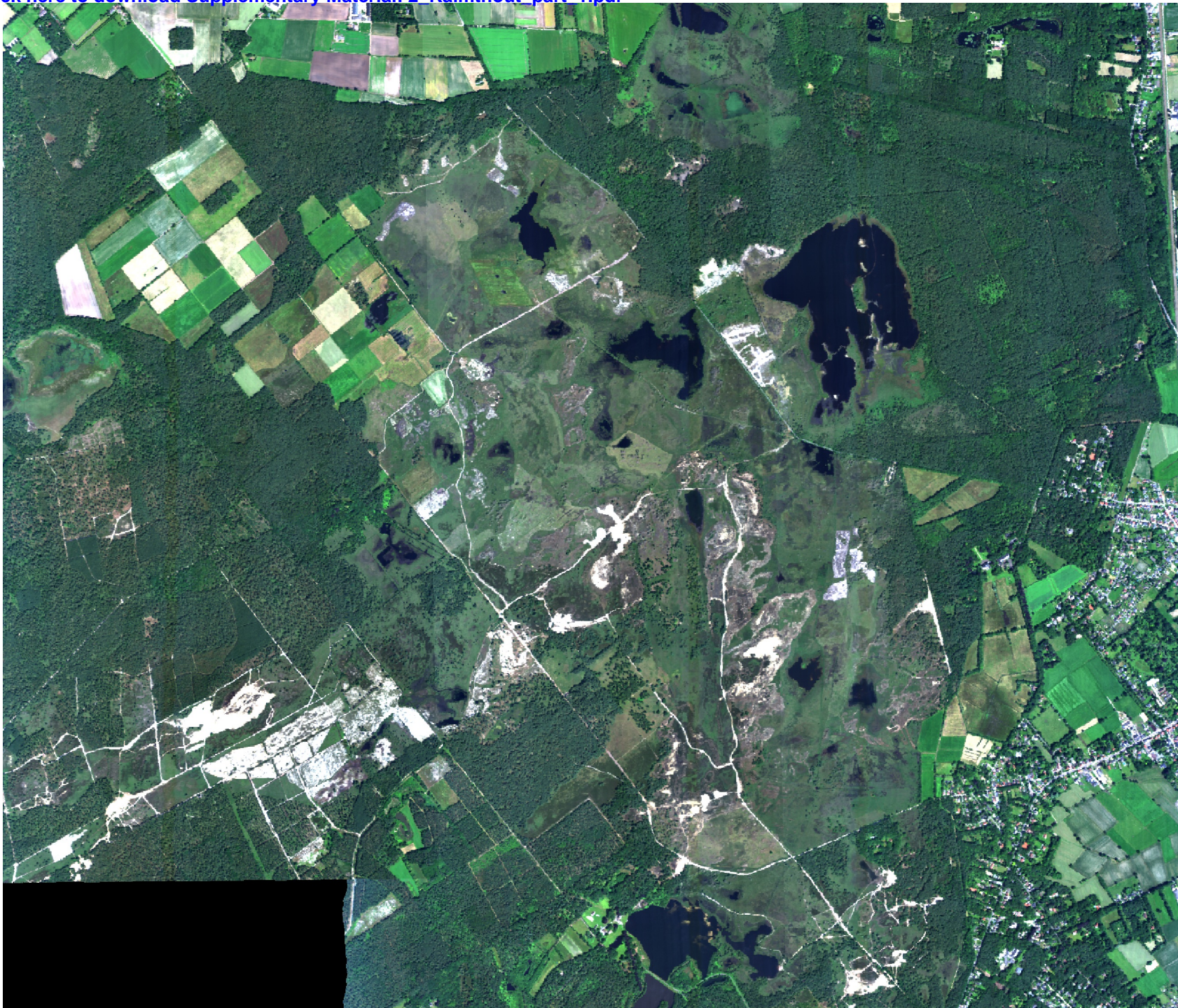
0 500 1000 m





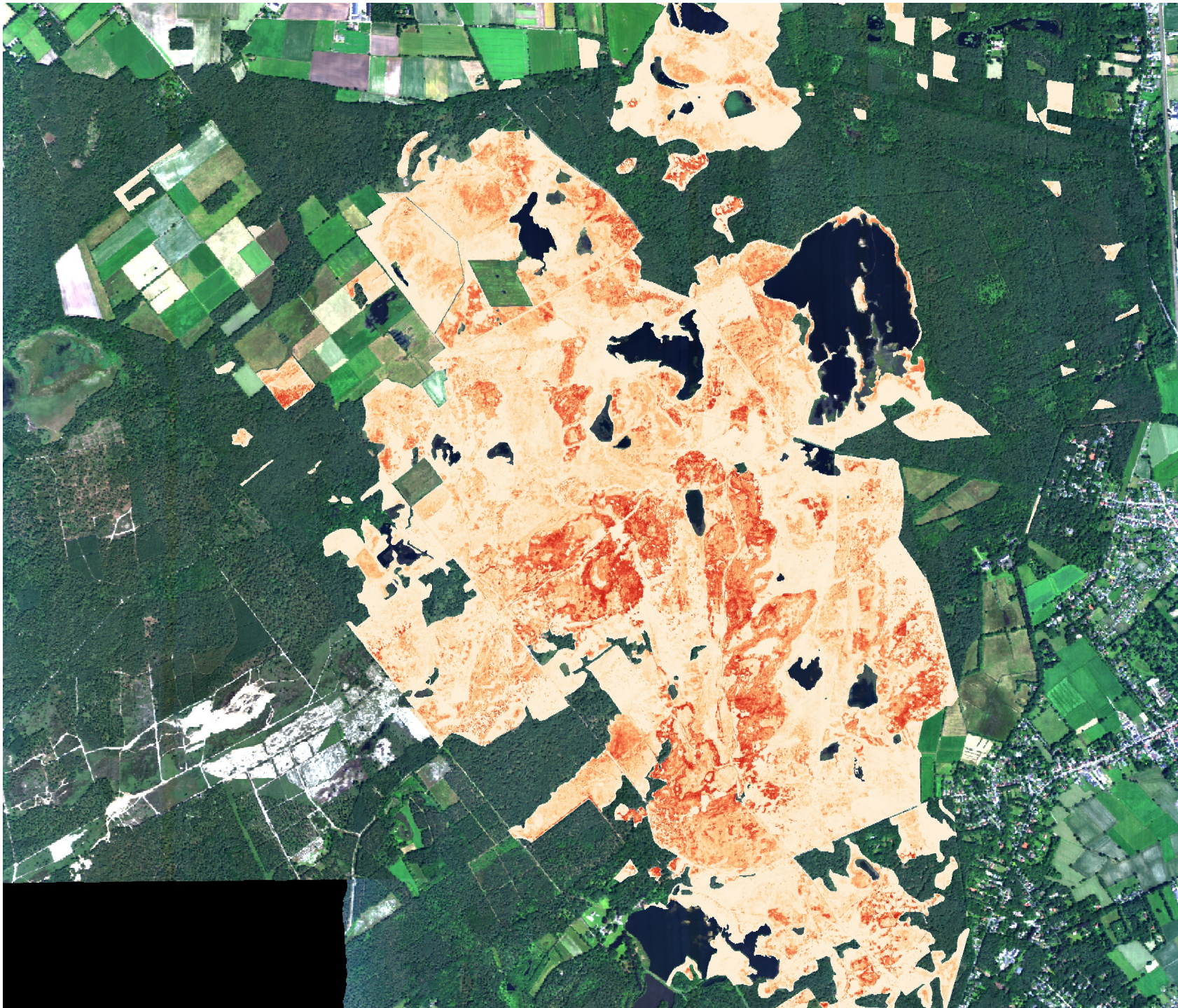


Kalmthout

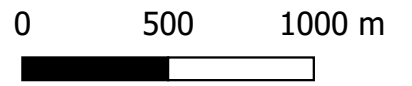
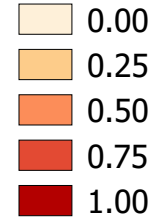


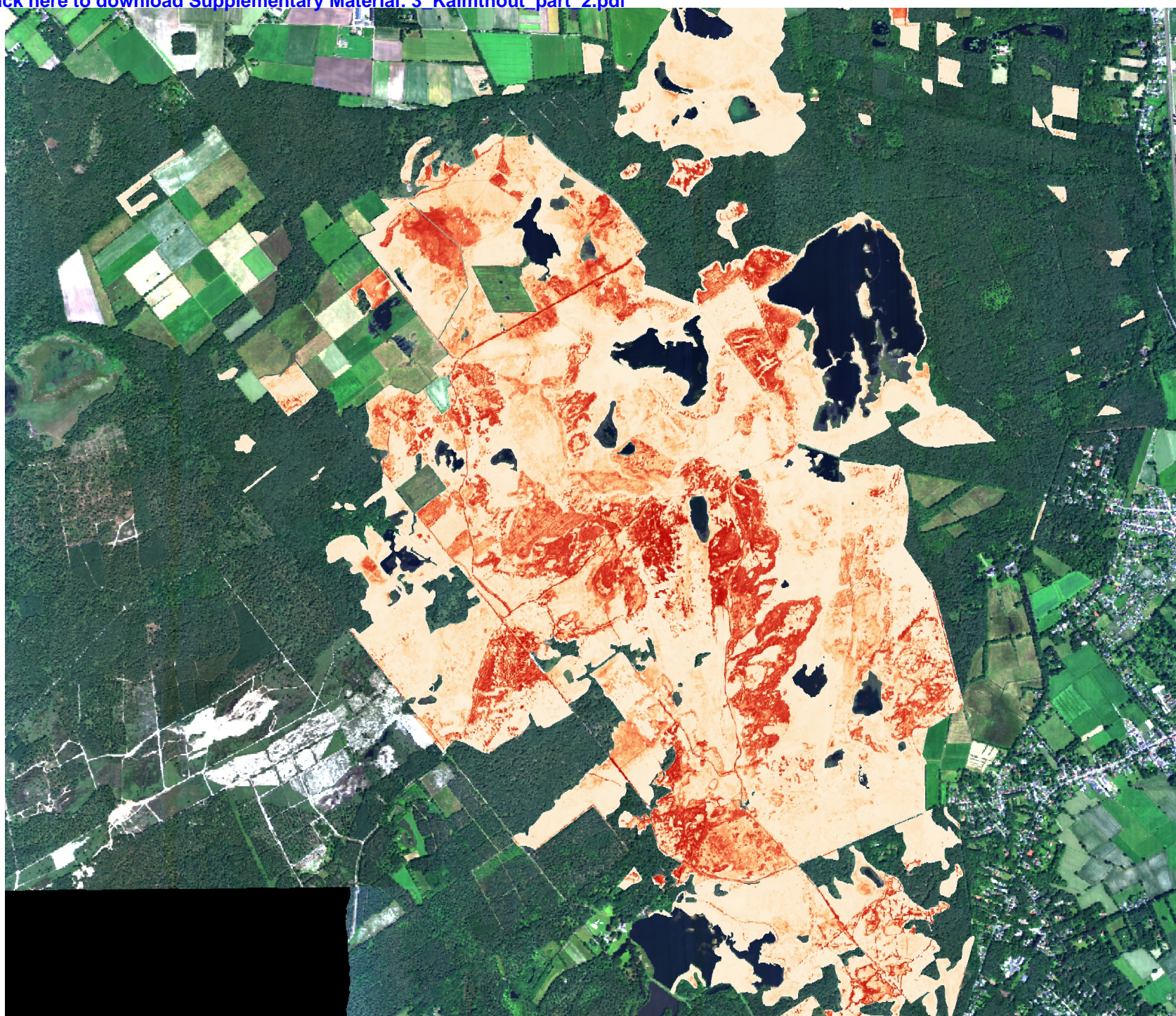
0 500 1000 m



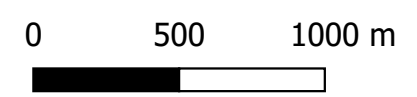
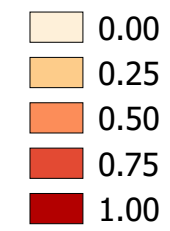


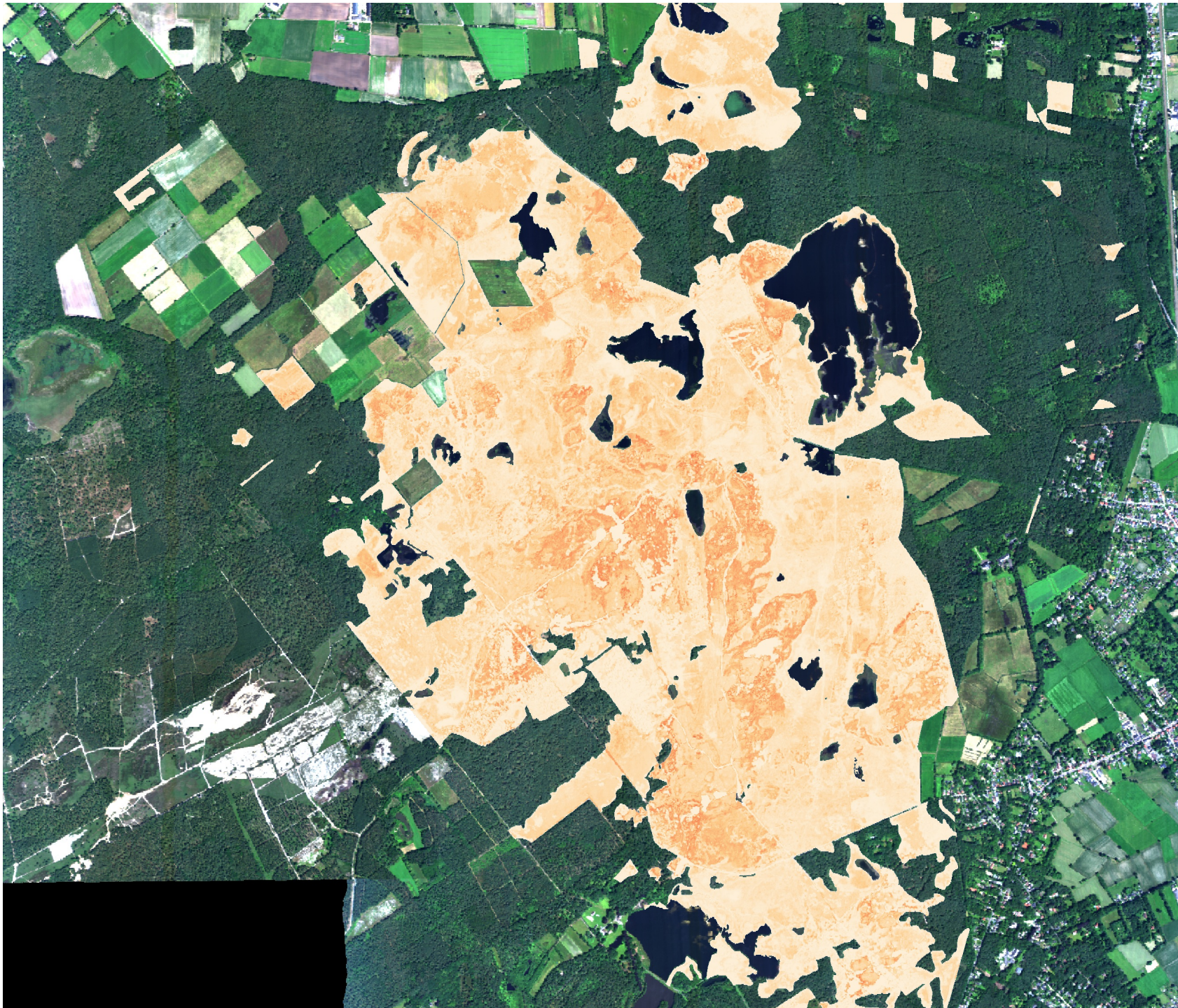
Probability Ka on Ka



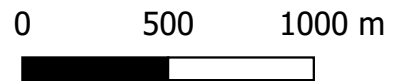
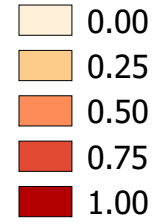


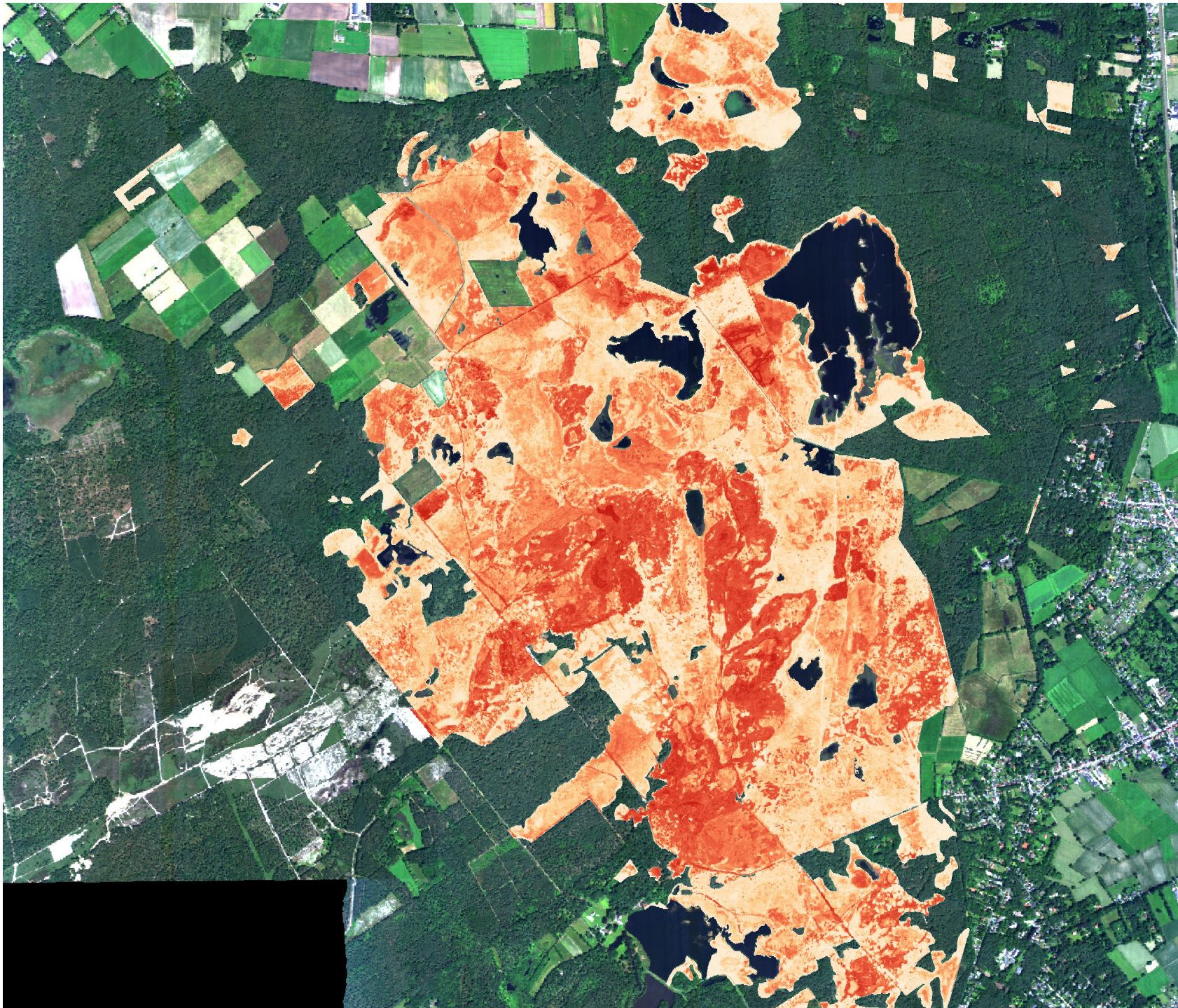
Probability Sy on Ka



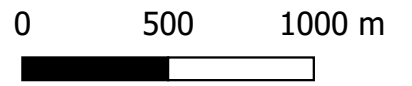
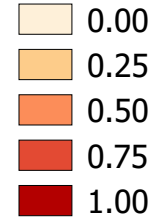


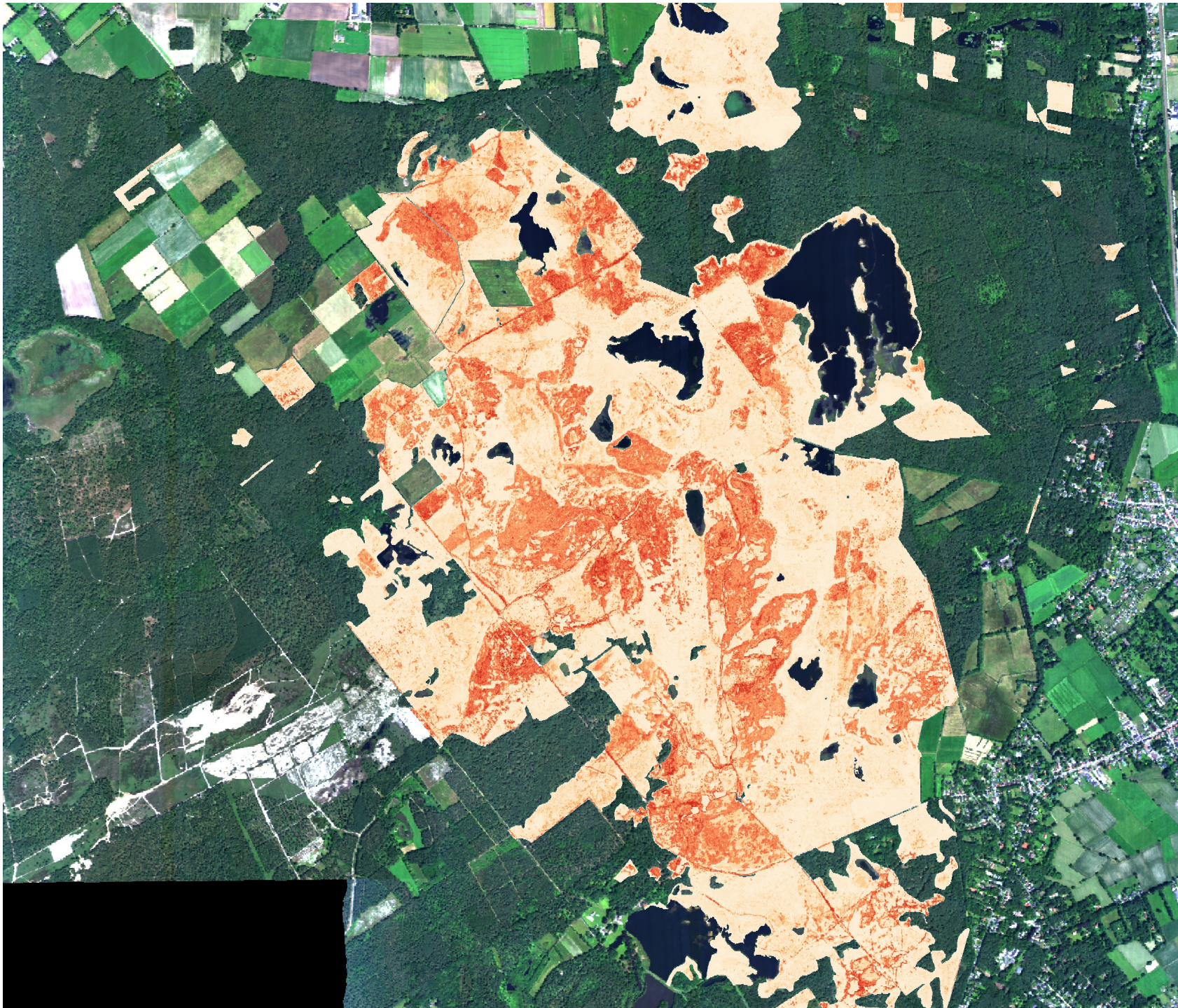
Probability Av on Ka










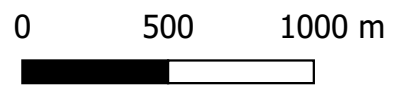
Probability Li on Ka

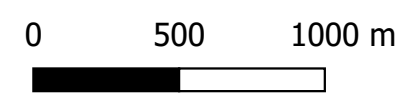
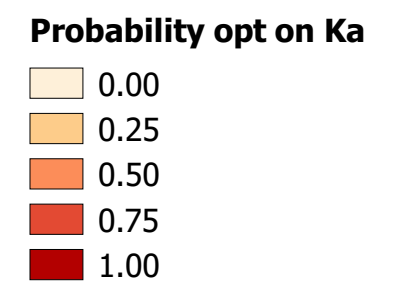
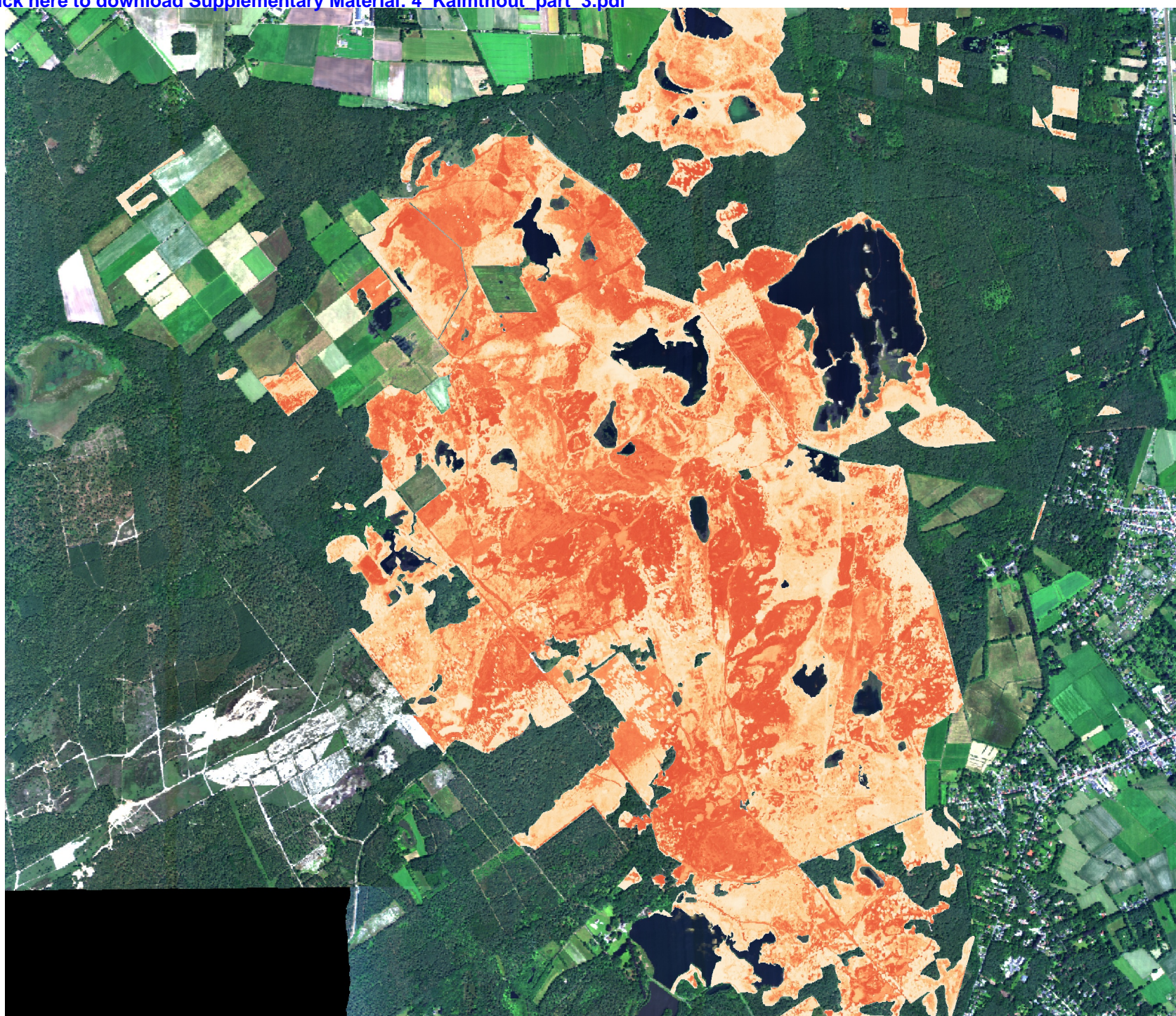




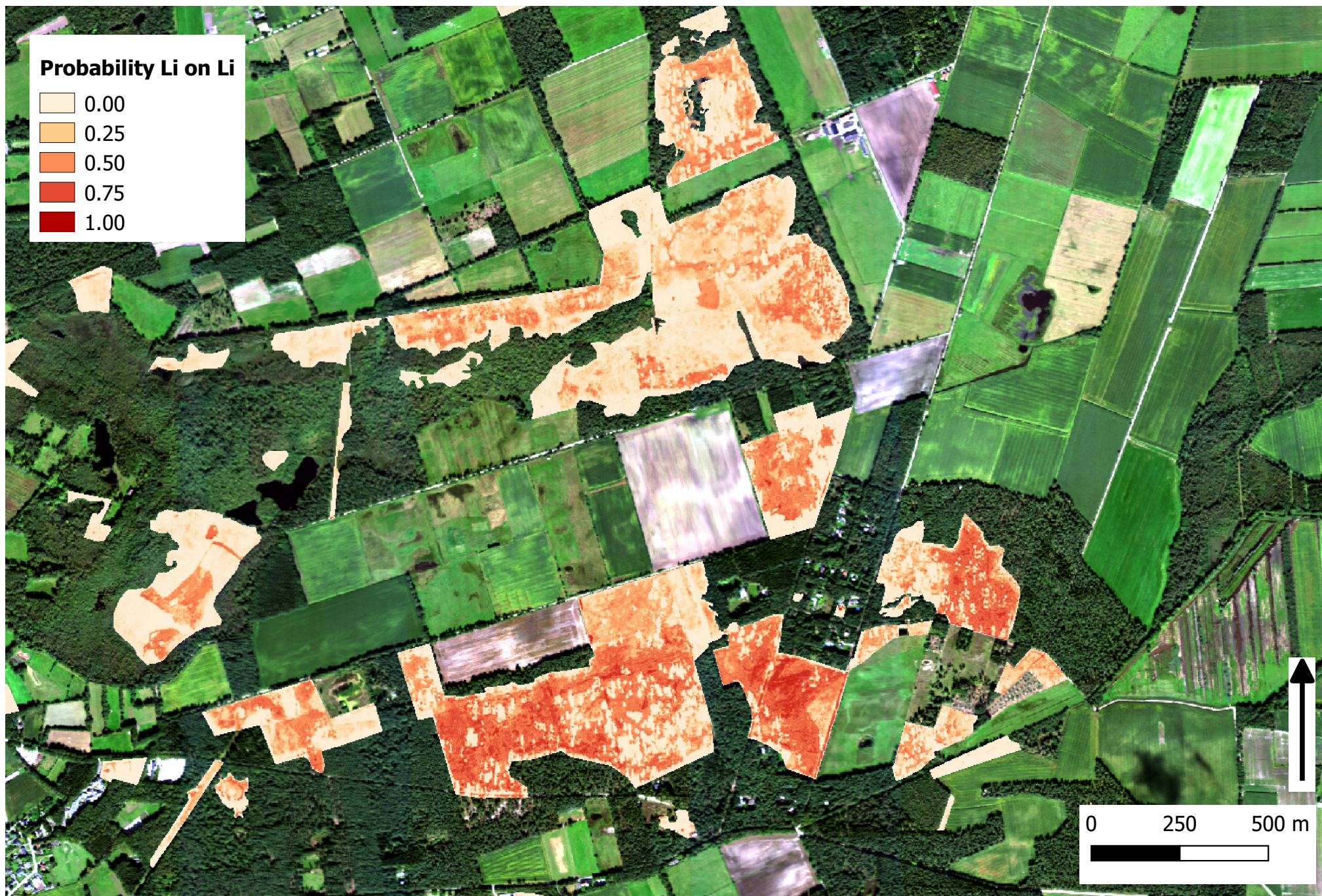
Probability def on Ka

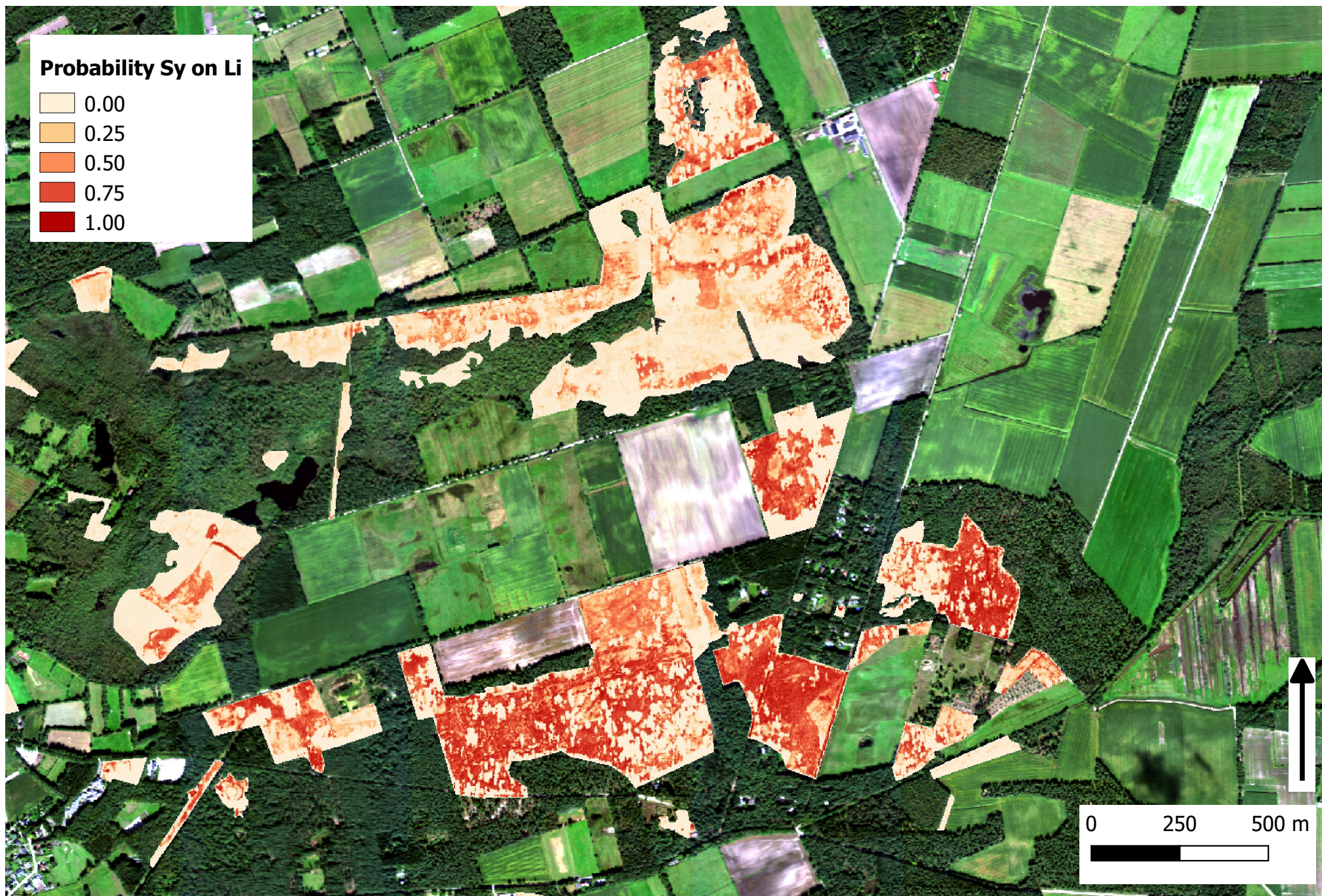
-  0.00
-  0.25
-  0.50
-  0.75
-  1.00





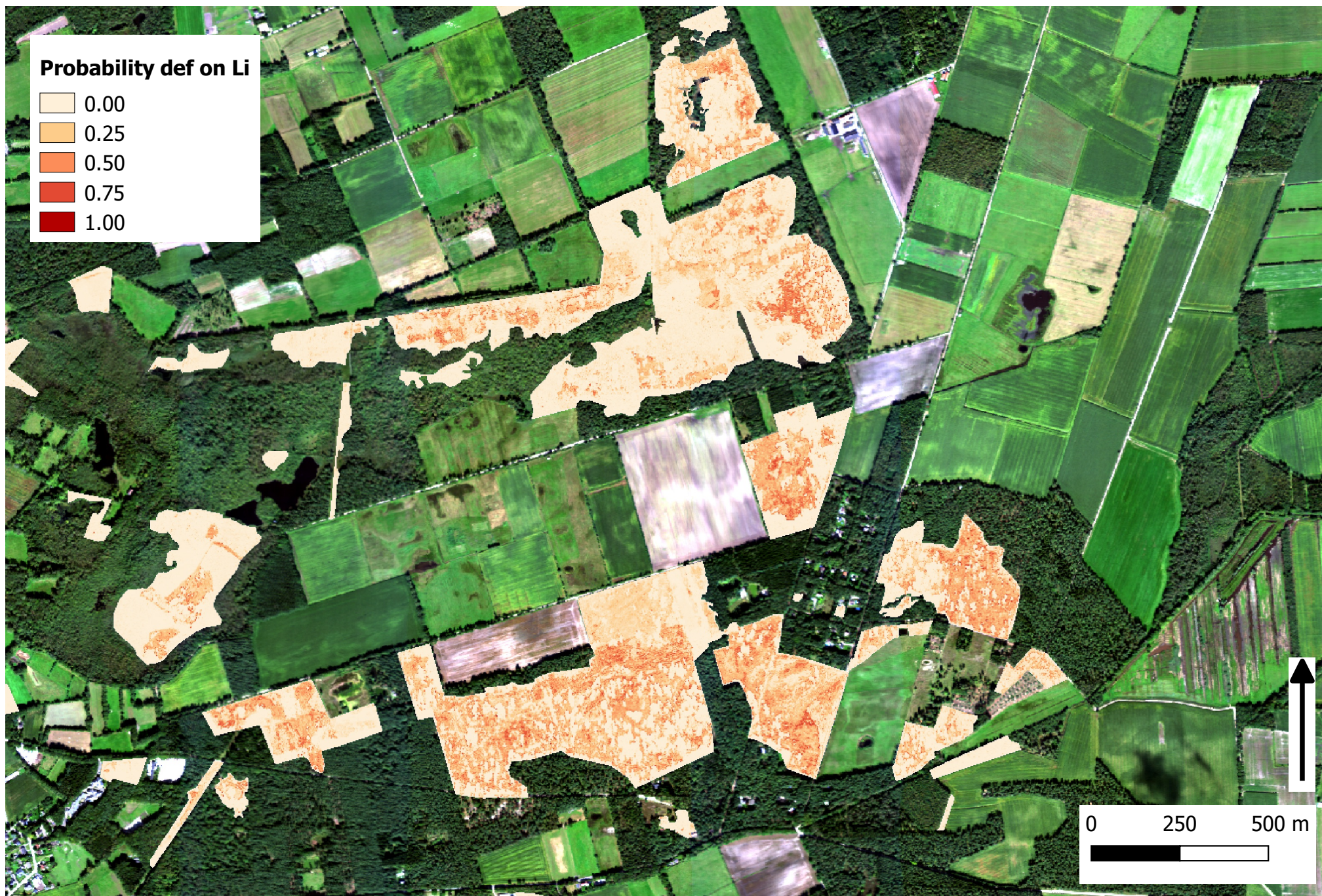










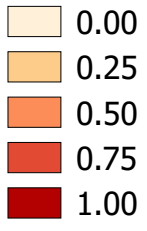




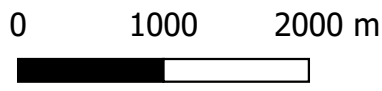
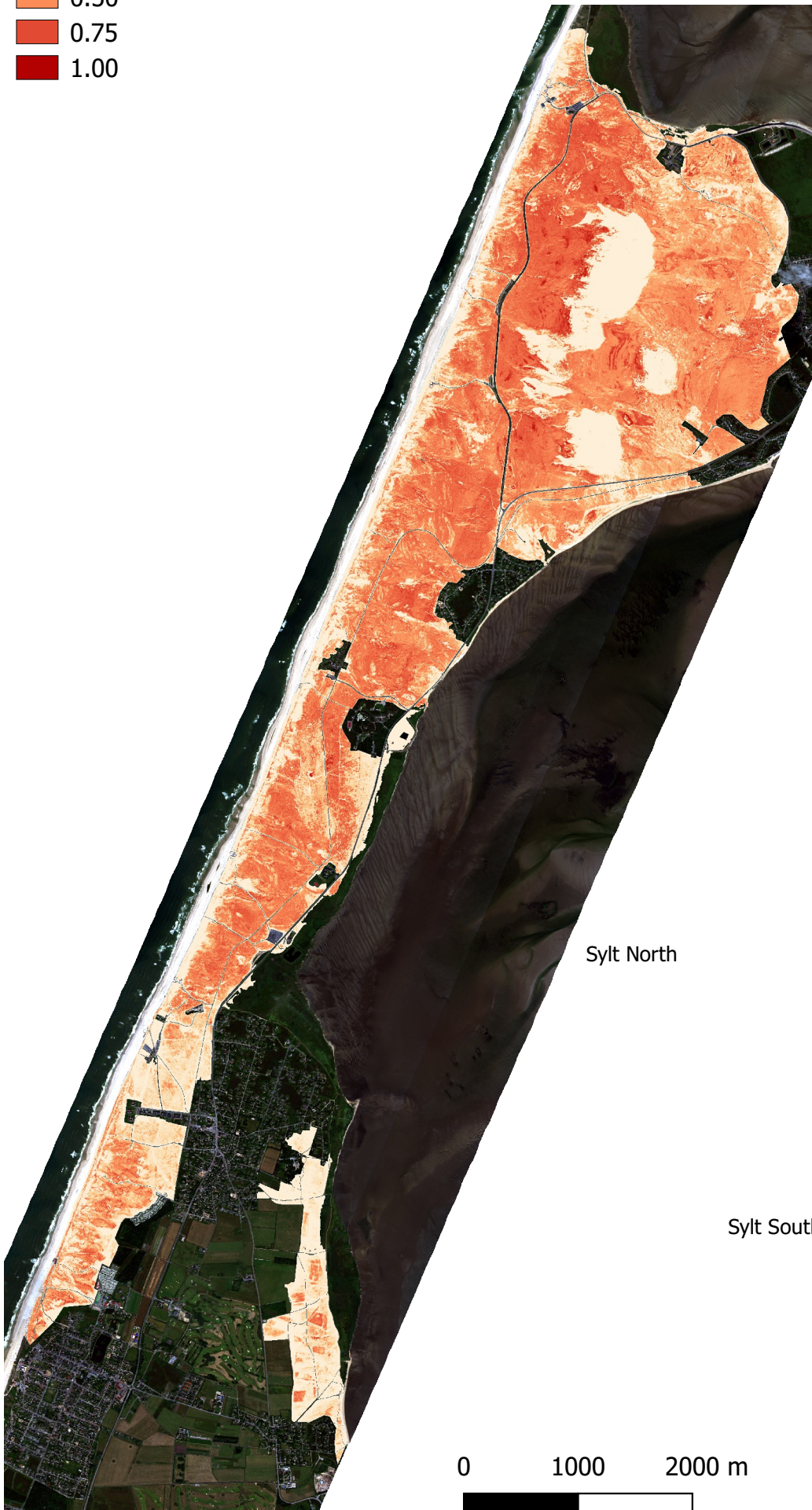
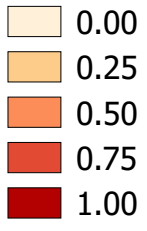
Sylt



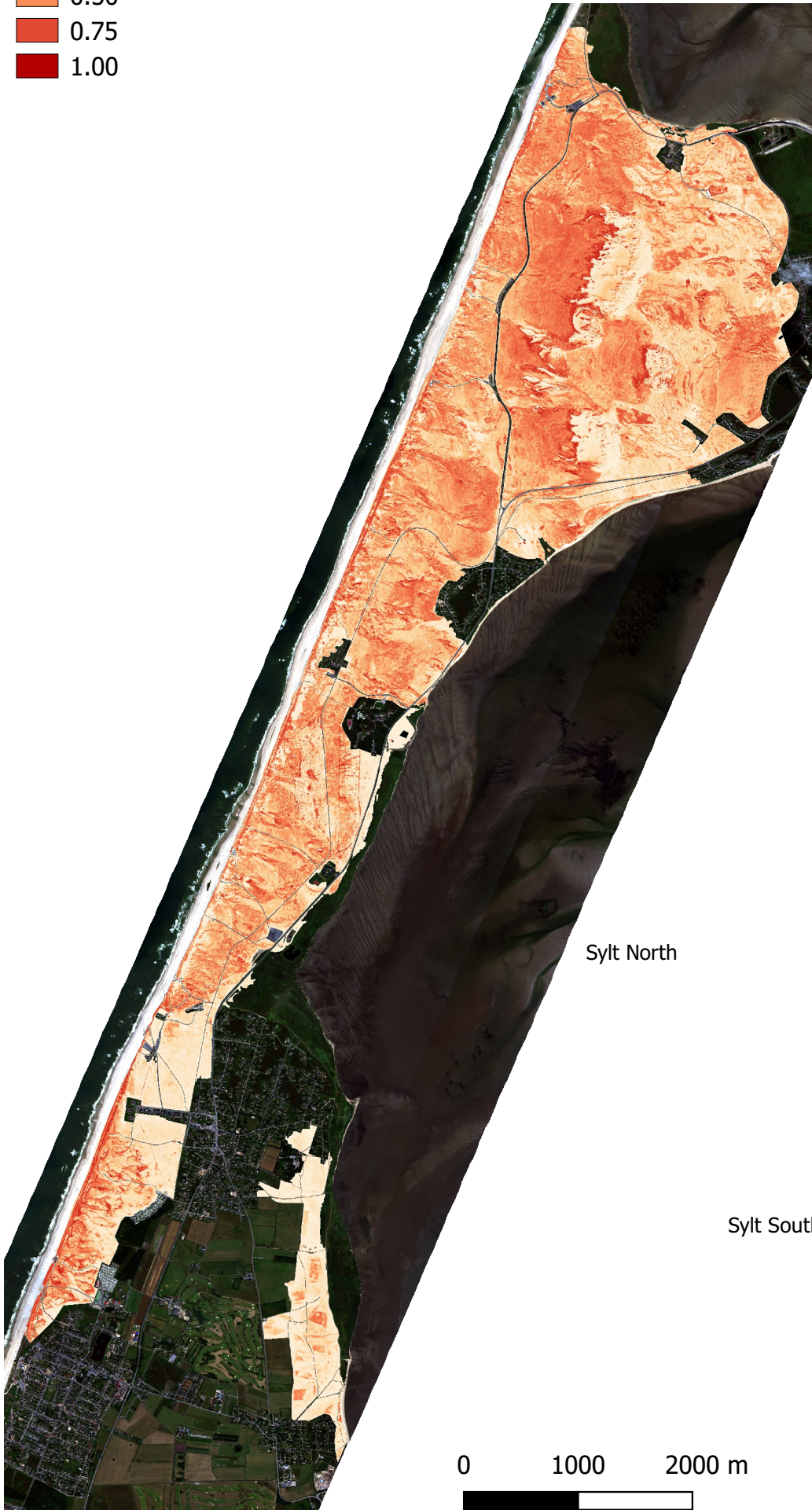
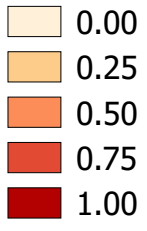
Probability Sy on Sy



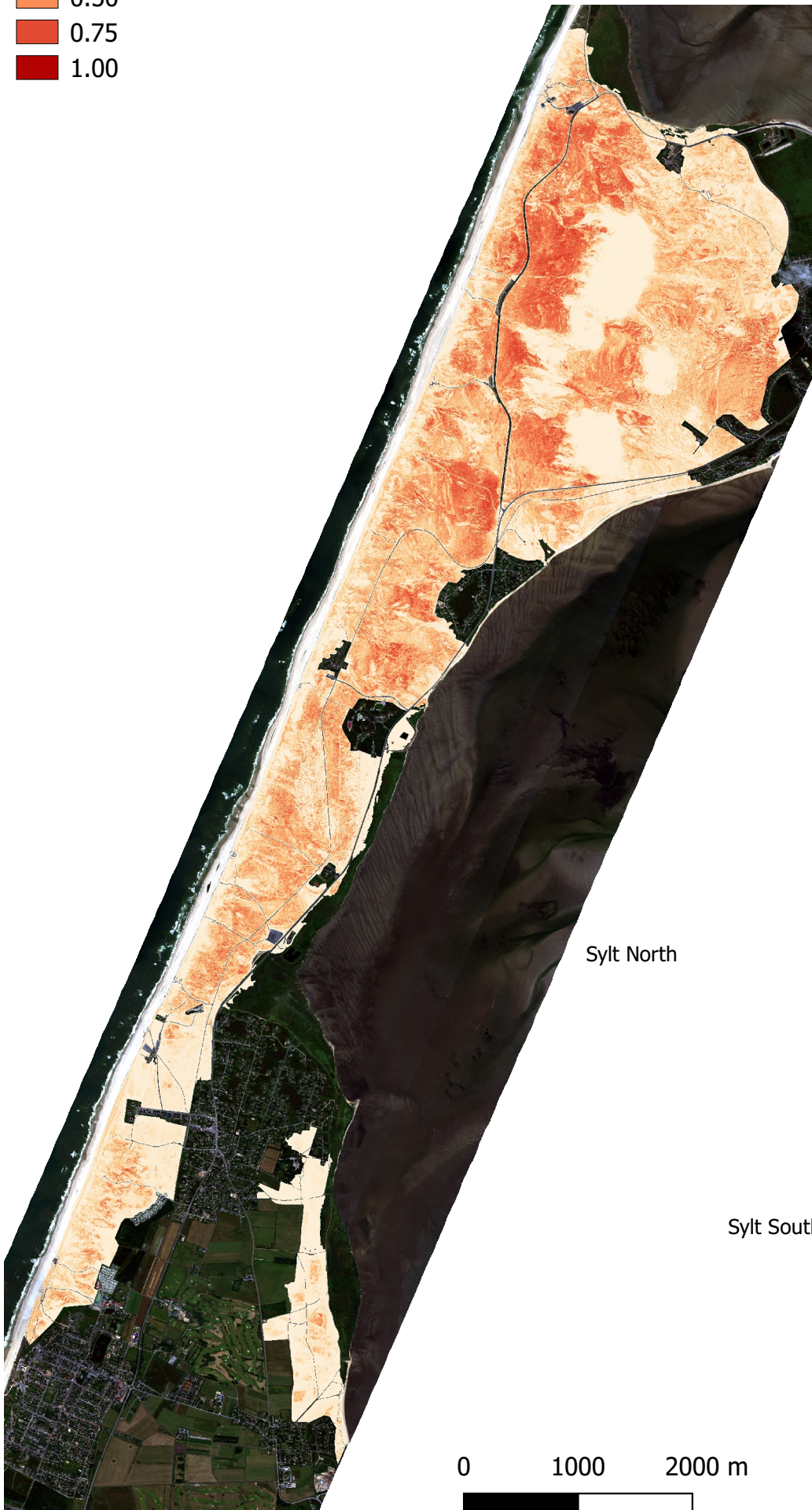
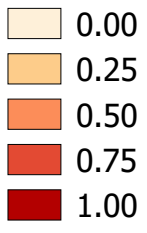
Probability Av on Sy



Probability Li on Sy



Probability Ka on Sy



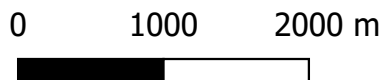
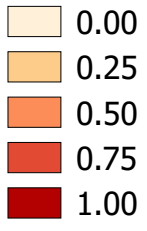
Sylt North

Sylt South

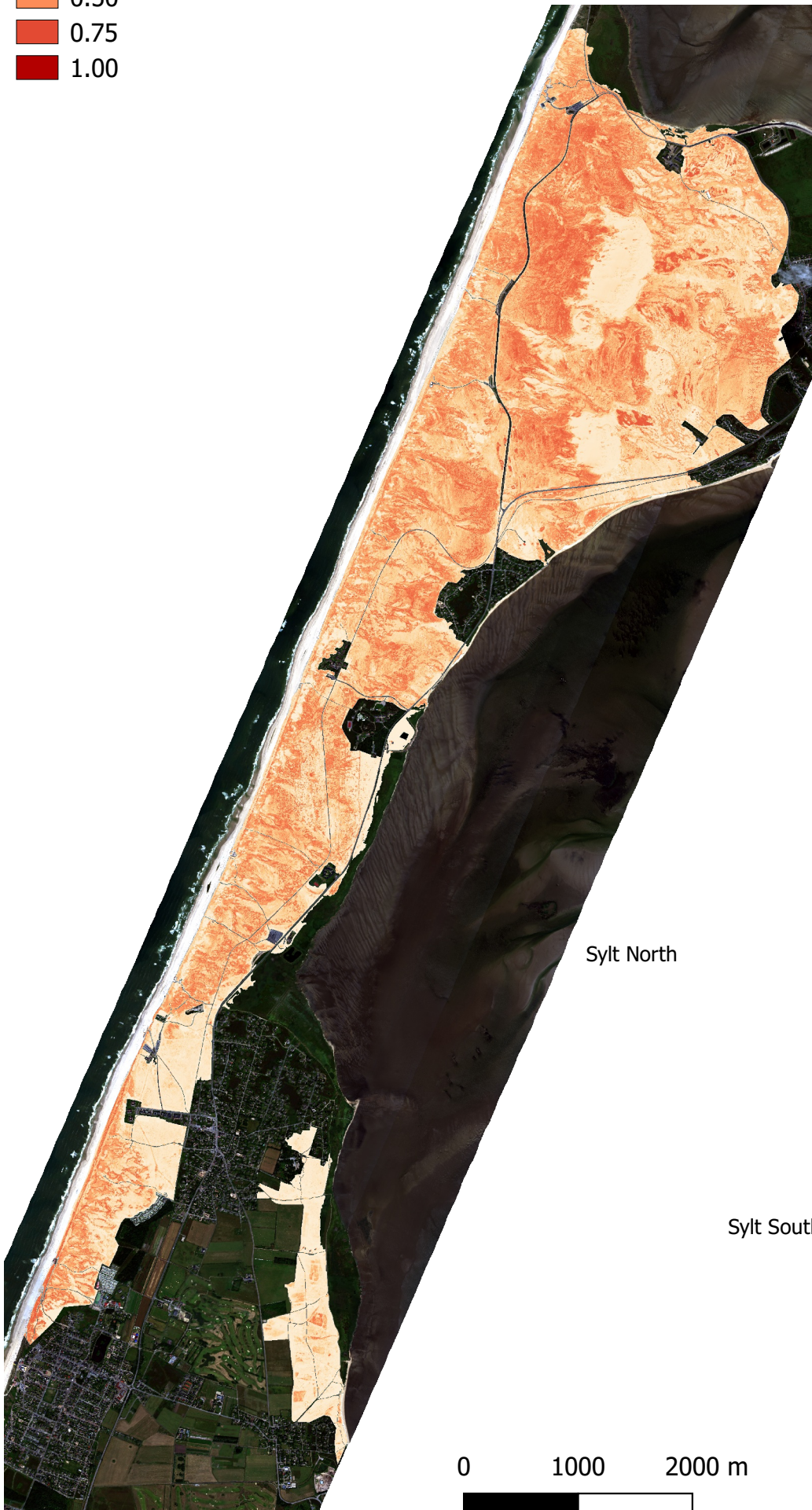
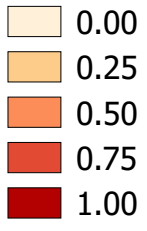
0 1000 2000 m



Probability def on Sy



Probability opt on Sy



Sylt North

Sylt South

0 1000 2000 m



Supplement 2

Reflectances of the calibration and background plots for all four study sites: A – Averbode, B – Kalmthout, C – Liereman and D – Sylt. The yellow line shows the mean reflectance of the calibration plots for each site.

