# 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.

Elsevier Editorial System(tm) for
International Journal of Applied Earth Observation and Geoinformation
Manuscript Draft

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:

***Cover Letter**

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)

***Response to Reviewers**

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** |
|---|---|---|
| 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. | 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. | Added sentence: 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. |
| Another question related to table 1. You have sampled "absence" plots, however Maxent generates absence data itself, right? These are | 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 | Added sentence: In all four study areas, we collected presence data to calibrate the model and presence/absence data to validate the prediction. |

| | | |
|---|---|---|
| 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? | absences. We added a sentence to make this clearer. | |
| 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? | 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. | We added "for each band and" on line 203 to clarify how this spectral variation was calculated |
| 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? | 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 3x10,000=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 3x3,000= 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. | We added the following explanations:<br><br>"The background sample for each study site consisted of a large number of random points located within the biotope types of each study area"<br><br>"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." |
| L 215: "For Step I, we also derived the confusion matrices using kappa | Yes, something got mixed up in that sentence. We did use kappa as a threshold to derive the presence- | "For Step I, we also derived the presence / absence map from the probability maps |

| | | |
|---|---|---|
| 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? | absence map (there are multiple thresholds in the “evaluate” function of the dismo package and kappa worked best). The sentence was revised. | using kappa as threshold. We then also derived the confusion matrices and compared the overall accuracies (OA).” |
| 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 | 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. | |
| 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. | 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. | Changed to “separate independent” |
| L217 OAC -> OA (I never saw OAC before anywhere) please correct throughout the manuscript | Yes, indeed. | Done |

| | | |
|---|---|---|
| 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 | 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. | Done, see new Figure 4 and Maps in Supplement |
| L243: remove blank space after "details)" | Ok | Done |
| L277: "calibration AUC values (between 0.94 and 0.96)" maybe you should state that these values can be found in the appendix? | This seems to be a misunderstanding, as there is no figure in the appendix that contains the calibration AUC values. | |
| L290: write four instead of 4 | Ok | Done |
| L291: were combined instead of was combined? | Ok | Done |
| L374: if these are lower...lower then what? please explain | 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. | Changed “lower” to “low” |
| L396: please write out lq and lqhp - it is not clear here what these mean. | Ok | Done, “mainly lq (linear and quadratic) and lqhp (linear, quadratic, hinge and product)” |
| 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. | 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. | Changed to “mitigate sampling bias”.<br><br>Added sentence:<br>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. |

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).

| | | |
|---|---|---|
| 1. Abstract: It should be continuous (without paragraphs). | Ok | Done |
| 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 | 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.<br><br>We did not observe a significant phenological difference between the dates when the image data was acquired. *C. introflexus* 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.<br><br>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. | Added text<br>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.<br><br>Sentence added to section 4.4:<br>We did not observe a significant phenological difference between the dates when the image data was acquired.<br><br>and<br><br>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 |

| | | |
|---|---|---|
| the dates of image acquisition? Please, clarify in the text. | We thus estimate that neither the time gap between field campaign and image acquisition nor 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. | slight over or underestimation of the abundance of *C. introflexus* on single plots. |
| 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. | We agree and modified the descriptive paragraph accordingly. | 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 |
| 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. | We added the following description: | The data were geometrically and atmospherically corrected using the standard processing applied to APEX (Sterckx *et al.*, 2016; Vreys *et al.*, 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 |

| | | |
|---|---|---|
| | | from available spectral libraries. After atmospheric correction, |
| 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. | 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.<br>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.<br>In Skowronek et al. 2017 (Ecological Informatics) we also compared the performance of Maxent to SVM and BRT and found no major differences. | Sentences added:<br>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. |
| **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. | 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.<br><br>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.<br><br>For differences between backgrounds we also added more information, please see answers to comment 11.<br><br>Additionally we also remade the figure in the supplement by inserted units for the wavelength and adding a legend. | Text added to section 3.1<br><br>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, |

| | | |
|---|---|---|
| | | 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:<br><br>While the most abundant vegetation types on Sylt include *Empetrum nigrum* dominated heathland making up about ⅔ of the study area, other important vegetation types include grey dunes vegetation, *Erica-tetralix* |

| | | |
|---|---|---|
| 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. | 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. | and *Ammophilia arenaria* dominated areas. For the Belgian study sites, the most abundant biotope types sites include vegetation types dominated by *Caluna vulgaris*, *Molinia caeruela* and *Erica tetralix*.<br><br>One Figure was moved frome the supplementary material to the new section 3.1 and explanations were added, see answer to comment 6. |
| 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. | Ok | Done |
| 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. | 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.<br><br>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<br><br>Possible differences in phenological stages: The invasive bryophyte C. introflexus 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. | added sentences:<br>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). |
| 14. Line 458: Define EnMAP, if not done before. | Ok | Done, “EnMAP (Environmental Mapping and Analysis Program)” |
| 15. Please, it is very important to add a Conclusions section. | We added a conclusions section. | Added section:<br><br>5 Conclusions<br>In this study we successfully transferred species distribution models for *Campylopus introflexus* which were calibrated at different sites using airborne imaging spectroscopy as explanatory variables. Our results demonstrate that model |

| | | |
|---|---|---|
| | | 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. |
| 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:<br>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.<br><br>There are other references related to the topic in the | Thanks for those suggestions! The references were revised for missing information.<br><br>We integrated Robinson et al. 2016 and Fernandes et al. 2014. Müllerova et al. 2013 does not support our statement, we decided to cite Müllerova et al. 2017 instead as it fits best. | Sentence added to the discussion:<br>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). |

| | | |
|---|---|---|
| JAG journal that you can eventually consider in the literature review or discussion of results:<br>- 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.<br>- 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.<br>- 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.<br><br>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

**Sandra Skowronek[1], Ruben Van De Kerchove[2], Bjorn Rombouts[3], Raf Aerts[3,4], Michael Ewald[5], Jens Warrie[4], Felix Schiefer[1], Carol Garzon-Lopez[6], Tarek Hattab[6], Olivier Honnay[4], Jonathan Lenoir[6], Duccio Rocchini[7], Sebastian Schmidtlein[5], Ben Somers[3], Hannes Feilhauer[1]**

[1]Institute of Geography, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany*
[2]Flemish Institute for Technological Research (VITO), Mol, Belgium
[3]Division of Forest, Nature and Landscape, Department of Earth and Environmental Sciences, KU Leuven, Leuven, Belgium
[4]Ecology, Evolution and Biodiversity Conservation Section, Department of Biology, KU Leuven, Leuven, Belgium
[5]Institute of Geography and Geoecology, Karlsruhe Institute of Technology, Karlsruhe, Germany
[6]Unité de Recherche « Ecologie et dynamique des systèmes anthropisées » (EDYSAN, FRE 3498, CNRS-UPJV), Université de Picardie Jules Verne, Amiens, France
[7]Department of Biodiversity and Molecular Ecology, Fondazione Edmund Mach, San Michele all'Adige, Italy

*** Correspondence:**
Sandra Skowronek
Sandra.skowronek@fau.de

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

## 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.

**Declarations of interest:** none

**Funding**

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

**Acknowledgments**

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

# 1 Introduction

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

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

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

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

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

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

# 2 Materials and Methods

## 2.1 Study areas and study species

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

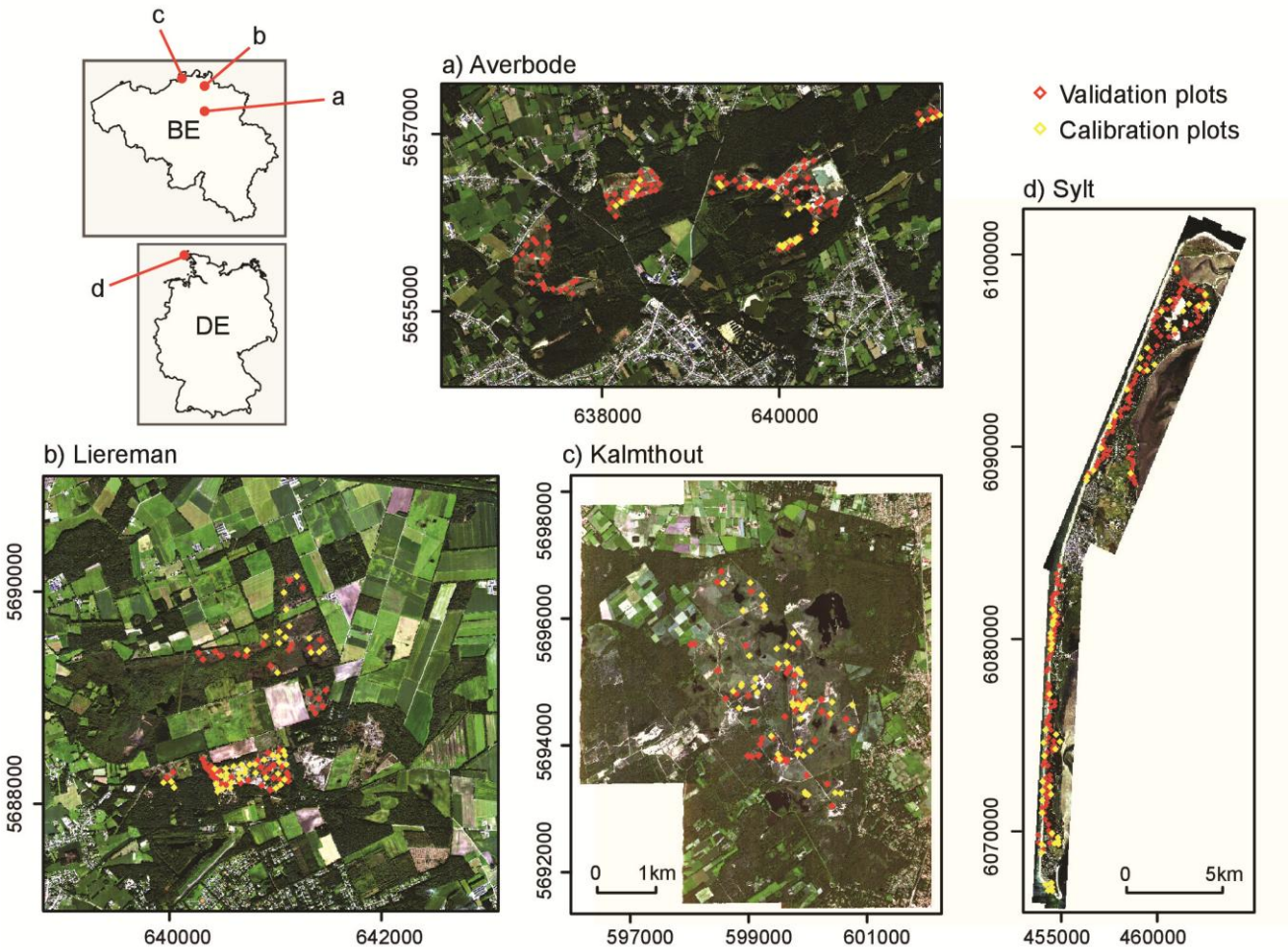



Figure 1: Study areas (a) Averbode Bos & Heide (Av); (b) Landschap de Liereman (Li); (c) Kalmthoutse Heide (Ka) and (d) Sylt (Sy). A true colour composite derived from the APEX data is used as background

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

abundant vegetation types on Sylt include *Empetrum nigrum* dominated heathland making up about ⅔ of the study area, other important vegetation types include grey dunes vegetation, *Erica-tetralix* and *Ammophilia arenaria* dominated areas. For the Belgian study sites, the most abundant biotope types sites include vegetation types dominated by *Caluna vulgaris*, *Molinia caeruela* and *Erica tetralix*. Sylt covers an area of 24.2 $km^2$ and the Kalmthout study site covers 8.0 $km^2$. The two other sites, Liereman and Averbode, are significantly smaller and cover 1.4 $km^2$ and 1.2 $km^2$, respectively.

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

## 2.2 Data acquisition

Field and remote sensing data were acquired between 2013 and 2015 (Table 1). In each of the study areas, a stratified sampling approach was used to lay out a set of 3 m x 3 m calibration (presence) plots, while a random sampling approach was used for laying out validation plots (Fig. 1, Table 1). In all four study areas, we collected presence data to calibrate the model and presence/absence data to validate the prediction. 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. For all plots, the cover of *C. introflexus* was recorded by dividing the plot in four equal parts and visually estimating and summing up the cover of *C. introflexus* on each of the subplots. For Liereman and Kalmthout, a differential GPS (Trimble GeoExplorer 6000) was used to determine the plot position and a differential correction was applied after data collection, while for Kalmthout and Averbode, no differential correction could be performed, as the device (Ashtech mobile mapper 10) did not allow for this feature. All positions are averages of at least 100 measurements.

Airborne imaging spectroscopy data, acquired by the Airborne Prism EXperiment (APEX) spectrometer were 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 by the Flemish Institute of Technology (VITO, Mol, Belgium) with different spatial 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 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 flights 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.

The data were geometrically and atmospherically corrected using the standard processing applied to APEX (Sterckx et al., 2016; Vreys et al., 2016) at VITO's Central Data Processing Center. The processing chain is based on the MODTRAN4 software (Berk et al., 1999) 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 from available spectral libraries. After atmospheric correction, bands from both ends of the spectra and bands disturbed by water absorption were removed (bands between 1320-1447 nm and 1762-1988 nm: selected based on visual interpretation, i.e. noisy profile). Thus, a total of 244 spectral bands (between 426 nm and 2425 nm) were used in the subsequent analyses.

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

| **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<br>(48 p, 102 a) | 93<br>(66 p, 27 a) | 51<br>(28 p, 23 a) | 50<br>(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 | | | |

**2.3 Data analysis**

All species distribution models were built with Maxent (Phillips et al., 2004), a one-class classifier, which differentiates the target species from a background sample based on the comparison of probability densities. Maxent makes an estimate of the ratio between the conditional density of the predictors at the presence sites and the unconditional density of the predictors across the study area, where the distance between those densities is minimized. The logistic output of the model represents an estimate of the probability that the species is present in a certain location. For detailed information on the model, see Phillips et al. (2006) and Elith et al. (2011). 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. (2017a) 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.

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

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

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

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

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

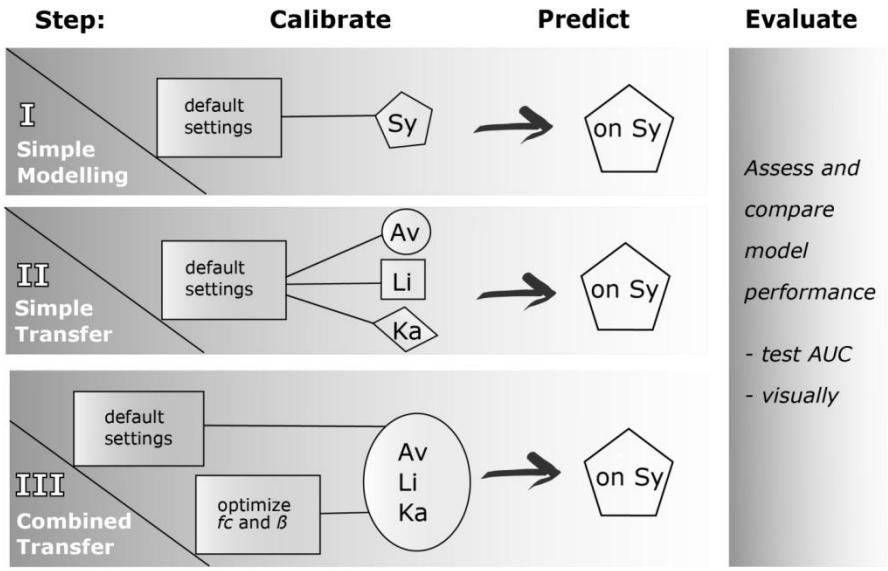



Figure 2: Workflow for each study area using one study site (Sylt) as an example

We made use of the default settings for Maxent (*ß*=1, *fc*=default, 10,000 background points) for Step I and II (simple modelling and simple transfer), whereas in Step III (combined transfer) we also tested the effect of varying the model parameters *fc* and *ß*, as our results using the default settings indicated a highly complex and possibly overfit model (see section 3.3). We tested *ß* values between 0.5 and 4 at 0.5 intervals, as values above the default have been found to produce better results (Radosavljevic and Anderson, 2014; Warren et al., 2014) as well as different combination of the feature classes linear (l), quadratic (q), hinge (h), product (p), and threshold (t), the model being restricted to the following feature classes: lq; lqp; h; qh; qhp; qhpt. 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. The Akaike information criterion (AIC) was used to select the best model (Warren and Seifert, 2010).

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

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

$$\mathrm{Tr}_{\mathrm{AUC}} = \frac{\mathrm{testAUC}_{\mathrm{Site1 \rightarrow Site2}}}{\mathrm{testAUC}_{\mathrm{Site2 \rightarrow Site2}}} \qquad (1)$$

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

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

# 3 Results

## 3.1 Spectral reflectance of the calibration and background plots across sites

Figure 3 shows the mean reflectance and the spectral variability for all four study sites for the calibration plots as well as the background plots. 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 showed the highest mean reflectance in the SWIR while 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, 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.

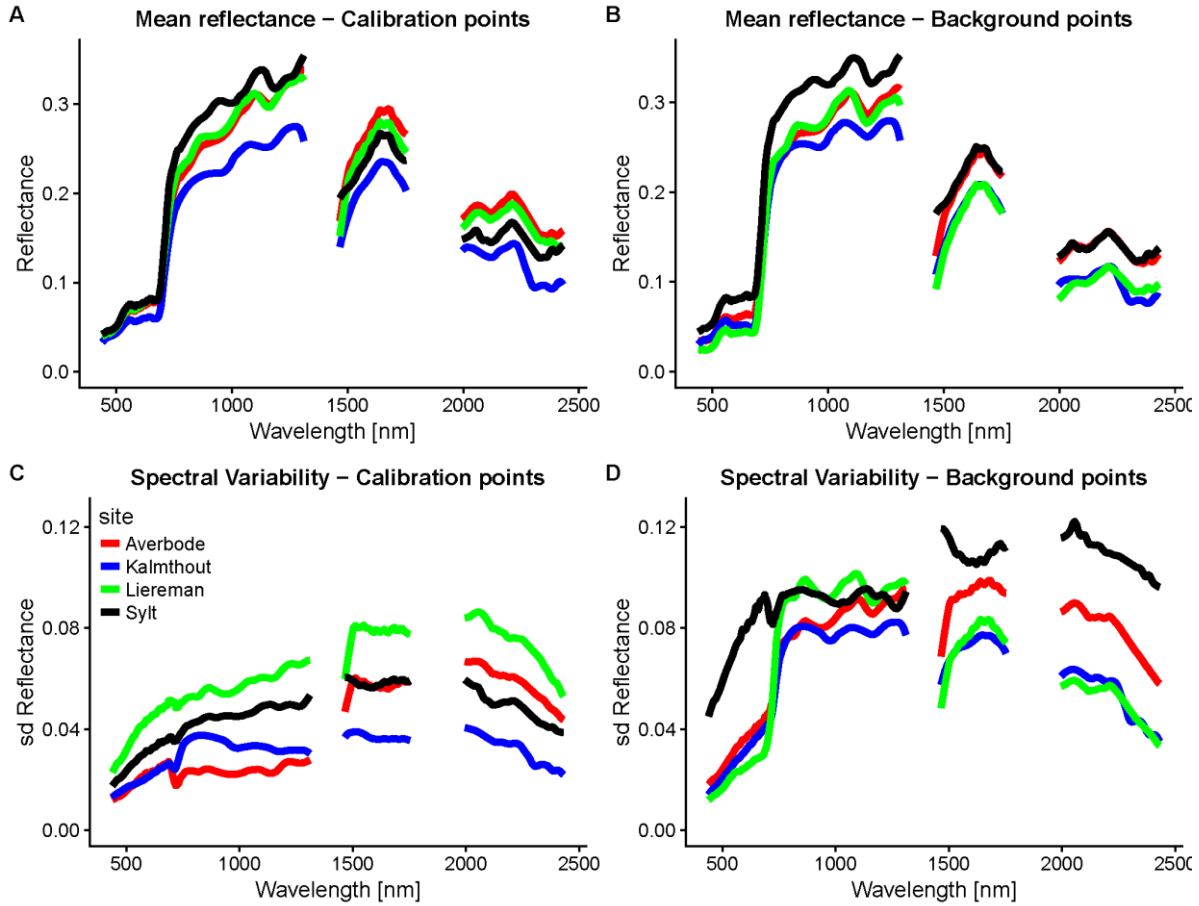


Figure 3: Mean reflectance and spectral variability of the calibration and the background data (measured as the standard deviation per band for all plots) in a spectral range of 380 – 2500 nm.

### 3.2 Simple modelling and band importance

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

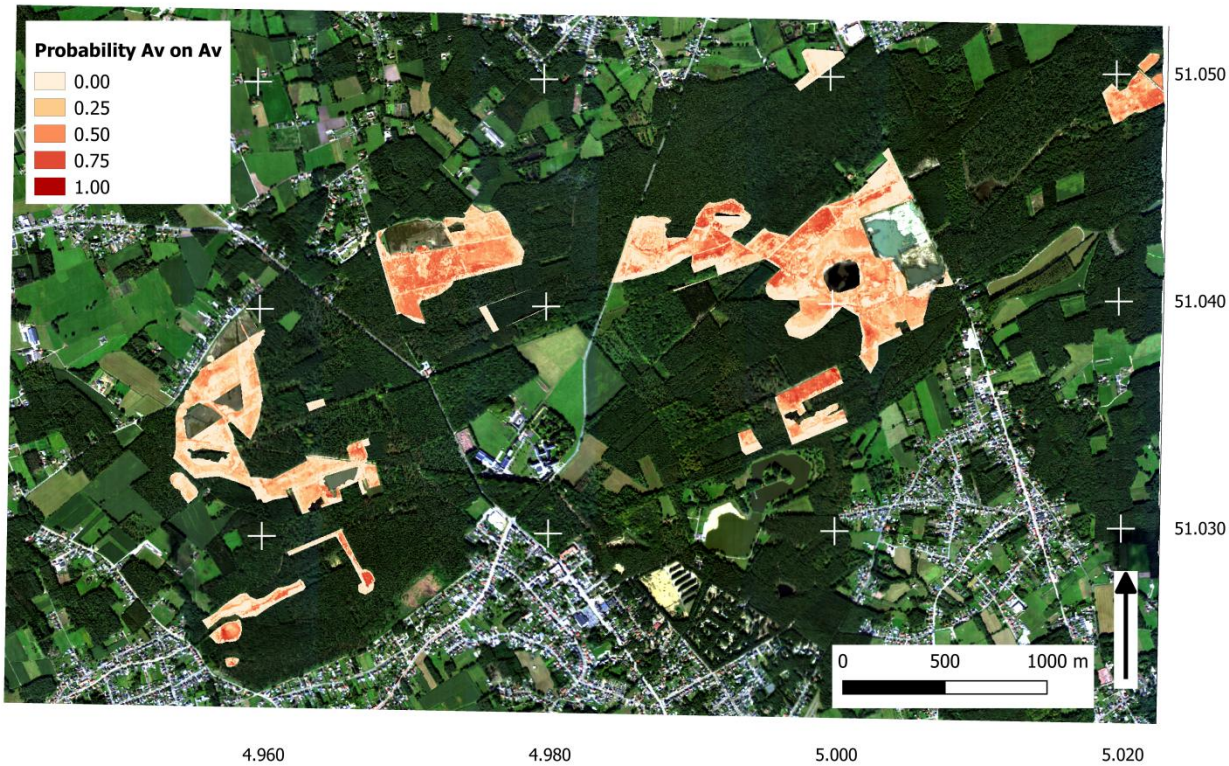



Figure 4: Predictions of the simple modelling (Step I) for Averbode showing the occurrence probability of *Campylopus introflexus*; see Supplement 1 for all predictions for Steps I, II and III.

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

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

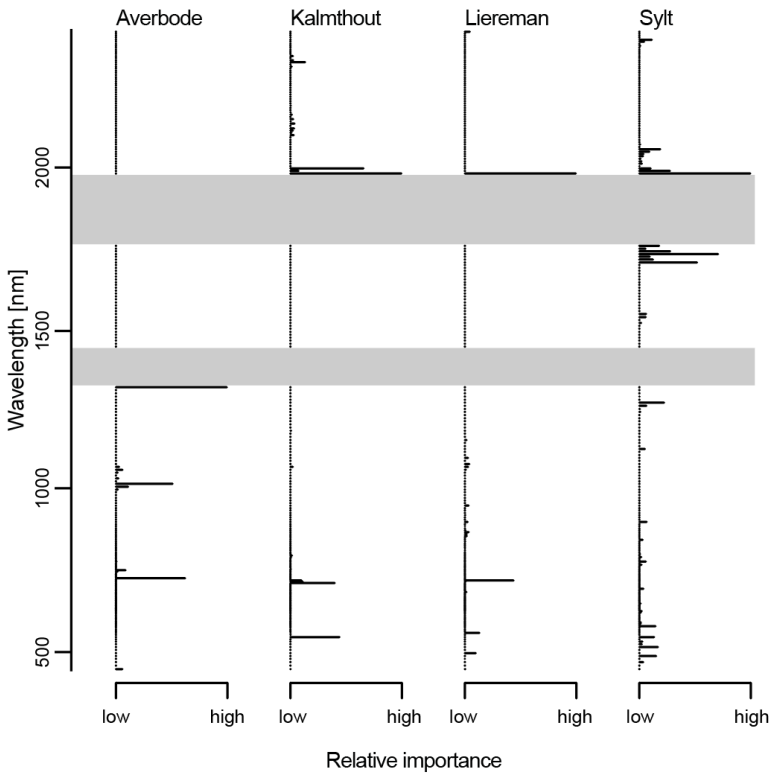



Figure 5: Relative band importance for the simple modelling (Step I) for each study site. Gray shaded areas indicate bands that were removed from the data set prior to the analyses.

### 3.3 Simple transfer

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

The visual evaluation confirmed that models with a $Tr_{AUC}$ around or above one displayed similar patterns and that maximum probabilities were within the same range as the respective original model for each area. The predictions of models calibrated for Sylt and Liereman were generally very similar compared to the 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 *ß* and *fc*, we observed the tendencies demonstrated in Figure 8. We found that *ß* 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

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

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

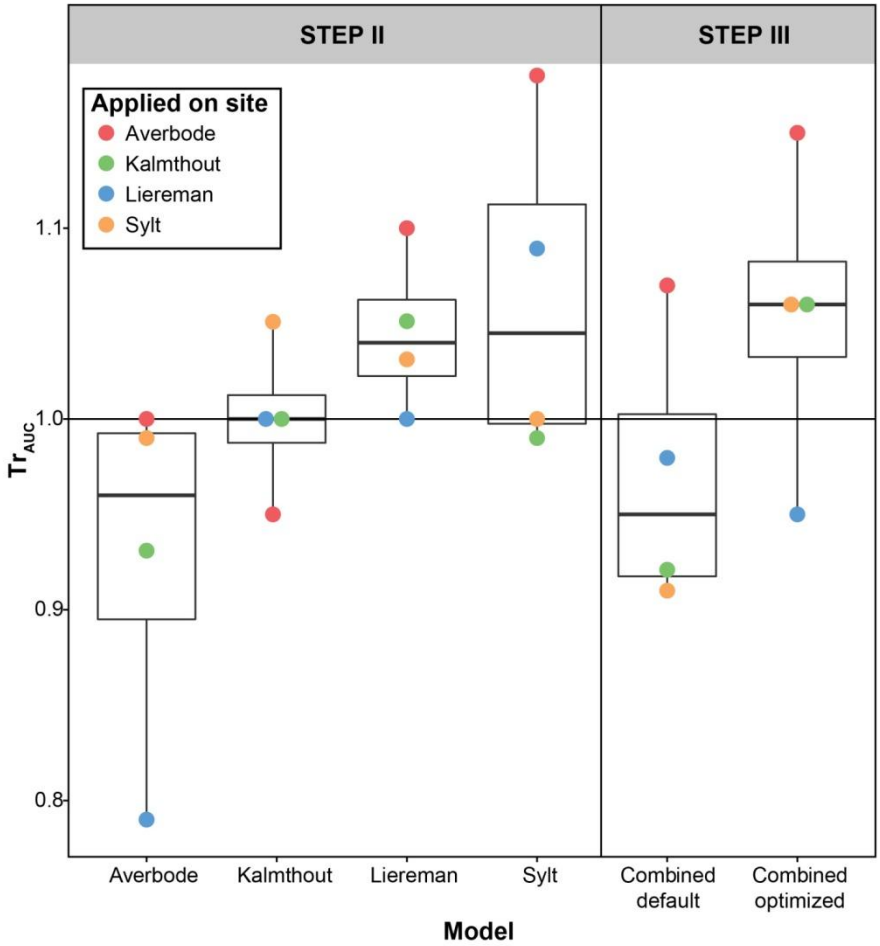



Figure 6: The transferability index ($Tr_{AUC}$) for the four different models for each study site applied on the respective three other study sites (STEP II) as well as the combined model (STEP III) using the default settings (combined default), and the combined model using the optimized parameter settings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

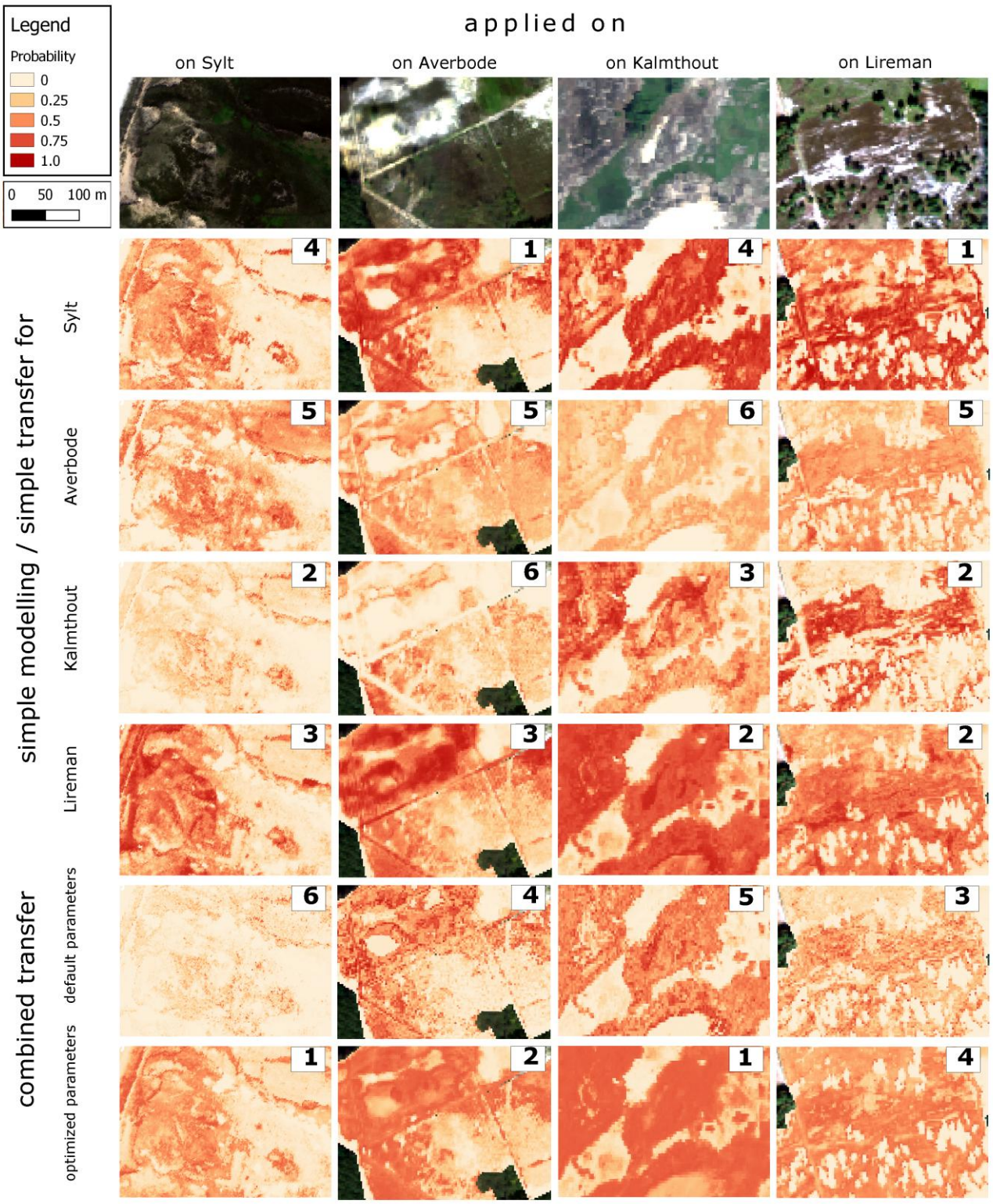



Figure 7: Model results for the four different study areas using three different approaches (Step I, II and III). The number in the right corner of each subset indicates the rank according to test AUC values, 1 being the highest and 6 the lowest rank. The test AUC indicates how well the model performs, while the probability

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

### 3.5 Comparison of simple modelling, simple transfer and combined transfer

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

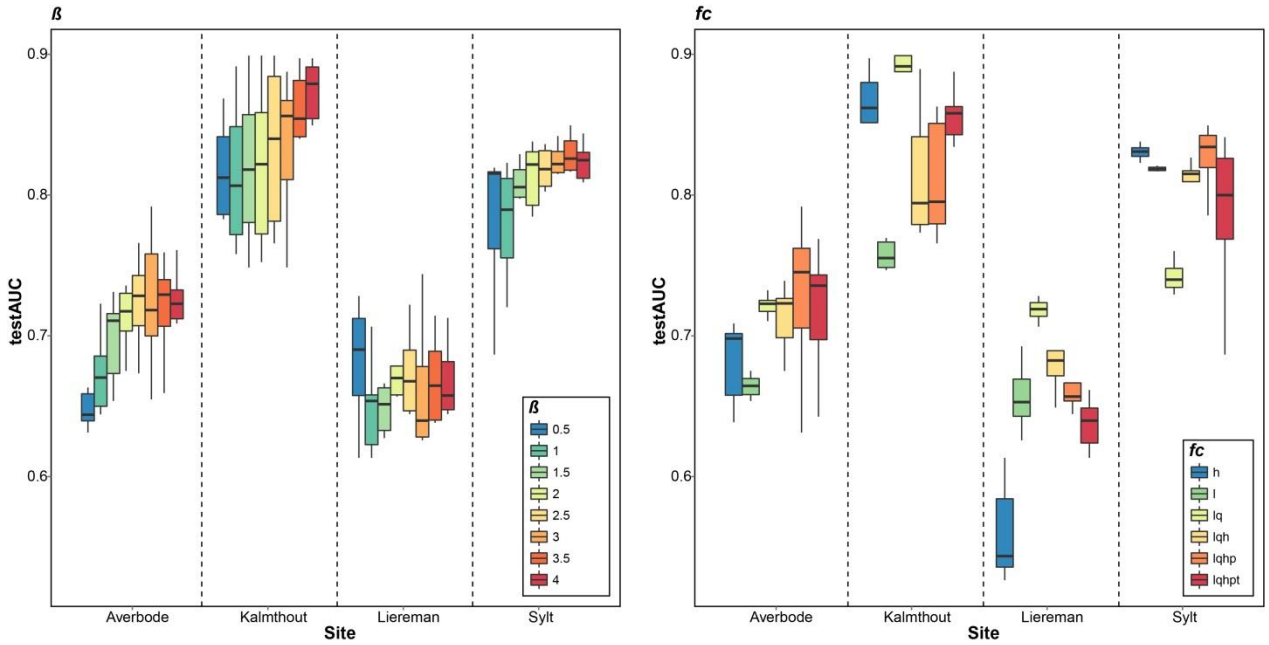



Figure 8: Effect of changing the Maxent parameter *ß* and feature class (*fc*) on the observed test AUC values for the combined models (Step III) created to map the invasive bryophyte *C. introflexus* in a dune habitat in four different study sites. Data from three study sites were combined to generate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

# 4 Discussion

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

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

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

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

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

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

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

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

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

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

### 4.3 Recommendations

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

Concerning the input data:

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

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

iii) One should always mitigate sampling bias.

Concerning analysis:

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

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

Concerning output evaluation:

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

### 4.4 Uncertainties, future research needs and potential applications

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

For all areas, there was a time lag of a few weeks to several months between the remote sensing acquisition and the fieldwork campaign, which may cause a slight under or overestimation in the species cover. This is especially true for Liereman and Kalmthout. However, 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).

Additional uncertainties may occur due to the different GPS devices that were used. While we used devices with differential correction for data collection in Liereman and Kalmthout, the devices used in Averbode and Kalmthout did not have this option, which may led to larger position uncertainties on those study sites in addition to the position uncertainties in the remote sensing data. Furthermore, our validation datasets, particularly for Kalmthout and Liereman, were relatively small. We chose to still work with these datasets as having to deal with a small amount of occurrences represents a real-world scenario for the (early) detection of invasive alien plant species, where informed decisions have to be made with a limited amount of data. However, collecting larger field datasets for validation might further enhance our understanding of the model performances and transferability success (Bean et al., 2012). Another important factor influencing the model are the soil reflectances, as some of the plots contain quite a high amount of bare soil. We did not separately assess the influence of the soil reflectance due to a lack of adequate data, but doing so could enhance the understanding of the different model performances. Finally, 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.

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

While we currently may not have very many situations where multiple imaging spectroscopy datasets are available for the same study species in different regions to build a combined model, this might change in the near future with the launch of hyperspectral satellite missions, such as EnMAP (Environmental Mapping and Analysis Program), where imaging spectroscopy data with a 30 m x 30 m resolution will be available worldwide. While this spatial resolution is certainly too coarse for mapping *C. introflexus*, it might be interesting for mapping larger species or vegetation types. 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. Fernandes et al., 2014; Robinson et al., 2016).

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

# 5 Conclusion

In this study we successfully transferred species distribution models for *Campylopus introflexus* which were calibrated at different sites using airborne imaging spectroscopy as explanatory variables. Our results demonstrate that model 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.

# 6 References


Anderson, R.P., Gonzalez, I., 2011. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. Ecol. Modell. 222, 2796–2811. https://doi.org/10.1016/j.ecolmodel.2011.04.011

Andrew, M., Ustin, S., 2008. The role of environmental context in mapping invasive plants with hyperspectral image data. Remote Sens. Environ. 112, 4301–4317. https://doi.org/10.1016/j.rse.2008.07.016

Barbosa, A.M., Real, R., Mario Vargas, J., 2009. Transferability of environmental favourability models in geographic space: The case of the Iberian desman (Galemys pyrenaicus) in Portugal and Spain. Ecol. Modell. 220, 747–754. https://doi.org/10.1016/j.ecolmodel.2008.12.004

Bean, W.T., Stafford, R., Brashares, J.S., 2012. The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. Ecography (Cop.). 35, 250–258. https://doi.org/10.1111/j.1600-0587.2011.06545.x

Berk, A., Anderson, G.P., Bernstein, L.S., Acharya, P.K., Dothe, H., Matthew, M.W., Adler-Golden, S.M., Chetwynd, Jr., J.H., Richtsmeier, S.C., Pukall, B., Allred, C.L., Jeong, L.S., Hoke, M.L., 1999. MODTRAN4 radiative transfer modeling for atmospheric correction, in: Larar, A.M. (Ed.), . p. 348. https://doi.org/10.1117/12.366388

Biermann, R., Daniels, F.J.A., 1997. Changes in a lichen-rich dry sand grassland vegetation with special reference to lichen synusiae and Campylopus introflexus. Phytocoenologia 27, 257–273. https://doi.org/10.1127/phyto/27/1997/257

Bivand, R., Keitt, T., Rowlingson, B., 2016. rgdal: Bindings for the Geospatial Data Abstraction Library [WWW Document]. URL https://cran.r-project.org/package=rgdal

Boxel, J.H., Jungerius, P.D., Kieffer, N., Hampele, N., 1997. Ecological effects of reactivation of artificially stabilized blowouts in coastal dunes. J. Coast. Conserv. 3, 57–62. https://doi.org/10.1007/BF02908179

Bradley, B. a., 2013. Remote detection of invasive plants: a review of spectral, textural and phenological approaches. Biol. Invasions 16, 1411–1425. https://doi.org/10.1007/s10530-013-0578-9

Cheng, Y.-B., 2007. Mapping an invasive species, kudzu (Pueraria montana), using hyperspectral imagery in western Georgia. J. Appl. Remote Sens. 1, 13514. https://doi.org/10.1117/1.2749266

DAISIE, 2015. DAISIE - 100 of the Worst [WWW Document]. URL http://www.europe-aliens.org/speciesTheWorst.do (accessed 6.11.15).

Duque-Lazo, J., van Gils, H., Groen, T.A., Navarro-Cerrillo, R.M., 2016. Transferability of species distribution models: The case of Phytophthora cinnamomi in Southwest Spain and Southwest Australia. Ecol. Modell. 320, 62–70. https://doi.org/10.1016/j.ecolmodel.2015.09.019

Elith, J., Leathwick, J.R., 2009. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. Annu. Rev. Ecol. Evol. Syst. 40, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. Divers. Distrib. 17, 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Essl, F., Steinbauer, K., Dullinger, S., Mang, T., Moser, D., 2014. Little, but increasing evidence of impacts by alien bryophytes. Biol. Invasions 16, 1175–1184. https://doi.org/10.1007/s10530-013-0572-2

Feilhauer, H., Somers, B., van der Linden, S., 2017. Optical trait indicators for remote sensing of plant species composition: Predictive power and seasonal variability. Ecol. Indic. 73, 825–833.

https://doi.org/10.1016/j.ecolind.2016.11.003
Fernandes, M.R., Aguiar, F.C., Silva, J.M.N., Ferreira, M.T., Pereira, J.M.C., 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. Appl. Earth Obs. Geoinf. 32, 79–91. https://doi.org/10.1016/j.jag.2014.03.026
He, K.S., Bradley, B.A., Cord, A.F., Rocchini, D., Tuanmu, M.-N., Schmidtlein, S., Turner, W., Wegmann, M., Pettorelli, N., 2015. Will remote sensing shape the next generation of species distribution models? Remote Sens. Ecol. Conserv. 1, 4–18. https://doi.org/10.1002/rse2.7
He, K.S., Rocchini, D., Neteler, M., Nagendra, H., 2011. Benefits of hyperspectral remote sensing for tracking plant invasions. Divers. Distrib. 17, 381–392. https://doi.org/10.1111/j.1472-4642.2011.00761.x
Heikkinen, R.K., Marmion, M., Luoto, M., 2012. Does the interpolation accuracy of species distribution models come at the expense of transferability? Ecography (Cop.). 35, 276–288. https://doi.org/10.1111/j.1600-0587.2011.06999.x
Hijmans, R.J., 2016. raster: Geographic Data Analysis and Modeling [WWW Document]. URL https://cran.r-project.org/package=raster
Hijmans, R.J., 2012. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. Ecology 93, 679–688. https://doi.org/10.1890/11-0826.1
Hijmans, R.J., Elith, J., 2015. Species distribution modeling with R. R CRAN Proj. 79 pp.
Hijmans, R.J., Phillips, S., Leathwick, J., Elith, J., 2016. dismo: Species Distribution Modeling [WWW Document]. URL https://cran.r-project.org/package=dismo
Huang, C., Asner, G.P., 2009. Applications of Remote Sensing to Alien Invasive Plant Studies. Sensors 9, 4869–4889. https://doi.org/10.3390/s90604869
Instituut voor Natuurbehoud, 2016. BIOLOGISCHE WAARDERINGSKAART [WWW Document]. URL http://www.geopunt.be/download?container=bwk2&title=Biologische waarderingskaart - Natura 2000 Habitatkaart (accessed 2.2.18).
Jiménez-Valverde, A., Peterson, A.T., Soberón, J., Overton, J.M., Aragón, P., Lobo, J.M., 2011. Use of niche models in invasive species risk assessments. Biol. Invasions 13, 2785–2797. https://doi.org/10.1007/s10530-011-9963-4
Kempeneers, P., 2016. pktools 2.6.7 Processing Kernel for geospatial data [WWW Document]. URL http://pktools.nongnu.org (accessed 2.2.18).
Ketner-Oostra, R., Sykora, K. V., 2000. Vegetation succession and lichen diversity on dry coastal calcium- poor dunes and the impact of management experiments. J. Coast. Conserv. 6, 191–206. https://doi.org/10.1007/BF02913815
Ketner-Oostra, R., Sýkora, K. V., 2004. Decline of lichen-diversity in calcium-poor coastal dune vegetation since the 1970s, related to grass and moss encroachment. Phytocoenologia 34, 521–549. https://doi.org/10.1127/0340-269X/2004/0034-0521
LEGUAN, 2012. Kartierung der Salzwiesen und Dünen an der Westküste von Schleswig-Holstein 2011-2012 – Biotopkartierung Sylt.
Mateo, R.G., Broennimann, O., Petitpierre, B., Muñoz, J., van Rooy, J., Laenen, B., Guisan, A., Vanderpoorten, A., 2015. What is the potential of spread in invasive bryophytes? Ecography (Cop.). 38, 480–487. https://doi.org/10.1111/ecog.01014
Merow, C., Smith, M.J., Silander, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. Ecography (Cop.). 36, 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x
Michez, A., Piégay, H., Jonathan, L., Claessens, H., Lejeune, P., 2016. Mapping of riparian invasive 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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Figure 1
Click here to download high resolution image

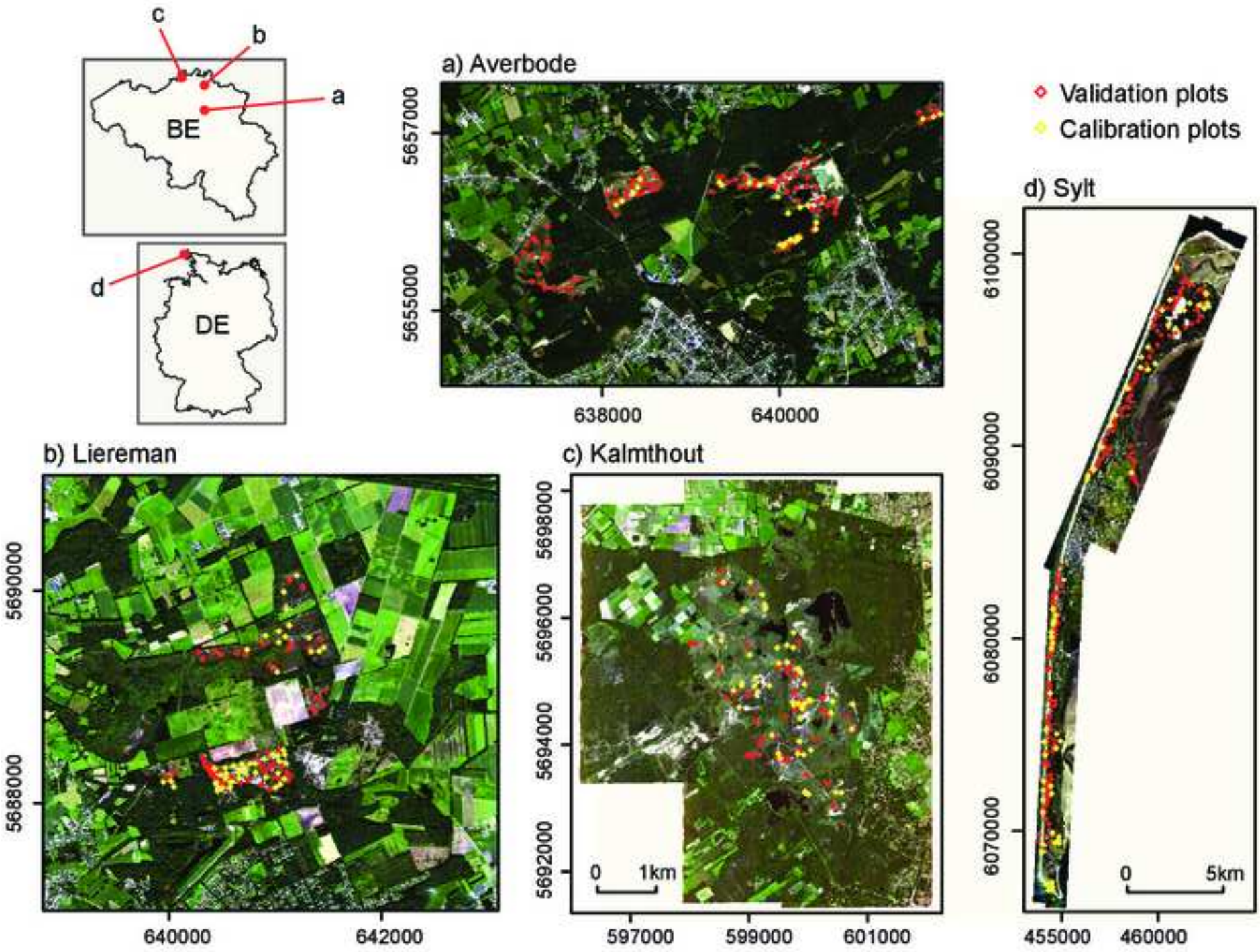


Figure 2

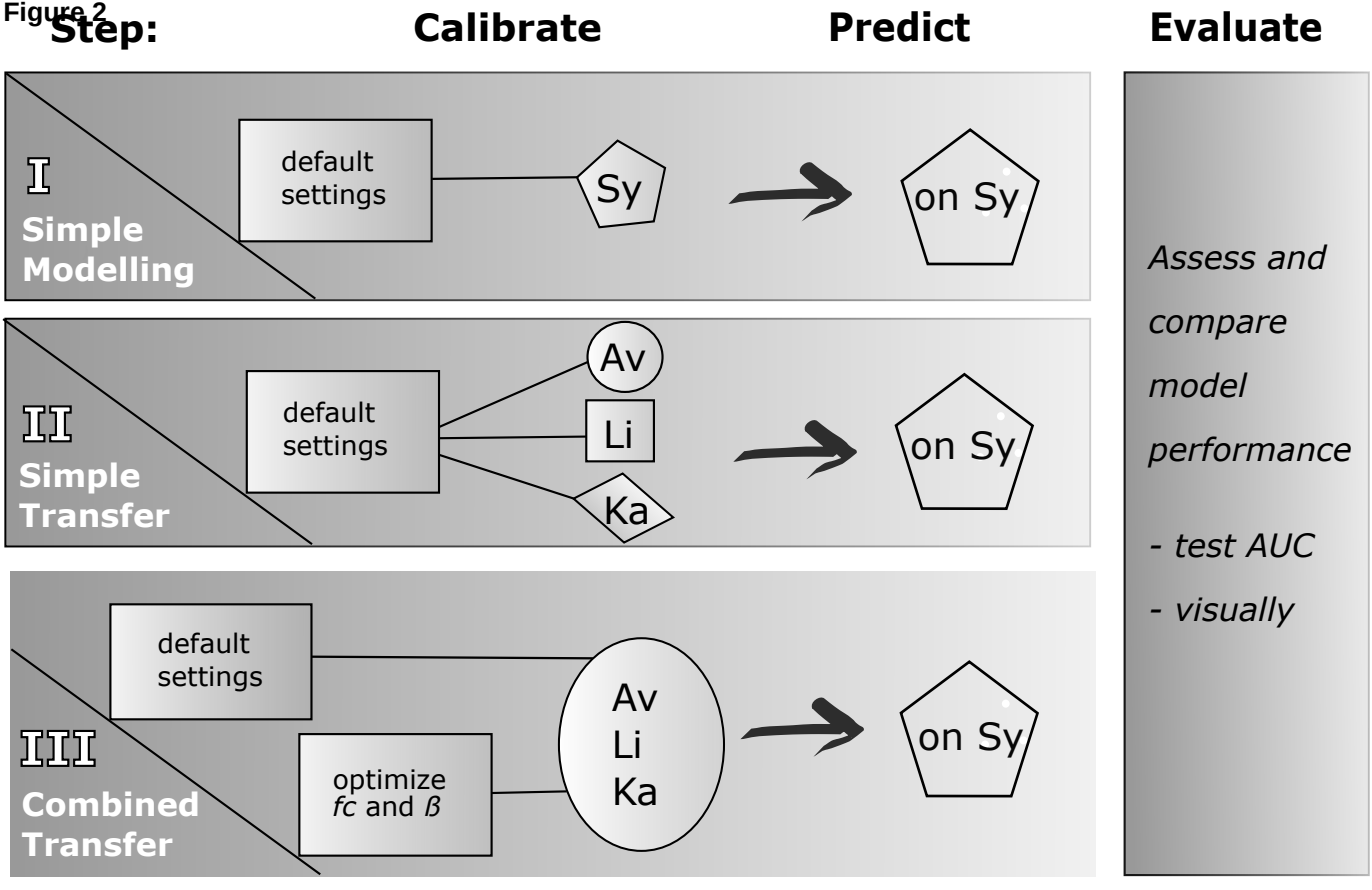

Step:
Calibrate
Predict
Evaluate
I
Simple Modelling
default settings
Sy
on Sy
II
Simple Transfer
default settings
Av
Li
Ka
on Sy
III
Combined Transfer
default settings
optimize fc and ß
Av
Li
Ka
on Sy
Assess and compare model performance
- test AUC
- visually

Figure 3

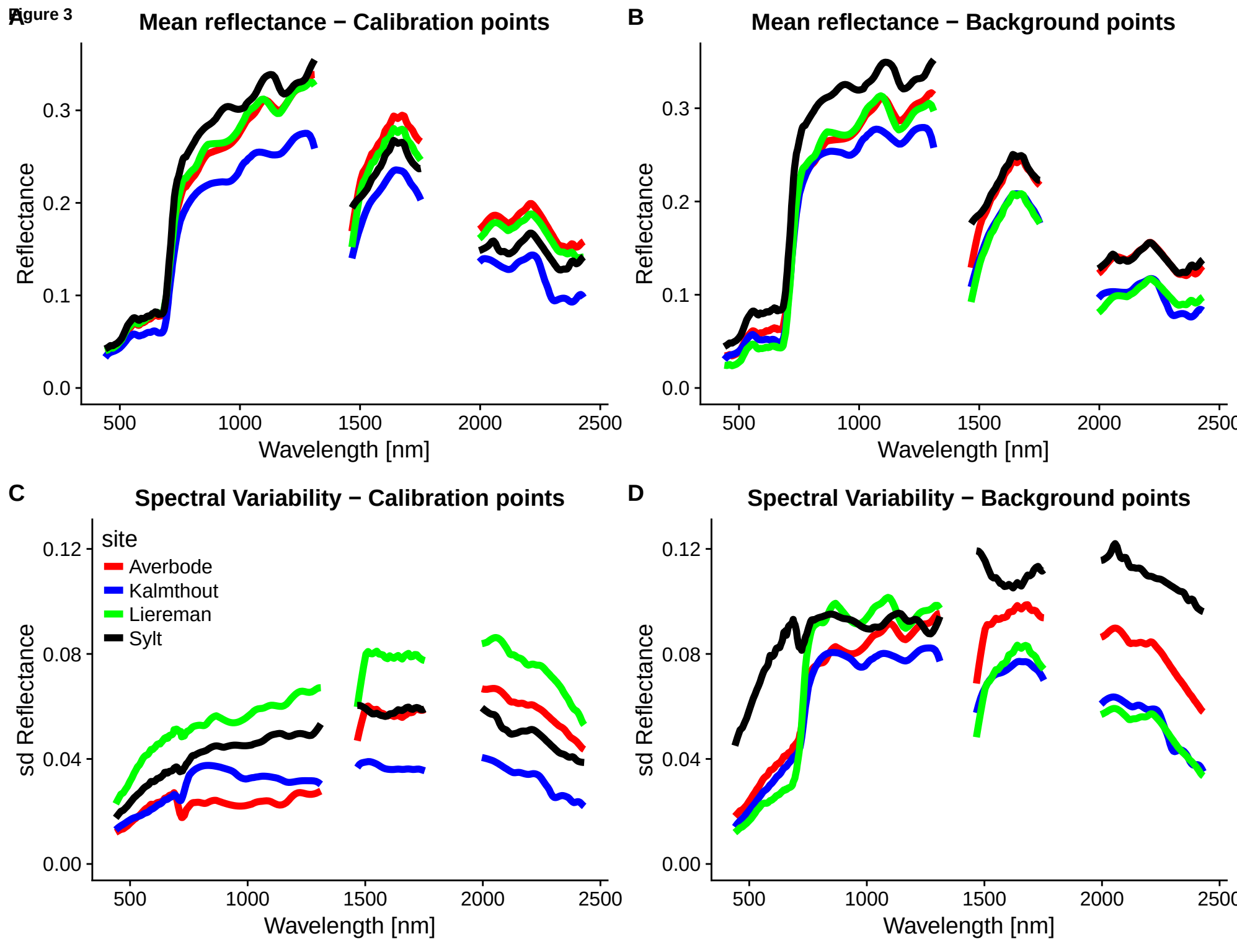
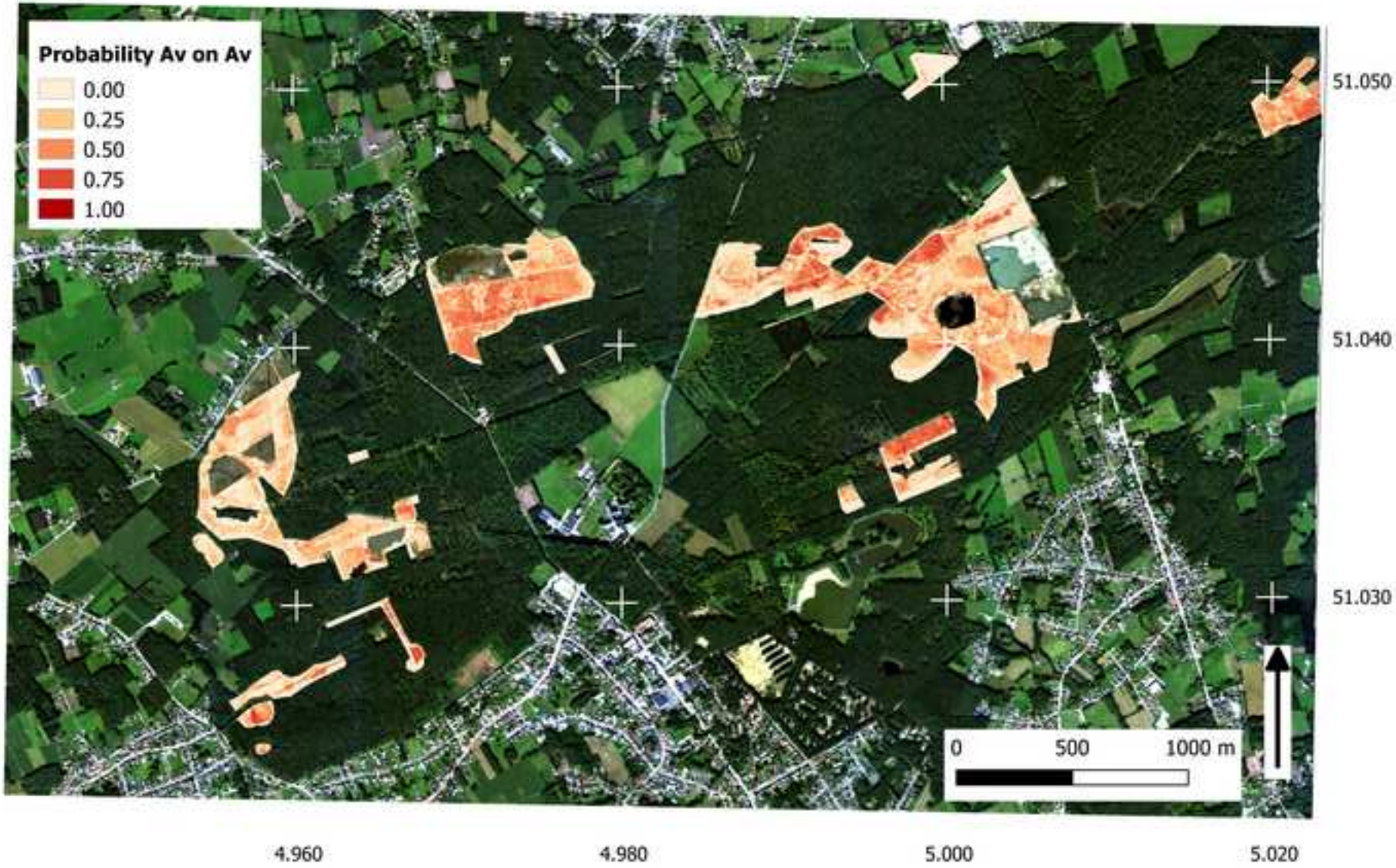

A
Mean reflectance − Calibration points
Reflectance
0.0
0.1
0.2
0.3
500
1000
1500
2000
2500
Wavelength [nm]
B
Mean reflectance − Background points
Reflectance
0.0
0.1
0.2
0.3
500
1000
1500
2000
2500
Wavelength [nm]
C
Spectral Variability − Calibration points
sd Reflectance
0.00
0.04
0.08
0.12
site
Averbode
Kalmthout
Liereman
Sylt
500
1000
1500
2000
2500
Wavelength [nm]
D
Spectral Variability − Background points
sd Reflectance
0.00
0.04
0.08
0.12
500
1000
1500
2000
2500
Wavelength [nm]

**Figure 4**
**Click here to download high resolution image**



# Figure 5

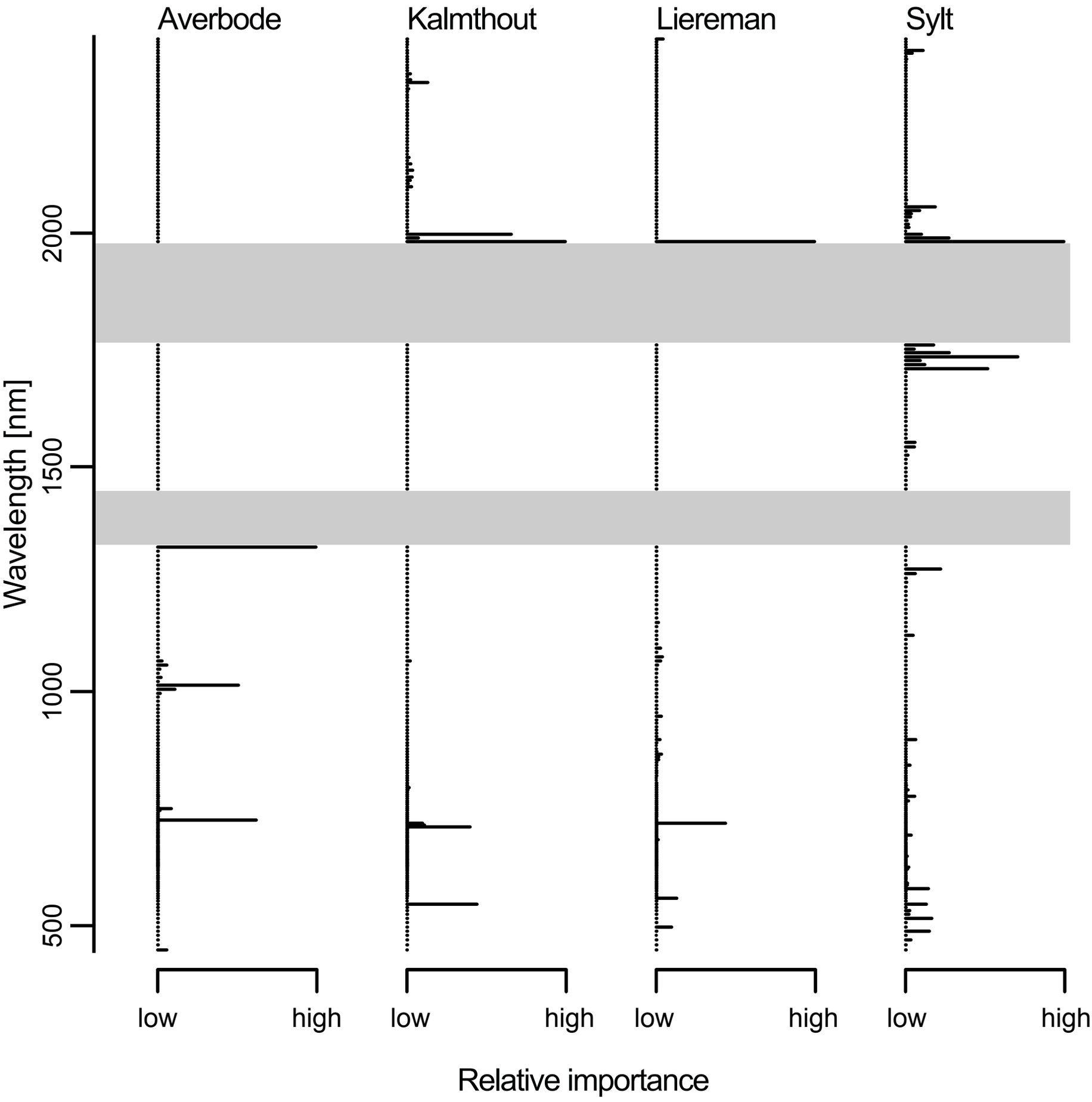

Averbode
Kalmthout
Liereman
Sylt
Wavelength [nm]
2000
1500
1000
500
low
high
low
high
low
high
low
high
Relative importance

Figure 6

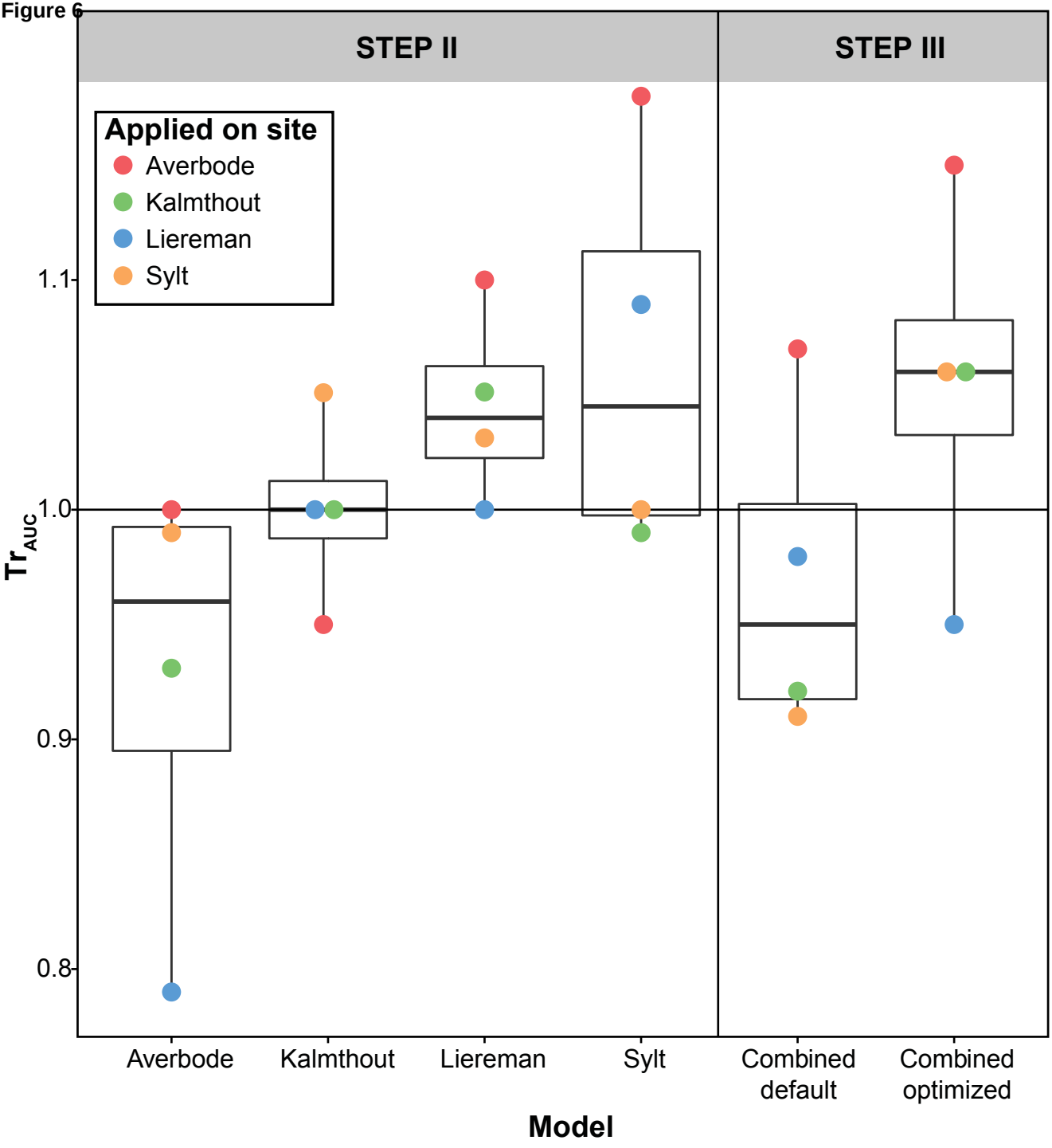

STEP II
STEP III
Applied on site
Averbode
Kalmthout
Liereman
Sylt
1.1
1.0
0.9
0.8
$Tr_{AUC}$
Averbode
Kalmthout
Liereman
Sylt
Combined default
Combined optimized
Model

**Figure 7**
**Click here to download high resolution image**

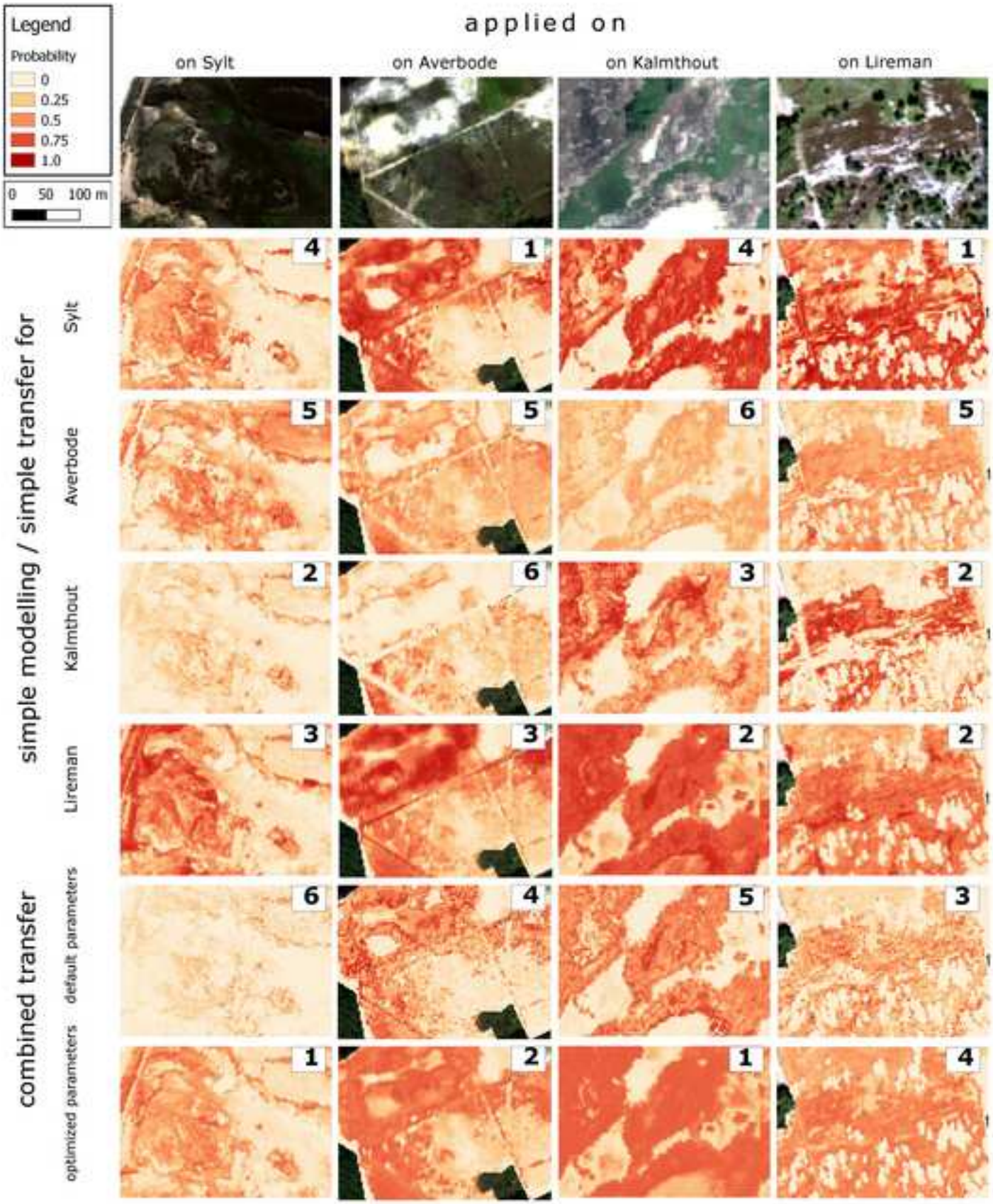
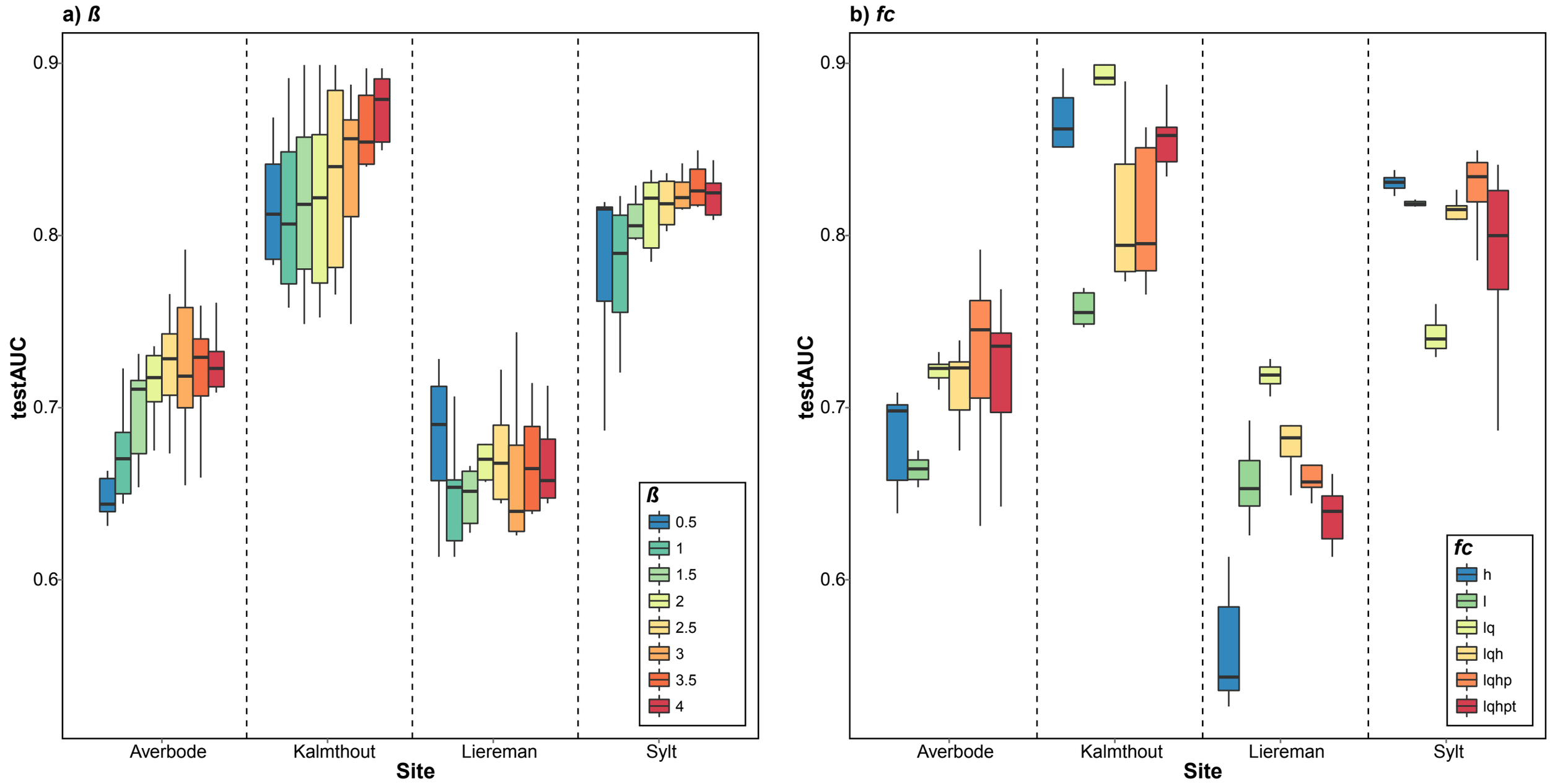


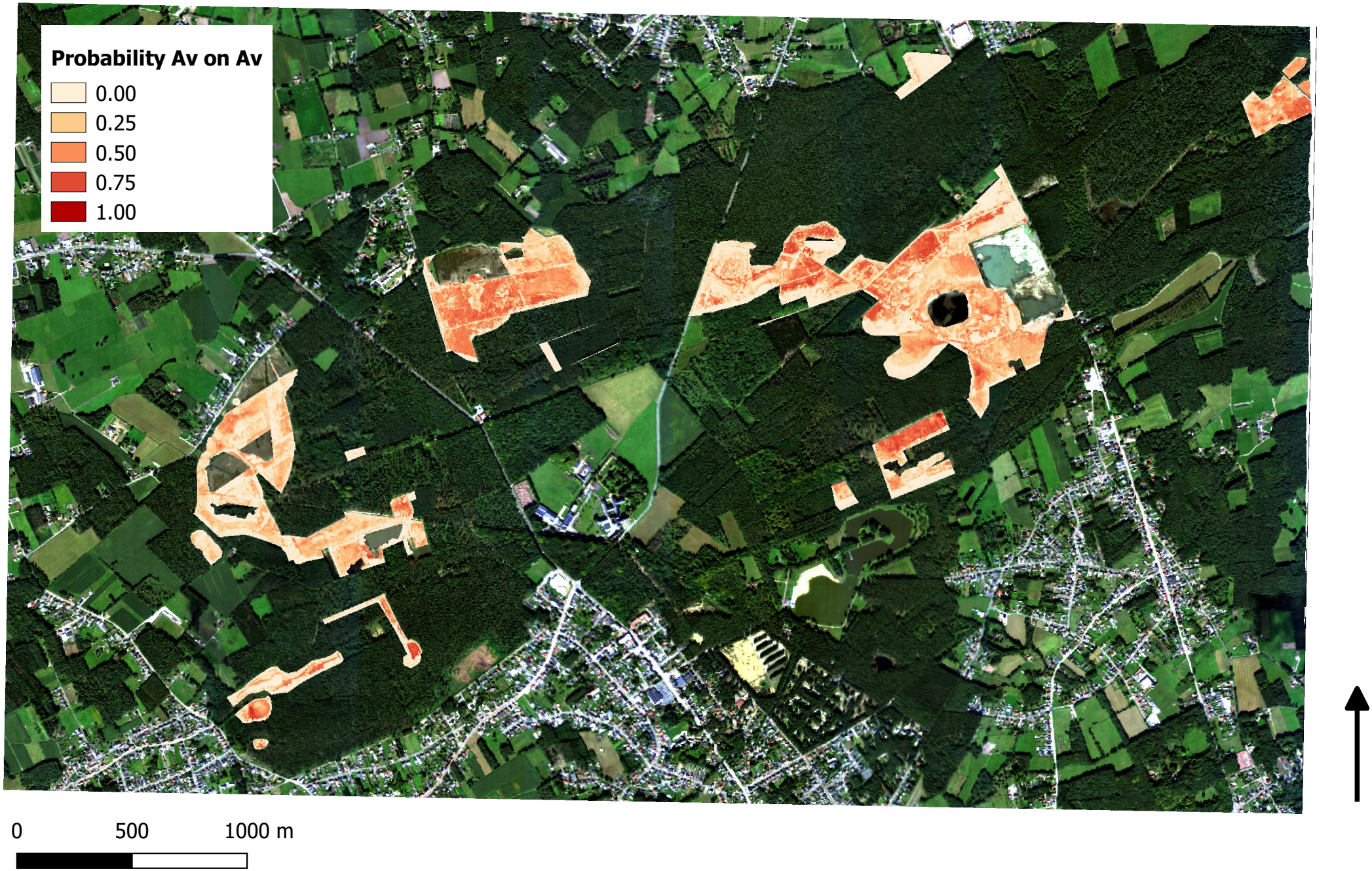
Figure 8

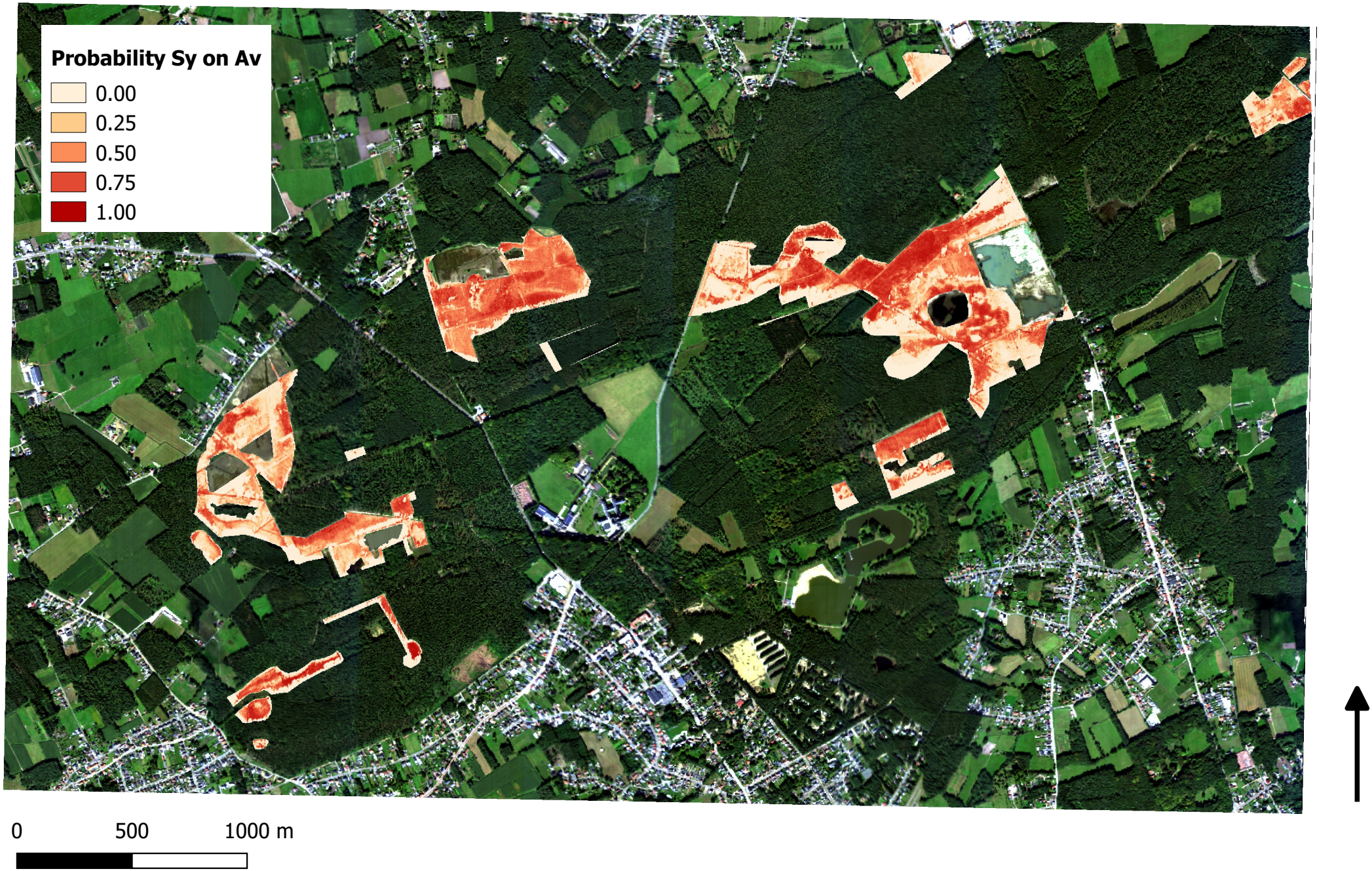

a) ß
b) fc
testAUC
0.9
0.8
0.7
0.6
Averbode
Kalmthout
Liereman
Sylt
Site
ß
0.5
1
1.5
2
2.5
3
3.5
4
fc
h
l
lq
lqh
lqhp
lqhpt

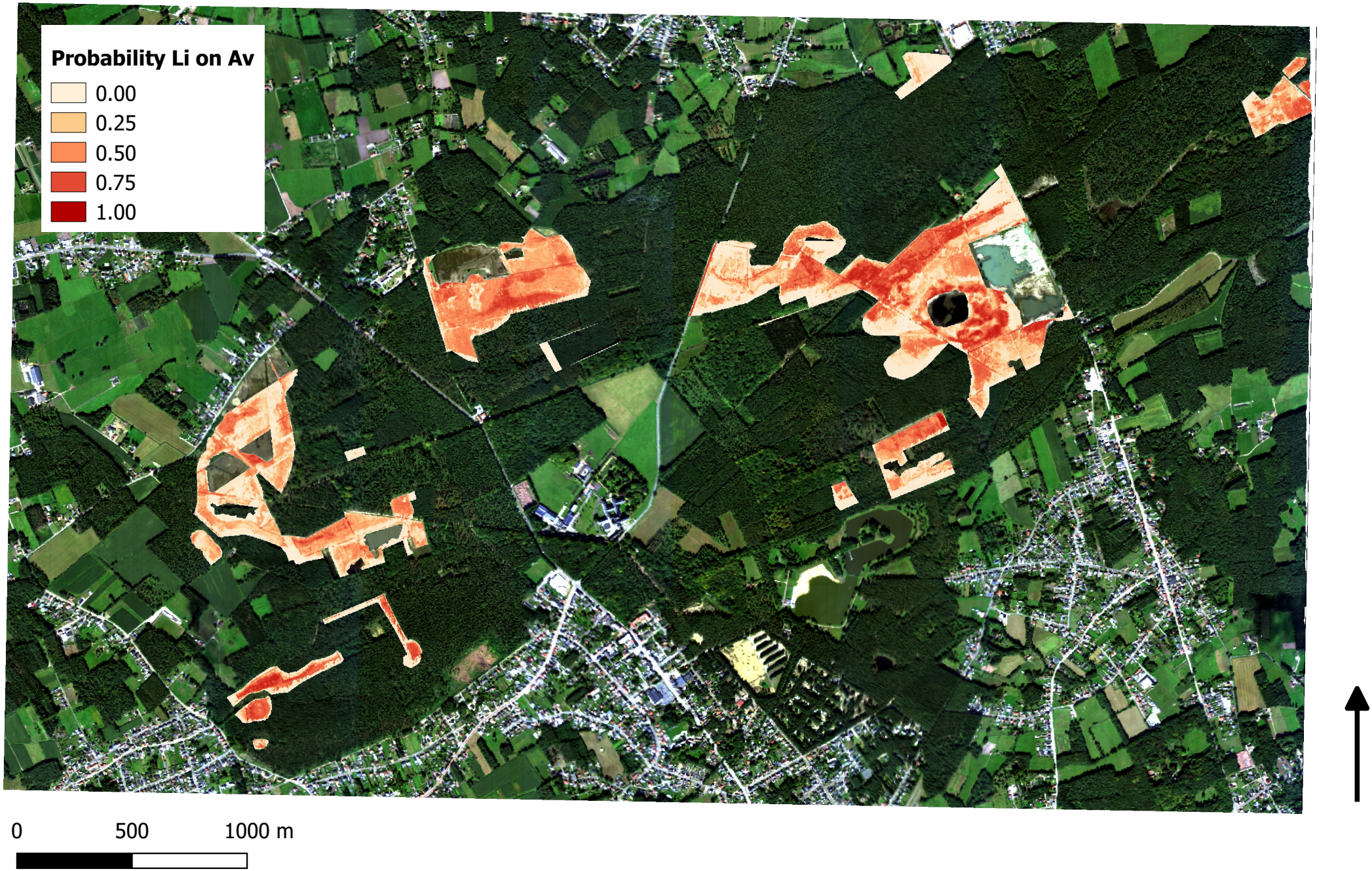
# Supplement 1

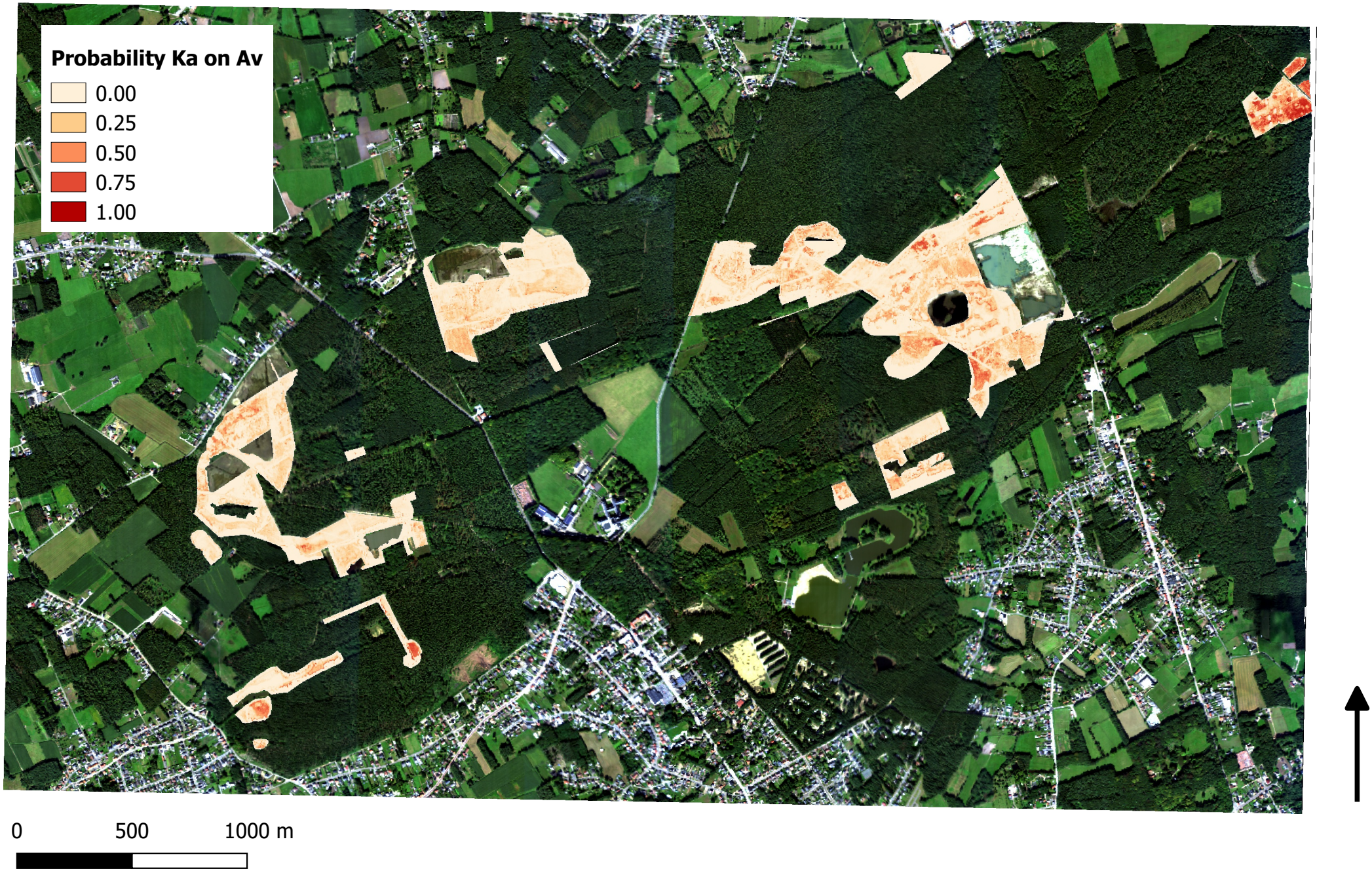
Occurrence probability of *Campylopus introflexus* according to the model predictions made in Step I (simple modelling), Step II (simple transfer) and Step III (combined transfer)

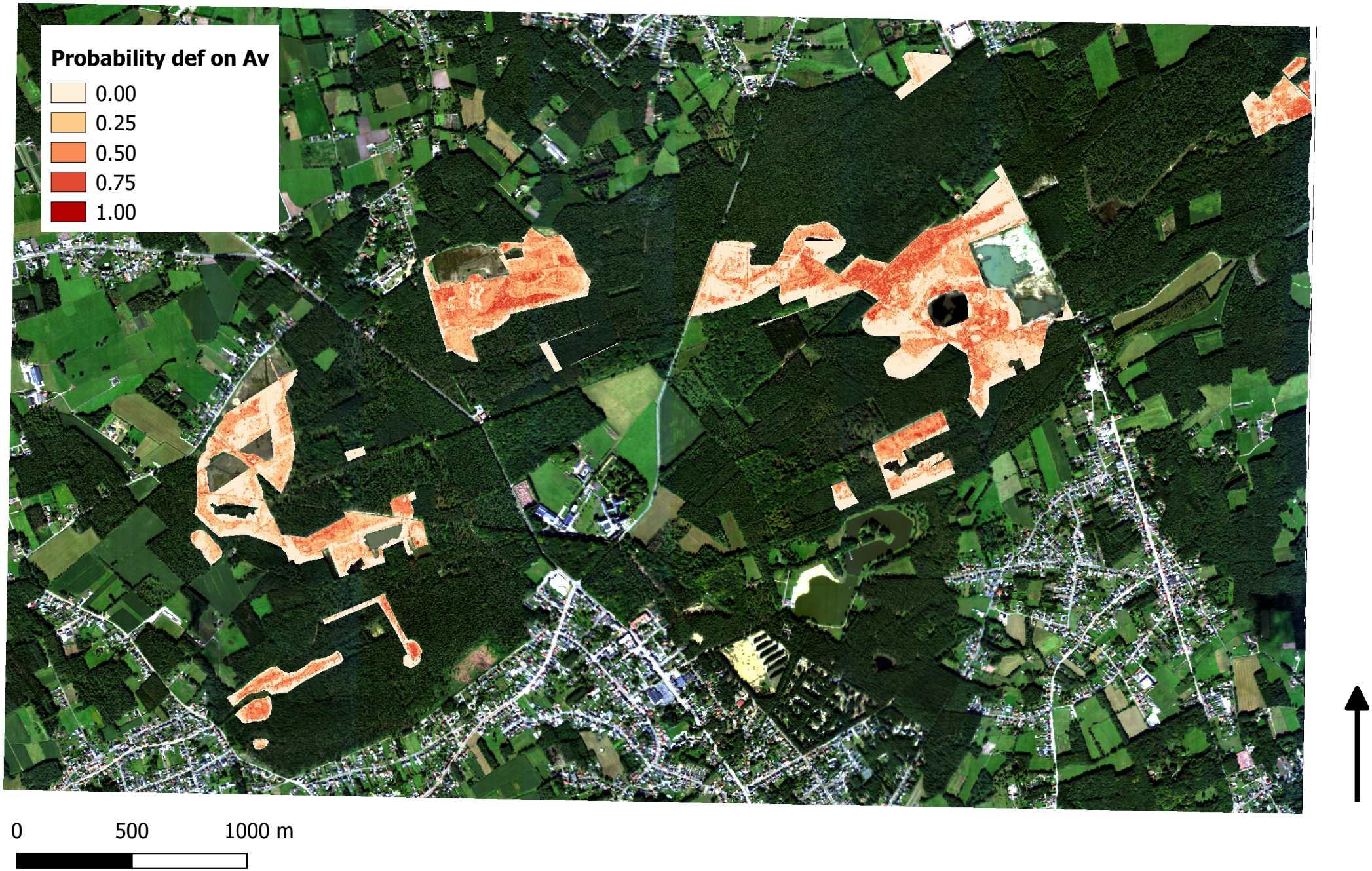

Averbode
0
500
1000 m

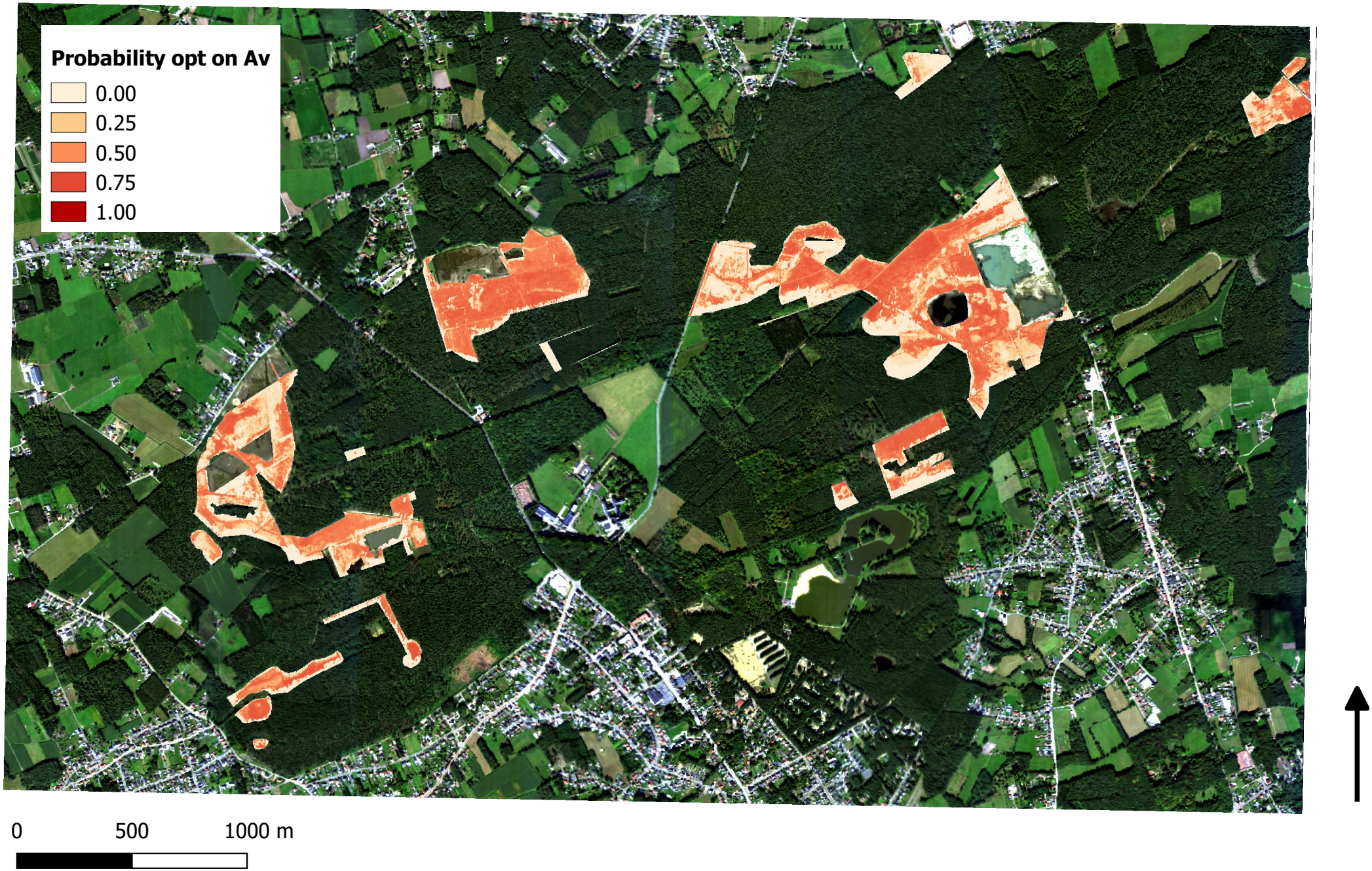

Probability Av on Av
0.00
0.25
0.50
0.75
1.00
0
500
1000 m


Probability Sy on Av
0.00
0.25
0.50
0.75
1.00
0
500
1000 m


Probability Li on Av
0.00
0.25
0.50
0.75
1.00
0
500
1000 m

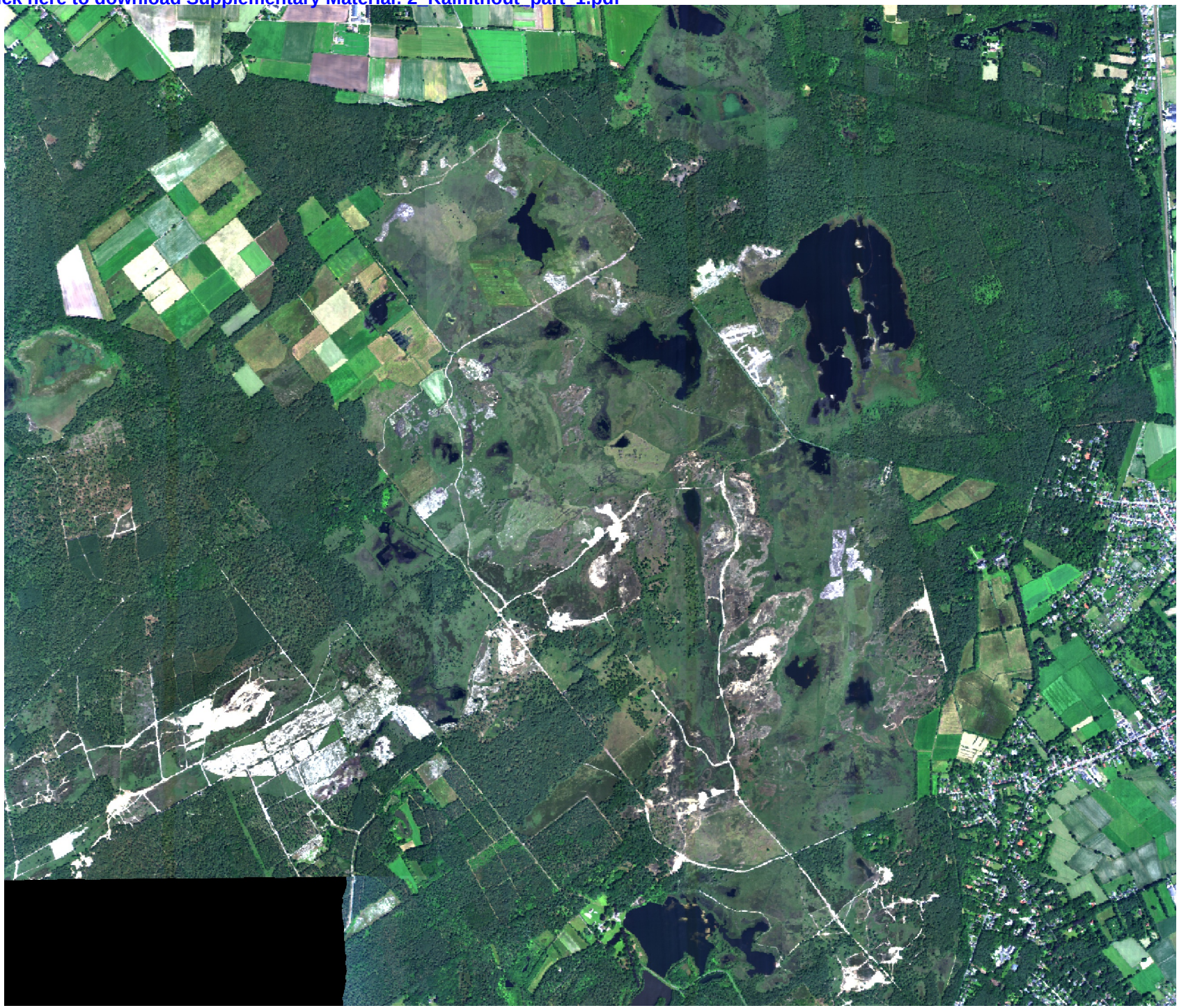

Probability Ka on Av
0.00
0.25
0.50
0.75
1.00
0
500
1000 m


Probability def on Av
0.00
0.25
0.50
0.75
1.00
0
500
1000 m


Probability opt on Av
0.00
0.25
0.50
0.75
1.00
0
500
1000 m

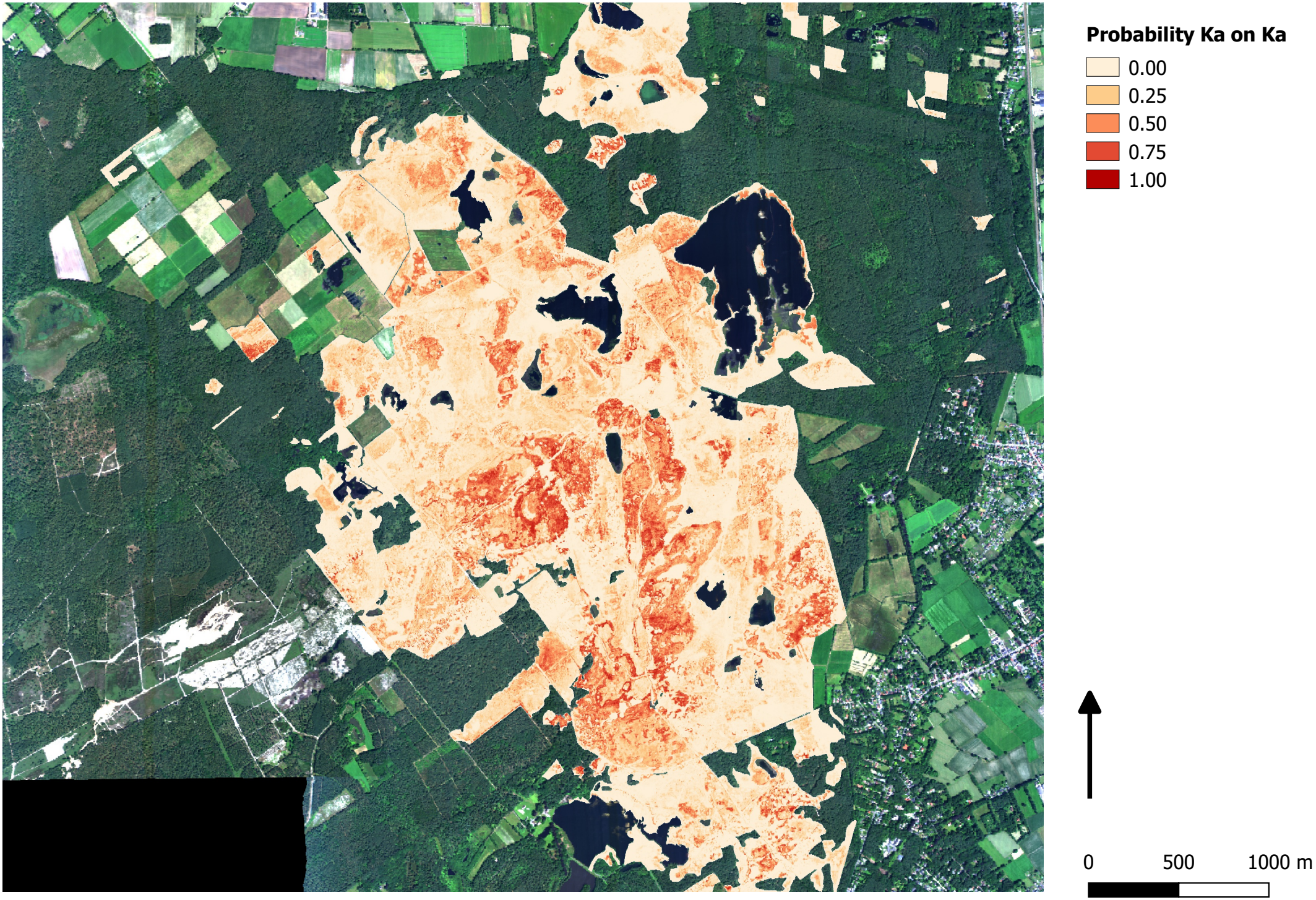
**Supplementary Material 1**

Click here to download Supplementary Material: 2_Kalmthout_part_1.pdf



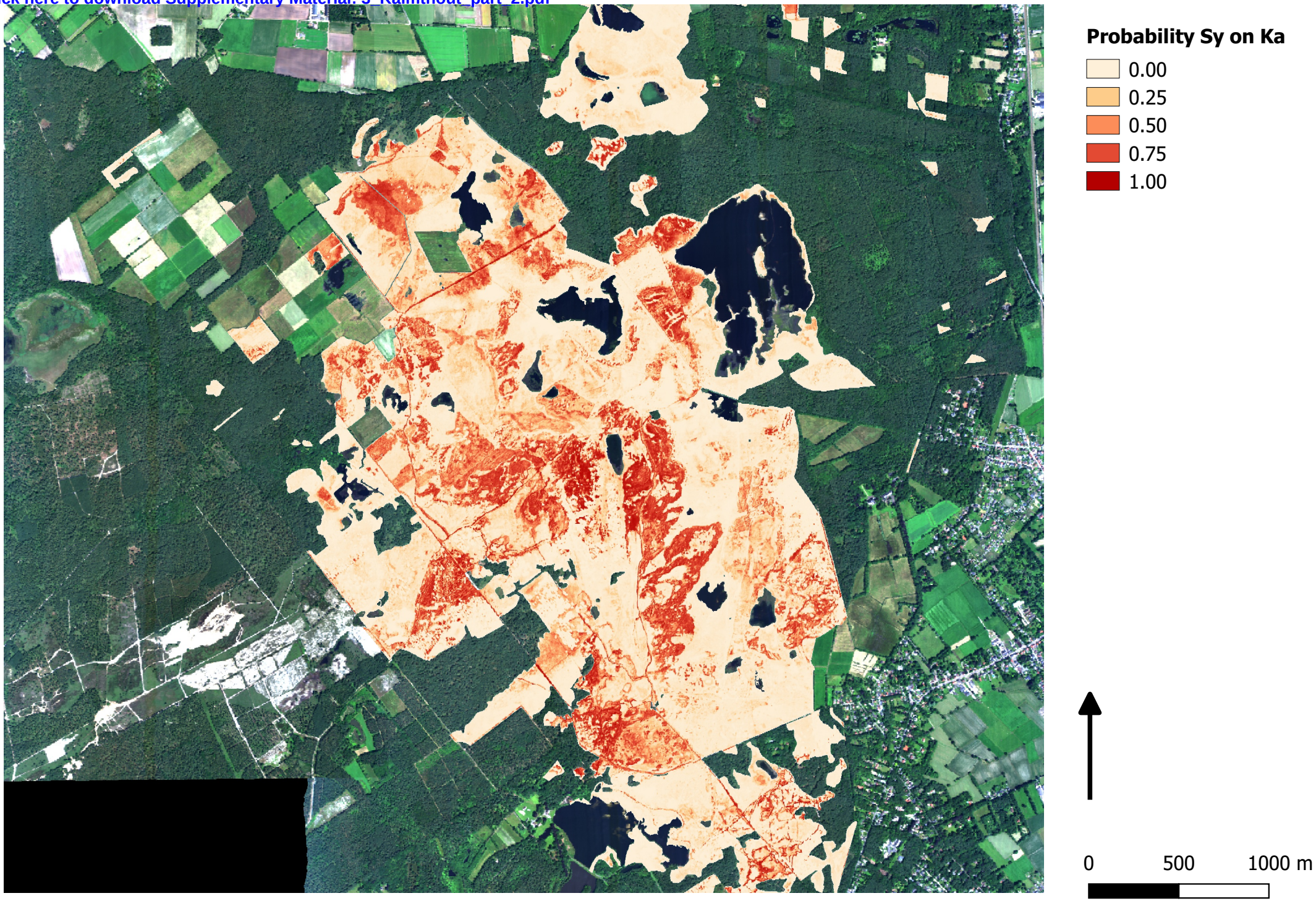
**Kalmthout**

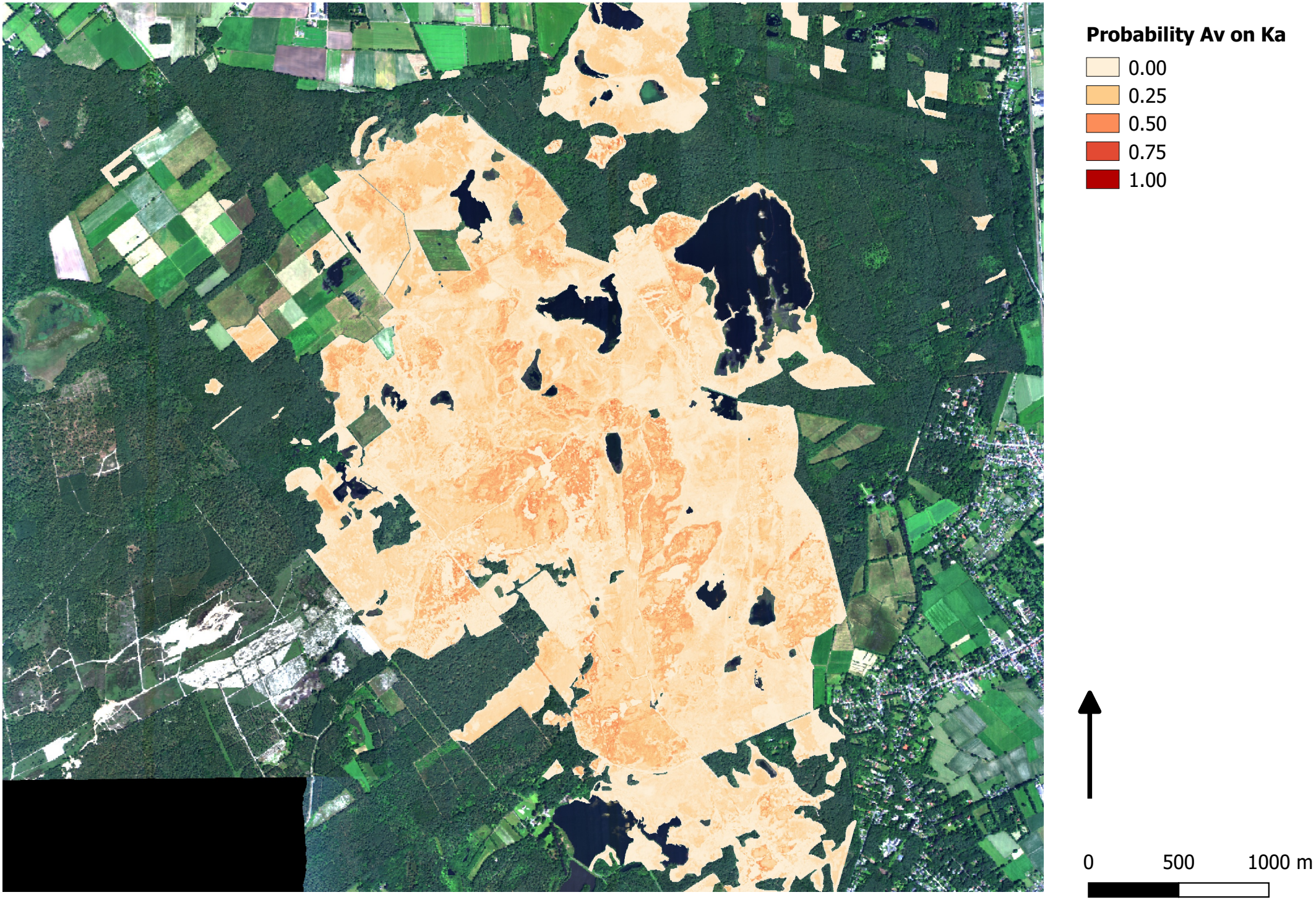
0 500 1000 m

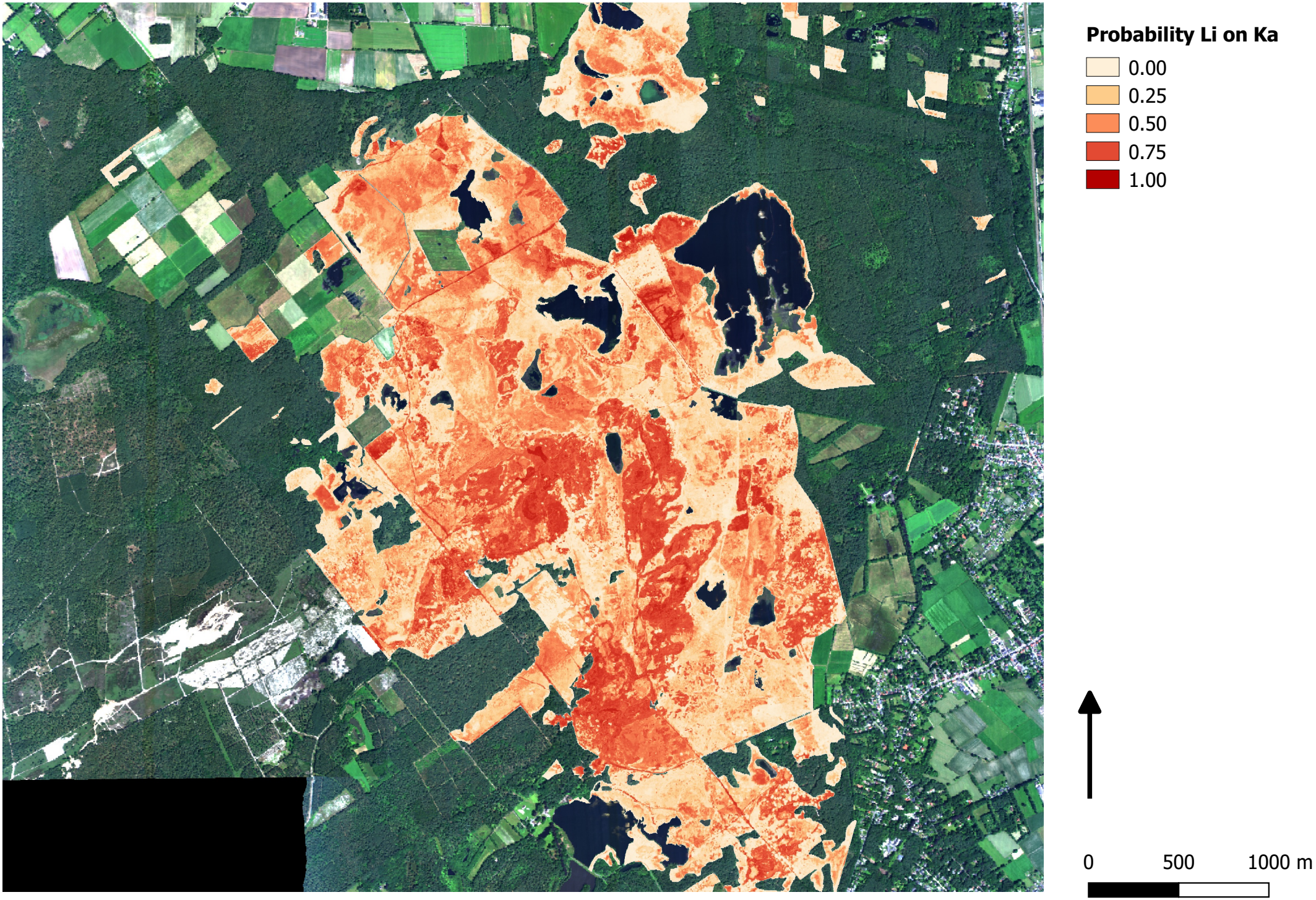

Probability Ka on Ka
0.00
0.25
0.50
0.75
1.00
0
500
1000 m

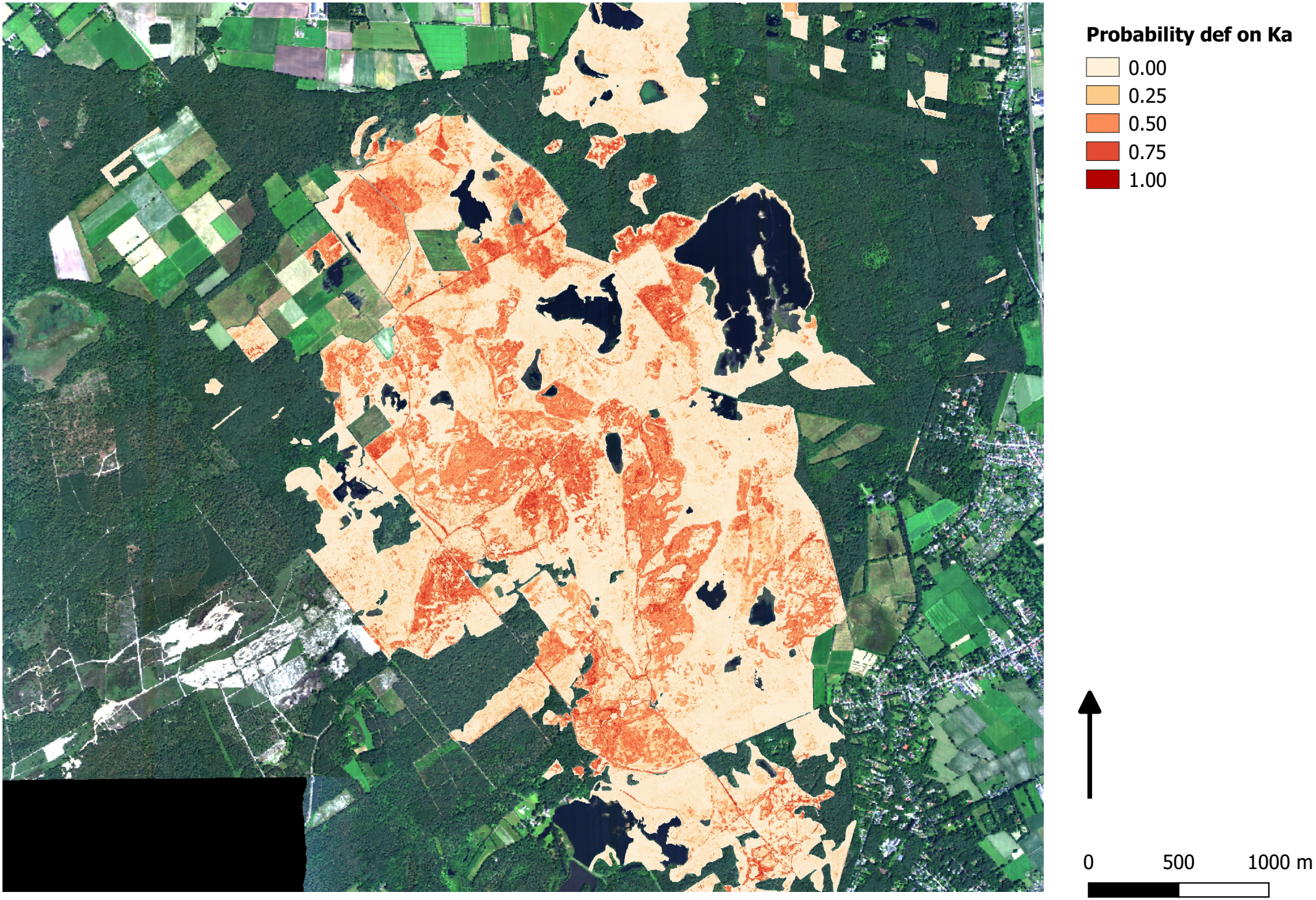
**Supplementary Material 1**

Click here to download Supplementary Material: 3_Kalmthout_part_2.pdf



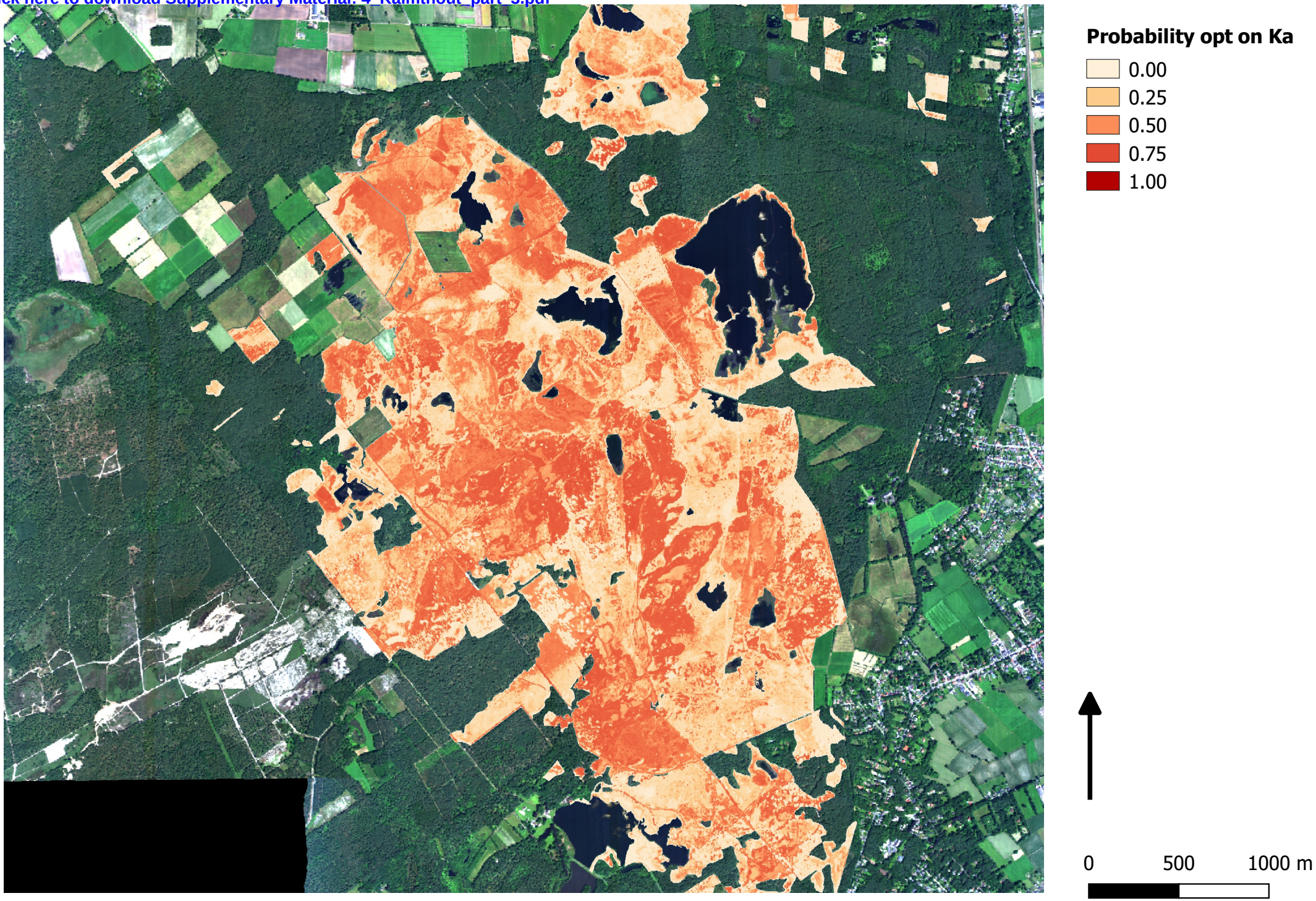

Probability Av on Ka
0.00
0.25
0.50
0.75
1.00
0
500
1000 m


Probability Li on Ka
0.00
0.25
0.50
0.75
1.00
0
500
1000 m


Probability def on Ka
0.00
0.25
0.50
0.75
1.00
0
500
1000 m

Supplementary Material 1

Click here to download Supplementary Material: 4_Kalmthout_part_3.pdf



**Supplementary Material 1**

**Click here to download Supplementary Material: 5_Liereman.pdf**


Probability Li on Li
0.00
0.25
0.50
0.75
1.00
0
250
500 m


Probability Sy on Li
0.00
0.25
0.50
0.75
1.00
0
250
500 m


Probability Av on Li
0.00
0.25
0.50
0.75
1.00
0
250
500 m


Probability Ka on Li
0.00
0.25
0.50
0.75
1.00
0
250
500 m


Probability def on Li
0.00
0.25
0.50
0.75
1.00
0
250
500 m

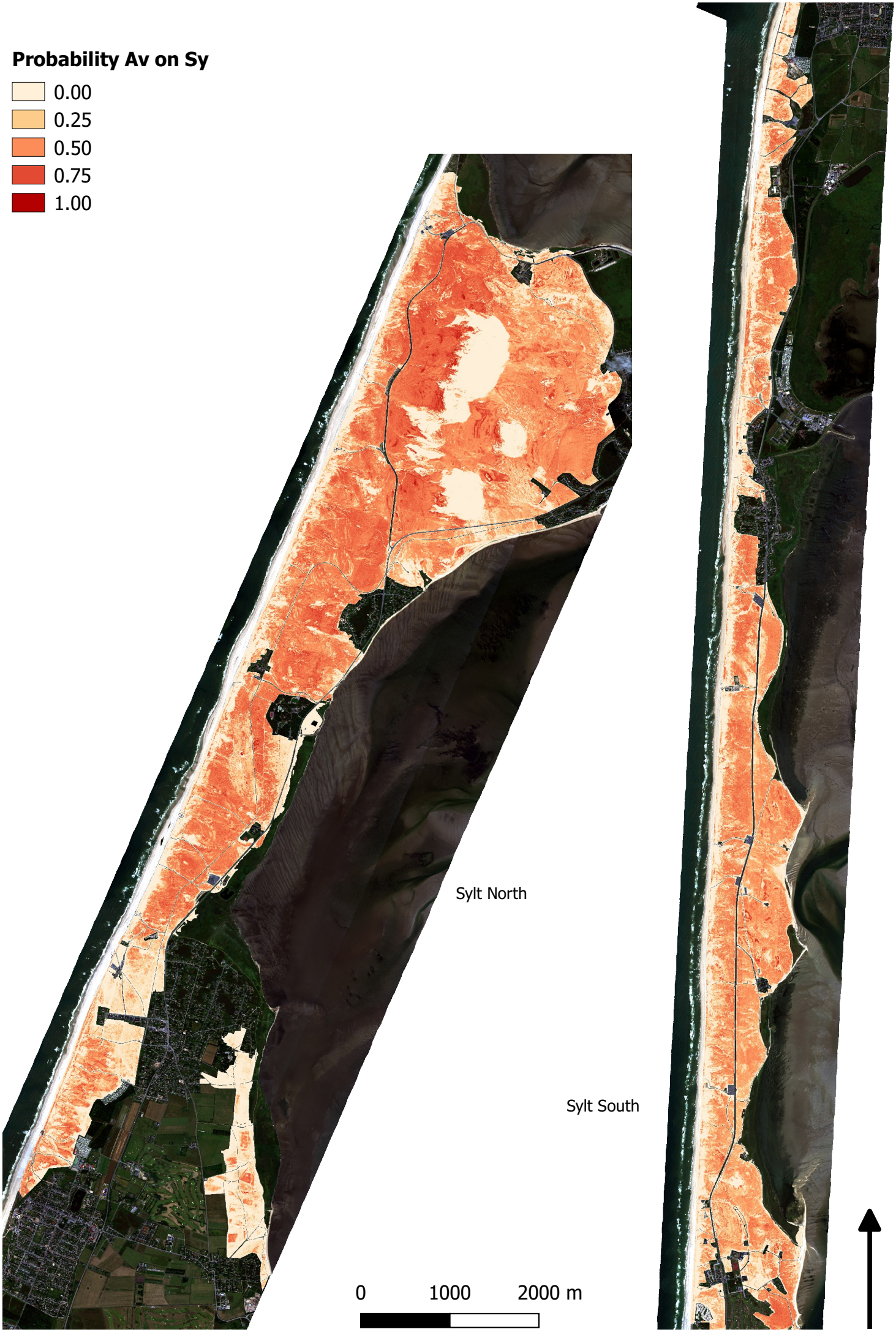

Probability opt on Li
0.00
0.25
0.50
0.75
1.00
0
250
500 m

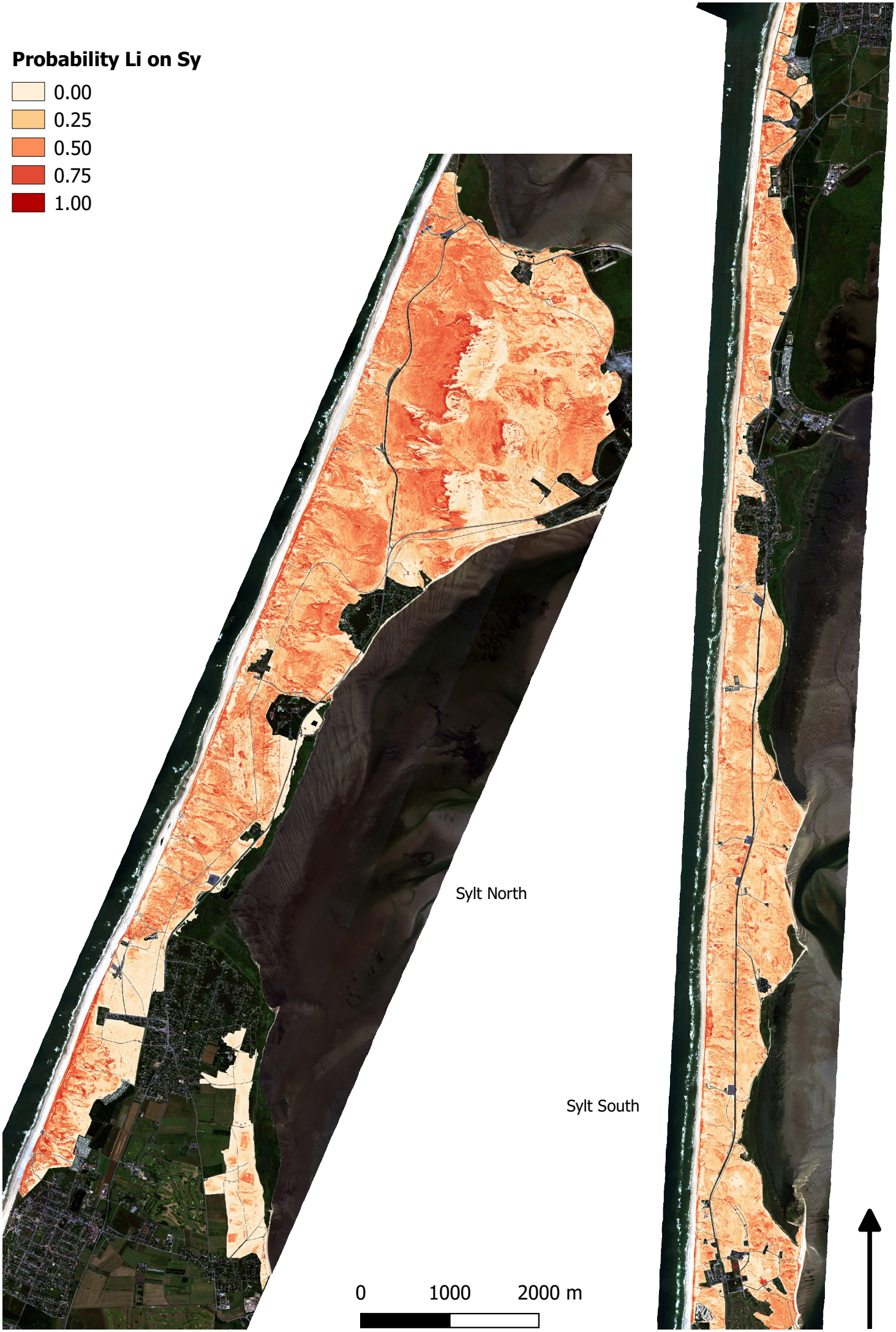
Supplementary Material 1

Click here to download Supplementary Material: 6_Sylt.pdf

## Sylt


Probability Sy on Sy
0.00
0.25
0.50
0.75
1.00
Sylt North
Sylt South
0
1000
2000 m


Probability Av on Sy
0.00
0.25
0.50
0.75
1.00
Sylt North
Sylt South
0
1000
2000 m


Probability Li on Sy
0.00
0.25
0.50
0.75
1.00
Sylt North
Sylt South
0
1000
2000 m

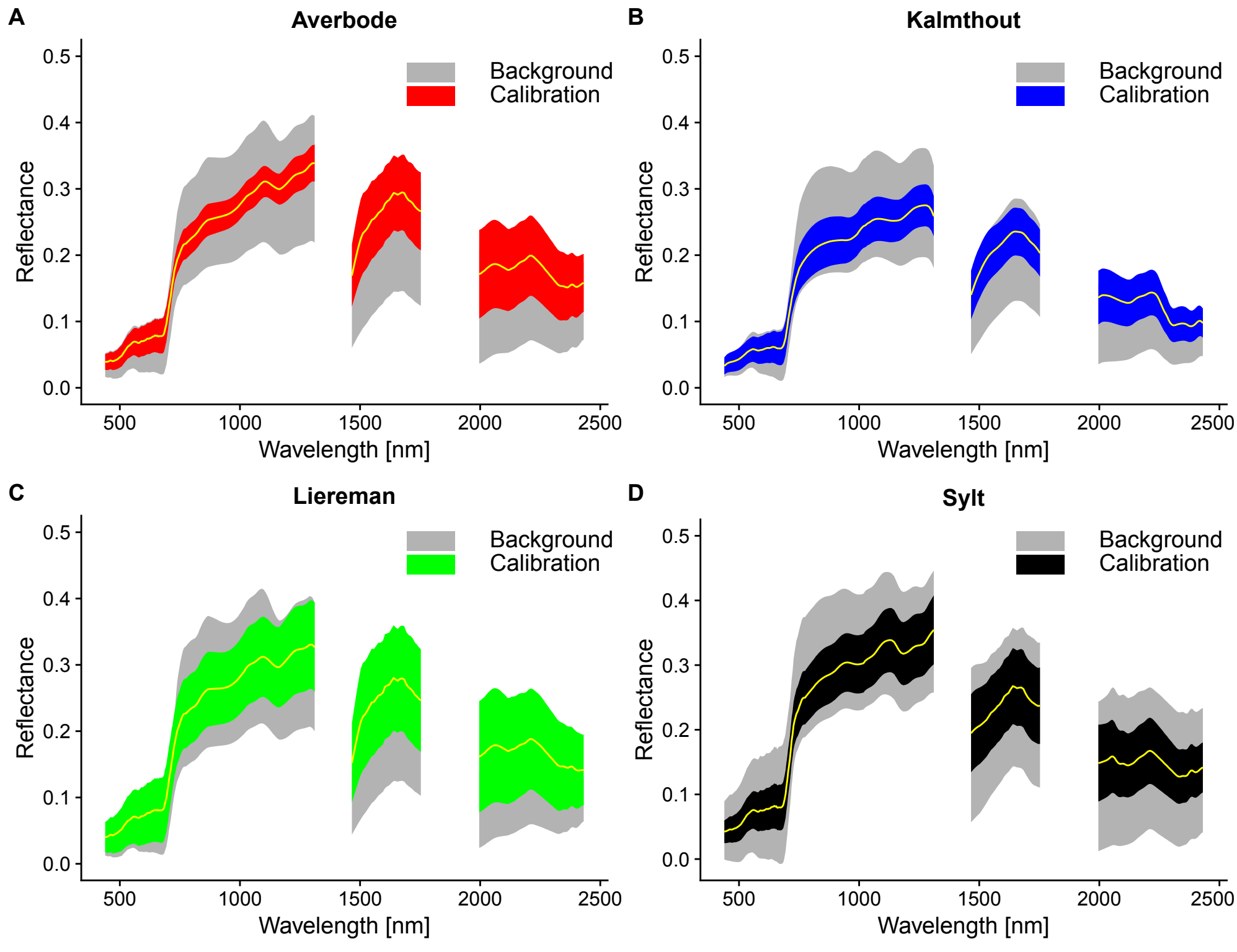

Probability Ka on Sy
0.00
0.25
0.50
0.75
1.00
Sylt North
Sylt South
0
1000
2000 m


Probability def on Sy
0.00
0.25
0.50
0.75
1.00
Sylt North
Sylt South
0
1000
2000 m


Probability opt on Sy
0.00
0.25
0.50
0.75
1.00
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.


A
Averbode
B
Kalmthout
C
Liereman
D
Sylt
Background
Calibration
Reflectance
Wavelength [nm]
0.0
0.1
0.2
0.3
0.4
0.5
500
1000
1500
2000
2500