biodiversa+
European Biodiversity Partnership

EUROPEAN PARTNERSHIP

MONITORING OF INVASIVE ALIEN SPECIES

Main findings from the first year of the Biodiversa+ pilot “Monitoring of invasive alien species with image-based methods”

Co-funded by
the European Union

Document Information

Grant Agreement number:	101052342
Project acronym	Biodiversa+
Project full name	The European Biodiversity Partnership
Biodiversa+ duration:	7 years
Biodiversa+ start date	Start date: 1st October 2021
More information about Biodiversa+	Website: www.biodiversa.eu Email: contact@biodiversa.eu @BiodiversaPlus Biodiversa+

Deliverable title:	D2.11 Mid-term report "Monitoring invasive alien species with image-based methods"
Authors:	Toke Thomas Høye, Lars Dalby, Kavi Mellerup, Asger Svenning, Nathan Pinoy (MoE_DK)
Contributors/reviewers:	Lars B. Pettersson (SEPA), Harriet Arnberg (SEPA), Arnaud Albert (OFB), Alix d'Audeville (OFB), Barbora Šingliarová (SAS), Carolina Parelho (FRCT), Helene Blasbichler (BOZEN), Magnus Tannerfeldt (Formas), Petteri Vihervaara (MoE_FI).
Work package:	WP2 Promote and support transnational biodiversity monitoring
Task or sub-task title:	Sub-task 2.6.2 sub-pilot on invasive alien species
Picture credits:	© Julia Seeber
Release date:	January 2024

Table of contents

What is Biodiversa+

Biodiversa+ is the new European co-funded biodiversity partnership supporting excellent research on biodiversity with an impact for policy and society. It was jointly developed by BiodivERsA and the European Commission (DG Research & Innovation and DG Environment) and was officially launched on 1 October 2021.

Biodiversa+ is part of the European Biodiversity Strategy for 2030 that aims to put Europe's biodiversity on a path to recovery by 2030.

The Partnership aims to connect science, policy and practise for transformative change. It currently gathers 80 research programmers and funders and environmental policy actors from 40 European and associated countries to work on 5 main objectives:

1. Plan and support research and innovation on biodiversity through a shared strategy, annual joint calls for research projects and capacity building activities
2. Set up a network of harmonised schemes to improve monitoring of biodiversity and ecosystem services across Europe
3. Contribute to high-end knowledge for deploying Nature-based Solutions and valuation of biodiversity in the private sector
4. Ensure efficient science-based support for policy-making and implementation in Europe
5. Strengthen the relevance and impact of pan-European research on biodiversity in a global context

More information at: https://www.biodiversa.eu/

Table of acronyms

AMI traps	"Automated Monitoring of Insects" traps based on a prototype from a research team at Aarhus University, Denmark and further developed into a standardised hardware system by UKCEH
API	Application Programming Interface
CamAlien	Camera system for monitoring invasive alien plant species along roadside verges. Produced by The AI Lab (https://theailab.dk/)
EUNIS	European Nature Information System (https://eunis.eea.europa.eu/)
ERDA	Electronic Research Data Archive, an open science data storage facility at Aarhus University, Denmark
EUNIS	European Nature Information System (https://eunis.eea.europa.eu/)
GBIF	Global Biodiversity Information Facility (www.gbif.org)
GDPR	General Data Protection Regulation. Regulation (EU) 2016/679
IAS	Invasive Alien Species
JRC	Joint Research Centre
STING	Science and Technology for Pollinating Insects
SPRING	European project on "Strengthening pollinator recovery through indicators and monitoring"

Executive Summary

The European Biodiversity Partnership, Biodiversa+, has established a series of pilot studies as a proof of concept for its tasks leading to the establishment of a transnational network of monitoring systems. In line with the Biodiversa+ priority on "Invasive Alien Species", the Biodiversa+ pilot on “Monitoring invasive alien species with image-based methods” was launched in 2023.

During 2023, the pilot involved 10 active partners across Europe, which participated in two modules: a plant module and an insect module. Both modules employ image-based monitoring approaches to the monitoring of invasive alien species. The plant module trials a camera system (CamAlien) mounted on a car to map invasive alien plant species along roads, while the insect module trials insect camera traps (AMI traps) for monitoring invasive alien moth species attracted to light. During the 2023 season, each partner has collected data in their own country. The data collection in the plant module has mostly progressed as planned and the most relevant subset of the resulting data has been processed using a newly developed data management and data processing pipeline developed in collaboration with the external partner Pl@ntNet. Data collection in the insect module has been challenged by delays related to the ordering, production, delivery and troubleshooting of the equipment. Despite these challenges, several partners have been able to collect data using this equipment and a preliminary analysis pipeline has been prepared. While there have been challenges in the first year, these have been overcome, and it is expected that data collection and analysis can be carried out according to plans in 2024, and that a plan for potential implementation of these tools in transnational monitoring of invasive alien plant and insect species will be presented as part of the final report of this pilot.

The pilot was coordinated by the Ministry of the Environment Denmark through Aarhus University as third party, and was conducted with ten active partners: The Autonomous Province of Bolzano, Italy (BOZEN), Denmark (MoE_DK), France (OFB), Sweden (SEPA), Czech Republic (NCA CZ), Slovakia (SAS), Bulgaria (ExEA), Croatia (MESD), Portugal, the Azores (FRCT), and Israel (MoEP).

1 Introduction

Evidence is mounting of widespread biodiversity decline across the globe. This gives a stark warning for the perilous state of our planet, yet the evidence base remains biassed to a few regions and a few species groups which are feasible to monitor. General conclusions on the status of biodiversity are complex given their diverse ecologies and high variability between and within taxon groups, over time and geographic regions. Repeatable sampling methods that can automate and expand the extent and resolution of biodiversity monitoring are urgently needed to provide robust estimates of long-term trends. Closing the knowledge gap has never been more important.

Biodiversa+, the European Biodiversity Partnership, aims at promoting and supporting transnational biodiversity monitoring, by building a transnational network of harmonised biodiversity monitoring schemes on specific priority topics (Basille et al. 2023). One of these Biodiversa+ priorities focuses on invasive alien species (IAS). As a way of supporting this harmonisation work on this priority, the Biodiversa+ partners agreed to launch in January 2023 a one-year pilot study on monitoring invasive alien species with image-based methods which was then extended by one extra year, hence lasting two years in total[1]. This pilot gathers ten countries: Italy - Autonomous Province of Bolzano with BOZEN, Denmark with MoE_DK, France with OFB, Sweden with SEPA, Czech Republic with NCA CZ, Slovakia with SAS, Bulgaria with ExEA, Croatia with MESD, Portugal (the Azores) with FRCT, and Israel with MoEP. The pilot was later extended by one year to include 2024 and will be joined by an additional new partner: Belgium - Flanders with VL O.

The aim of this pilot is to pave the way for and evaluate scalable methods using novel technologies for monitoring invasive alien species at the geographical scale of the partnership. More specifically, the pilot will: a) implement a coordinated international image-based monitoring scheme for invasive alien plant and insect species across biogeographical regions, b) improve image recognition for invasive alien plant and insect species, c) map benefits of and constraints to achieve a future real-time mapping of invasive alien plant and insect species as a decision support tool, and d) showcase the workflow of image-based monitoring for early detection of invasive alien species.

The initial specific objectives of the pilot are:

1) Collecting ten hours of image data along road sections with known occurrences of invasive alien plant species during the 2023 growing season for each partner.
2) Having three insect camera traps in operation for at least three months during the 2023 growing season for each partner.
3) Producing a database hosting all image data and associated metadata by the end of the pilot.
4) Training image classifiers for at least 15 plant species and 15 insect species of relevance to active partners, and applying them on image data from each active partner by the end of the project.
5) Producing visualisations of the locations, identity, and time of observation of at least 15 plant species and 15 insect species by the end of the project.

[1] Biodiversa+ biodiversity monitoring pilot: https://www.biodiversa.eu/biodiversity-monitoring/pilot/

This mid-term report describes the implementation steps taken in this pilot towards the transnational monitoring of invasive alien species, including a description of the equipment, taxonomic focus, site selection, data collection and image analysis. Further, the report summarises data management practices in the project and gives an overview of the first results and lessons learnt. The report ends with sections on the perspectives for implementation in long-term transnational monitoring. This report complements the living document describing the data management plan and the work plan for year two of the project.

2 Implementation

The IAS pilot consists of two modules: a plant module and an insect module. Both modules employ image recognition using computer vision and deep learning models for cost efficient and rapid detection of species of concern. These methods rely on training data of the species of concern to function effectively. Such data is available in the Global Biodiversity Information Facility (GBIF) as well as in databases of various national and regional pilot projects. The pilot program will be relevant in the context of introduction sites as well as for the collection of additional training data to improve image recognition models.

2.1 Equipment

2.1.1 Plant module

Invasive alien plants are monitored with a custom-built camera system mounted on a car. During driving, the camera records continuously and is capable of recording images without motion blur at driving speeds up to 110 km/h. A phone can be used to monitor camera settings in real time and view photos. An earlier version of the system used in the pilot is described in Dyrmann et al. (2021). The equipment used in this project is named CamAlien and is produced by the company The AI Lab. The instructions for the CamAlien system are available as Appendix 1 to this report. In addition, a series of Youtube videos give detailed descriptions of the operation of the system.

2.1.2 Insect module

Insects are monitored with Automated monitoring of insects (AMI) traps, which is a standardised camera-enabled light trap described in Bjerge et al. (2021). The AMI insect camera trap attracts nocturnal insects with UV light (LepiLED) and records insects landing on a white board. The AMI traps can be deployed with grid power or using solar panels and batteries. The traps run automatically on a predetermined schedule. The AMI trap images provide a rich basis for large-scale species inventories and abundance time series. The AMI trap manual is available as Appendix 2 to this report.

2.2 Taxonomic focus and site selection

2.2.1 Plant module

Each partner was tasked with listing five species of relevance to the country and five species for which data on occurrence exists per country. This was done to focus efforts on a manageable number of species. During the first year of the project, a collaboration with Pl@ntNet was established to build the image recognition pipeline around their facility. This development allowed for much greater flexibility in the taxonomic focus, and, as a result, the constraint on the number of species was relaxed. Nevertheless, the recording of images at high speed still requires that focal plants are clearly visible and identifiable from images recorded at a distance of 5–10 metres. This leads to the primary focus on larger herbaceous or climbing, shrubby or woody species.

The CamAlien camera system is designed for roadside verge monitoring. This is relevant in the context of the establishment and spread of invasive alien plant species, as roads and other transport networks are considered one of their primary dispersal routes. The camera system is capable of recording images along any road, but for consistency in the pilot scheme it was recommended to focus data collection on highways.

2.2.2 Insect module

Each partner was tasked with listing five species of relevance to the country and five species for which data on occurrence exists per country. This was done to focus efforts on a manageable number of species. From an image analysis perspective, the constraints on the taxonomic focus are determined by the computer vision approach and the capabilities of the deep learning models available. In this project, the classification models will allow for the species identification of moth species, which are identifiable from images recorded with the AMI trap cameras. This constrained the taxonomic focus to the subset of Lepidoptera species often referred to by naturalists as macro-moths. These are typically, but not always, larger and with more distinct morphological traits.

Regarding site selection, there is limited knowledge about the distribution of invasive alien insect species. As such, the AMI traps have the potential to reveal the occurrence of species not known to the region regardless where they are located. Ideally, potential establishment sites should be selected for the deployment of AMI traps. At the same time, there are logistical considerations with the deployment of AMI traps as they require a power source (e.g. grid or solar) and they are costly equipment with a risk of unwanted interference with the traps or even theft. Given these considerations, the work plan specified that primary sites should be botanical gardens or ports. For some partners, it was necessary to widen the scope to include other potential introduction sites for invasive alien insect species, such as plant nurseries or urban sites.

2.3 Data collection

2.3.1 Plant module

Instructions for the installation and operation of the CamAlien system are provided in Appendix 1. Each partner collects images during continuous recording for ten hours of driving. While driving, an expert in identification of invasive alien plant species (not the driver of the car) scans the roadside verge for the invasive alien plant species of concern. The camera equipment allows for tagging of image sequences when species of concern are observed. By pressing a button, a tag is added to images recorded during the previous few seconds. By tagging subsets of the images with observed species of interest, the co-pilot can help drastically minimise the search for data to evaluate model performance and ultimately feed additional data to the training of better image recognition models. Once satisfactory model performance has been achieved, future monitoring of invasive alien plant species will only require the driver and no expert in plant identification, and can thus be deployed at scale.

2.3.2 Insect module

Each active partner deploys three AMI traps. The AMI traps were provided with an illustrated manual for assembling the system and its operation (Appendix 2), in addition to the user manual for installation and deployment.

The ecologically relevant context of the sites will be described through vegetation surveys conducted in 2024. The goal is to assess the vegetation and habitats from each site with an AMI trap in 3 steps, all based on circles centred around the AMI trap: 1) record the occurrence of all vascular plant species and the cover of woody species above 1-m height within a circle with a 5-m radius; 2) record the occurrence of all woody species, and the cover of woody species above 1-m height within a circle with a 25-m radius; 3) record the cover of habitat types according to the EUNIS classification within a circle with a 100m radius.

2.4 Image analysis

2.4.1 Plant module

In this module, we are collaborating closely with Pl@ntNet. The image analysis pipeline has been integrated with the upload module to ensure that all images uploaded by the partners are managed automatically and in a consistent manner. First, the partners upload their data to the Electronic Research Data Archive (ERDA) at Aarhus University (erda.au.dk). After the user has uploaded an image, an API developed by AU is called and as a first step it will determine if the image name is parsable. Next, the metadata is stored in the database and the image moved to storage. Lastly, a call to Pl@ntnet is made and the results returned are written to the database. A detailed flow diagram of the steps involved in this process is presented in Fig. 1. Collaborators at Pl@ntNet, have prepared a dedicated adaptation of the Pl@ntNet application programming interface (API) involving custom-adapted parameters, which this pilot is the only user of so far. This process involves calling the API several hundred times for each image to assess the likelihood of occurrence of the species of interest in multiple regions of each image and summarising the outcomes of these calls into image-level metrics of confidence scores as the maximum score across all calls to the API.

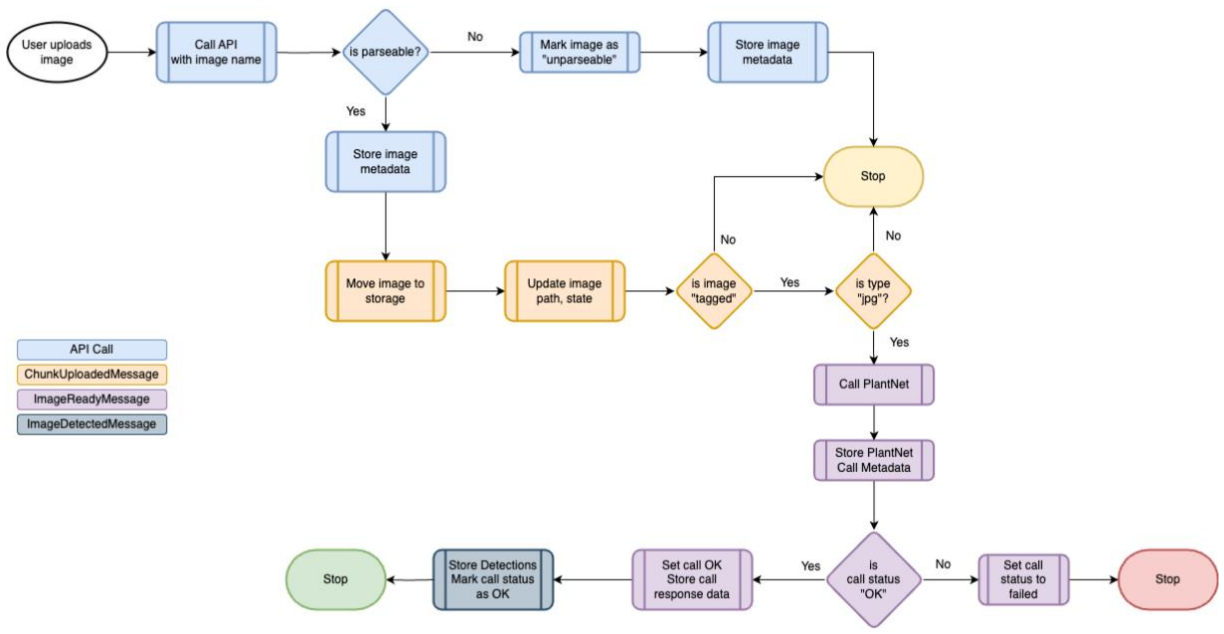

Figure 1. Flow diagram describing the process for each image recorded with the CamAlien system from the plant module uploaded to the Electronic Research Data Archive (ERDA) at Aarhus University.

2.4.2 Insect module

The extraction of species level data from the AMI trap images involves three key steps: localisation, filtering, and classification. All three steps of the pipeline are complex tasks for which currently three different deep learning models are involved. As part of ongoing collaborations, a framework for localising and classifying insects in images collected with AMI traps has been developed[2]. The localisation model is assumed to work generically for images from any region, but the current set of classification models are specific to particular biogeographic regions. For this project, additional work on improving the localisation model is carried out and a new regional classification model is being developed. A first version of both models is now available.

The new localisation algorithm for use on the AMI trap images is a YOLOv8 instance segmentation model. It has been pre-trained on images from several other automatic insect localization projects. To improve its performance, all insects from >150 AMI trap images have been manually annotated and used to train this model. The model will be further improved, but is already very promising (Fig. 2). Using instance segmentation for insect localization offers a major advantage compared to bounding boxes. When two or more insects are close to each other, bounding boxes can end up containing several insects. It is especially the case for the largest ones. Using polygons allows to extract each insect more precisely, which at the end, will facilitate the classification process.

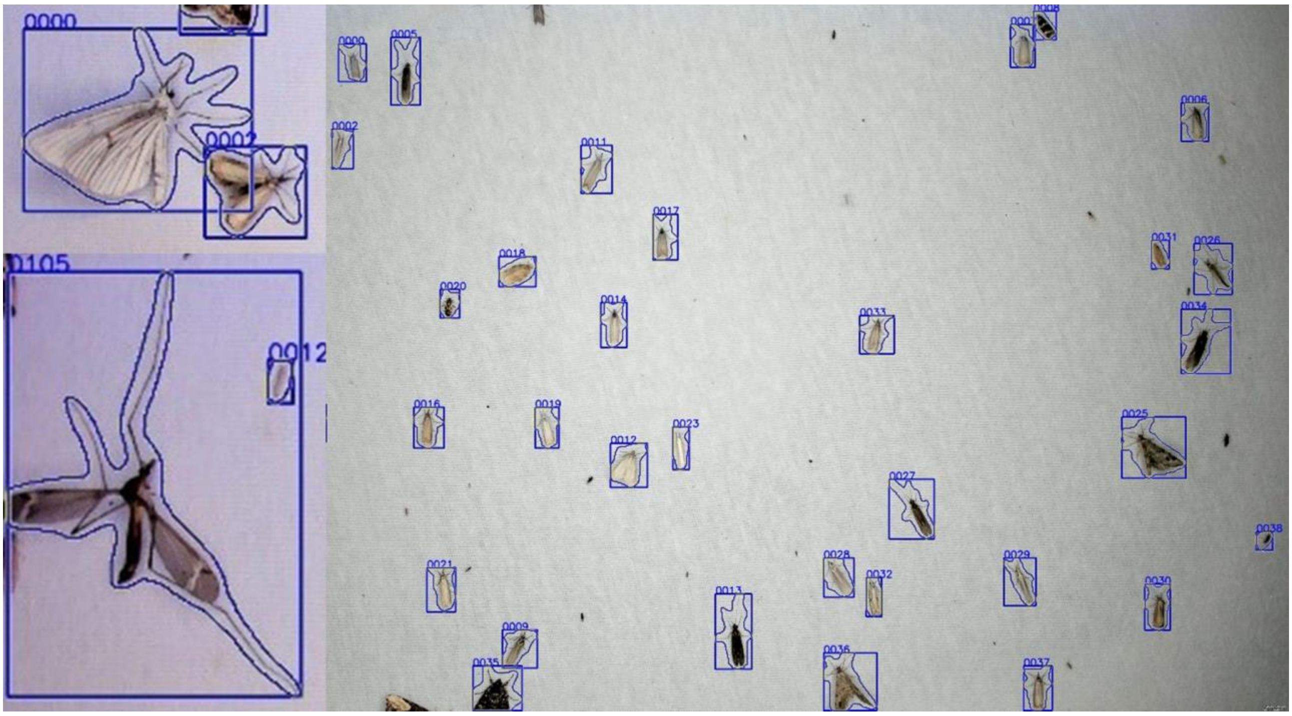

Figure 2. Example of an image recorded with the AMI trap. The new localisation model for AMI trap images draws a segmentation polygon as well as the rectangular bounding box around each detected animal. These are also shown in the main image as well as in two insets of zoomed in details from a similar image highlighting details for individual animals (Geissmann et al. unpublished data).

[2] https://github.com/RolnickLab/ami-data-companion

For the filtering and localisation tasks, a key advantage is that the insects found in the AMI trap images are also represented in global citizen science databases and GBIF. Although these images have many different inherent biases and are taken with different cameras against different backgrounds, they offer an important resource when training image recognition models. One of these opportunities in the context of this project, is to use these images as a dataset for the initial training phase. In this project, we use five steps to create the species classification model: (1) choosing a backbone model, here we used EfficientNet (Tan & Le, 2019), (2) applying an appropriate classification head, (3) collecting and training on citizen science images of the species-of-interest from GBIF, (4) manual annotations of a small number of localised objects from the AMI images and (5) fine-tuning on the manually annotated objects from the AMI images.

The GBIF dataset currently used is based on a list, kindly provided by the Dutch Butterfly Conservation, containing 3235 distinct macro-moth species found in Europe, and records of the countries, which are considered within their range. After conservative name resolution with the GBIF Backbone Taxonomy (i.e. the taxonomic reference that GBIF uses to consistently name species from multiple sources), the GBIF database was queried for observations (with images) for every species-range combination, yielding at least one observation for 2331 of the 3235 moth species and a total of 2,574,470 observations. This dataset was then filtered to exclude images of larvae, pupae, or large images where the moth takes up a small proportion of the image. The remaining dataset consisted of 1,873,998 observations of clearly visible adult moths. Lastly, the distribution of species frequencies was capped to remove the excessive amount of observation related to a few common and charismatic species: we chose to limit the maximum number of observations to 5,648 (from 41,117), leading to a retention of 80 % of the images and a final dataset size of 1,499,199 images.

During the pilot, the output format of data from the models will be adapted to comply with the newly published CamTrap DP model (see: Bubnicki et al. 2023) and will deliver ecological data in the form of Essential Biodiversity Variables[3] on species counts and abundance estimates.

[3] For more information on Essential Biodiversity Variables, see the Biodiversa+ report by Silva del Pozo et al. 2023: https://www.biodiversa.eu/wp-content/uploads/2023/10/Biodiversa_Best-practices_2023_v5_WEB.pdf p.18

3 Data management

3.1 Data description

For both modules, the species lists provided by the partners included GBIF IDs to ensure smooth integration with classification algorithms using these same GBIF IDs for training and prediction species identities.

The CamAlien system records images (4096 × 3000 pixels in tiff format as well as a compressed version of each in image in jpg format) in two different user defined modes. During the IAS pilot, the recording is continuous when the system is powered on. It is also possible to run the system in a mode, where recording is only done when a remote-control button is pushed. During continuous recording, five images are recorded per second. Each image is stored in a buffer on the computer, which is part of the system. The location (latitude, longitude and altitude), speed during recording, exposure time, gain, date and time of recording are stored as part of the filename of each image. In addition, the user can add a tag stored in the filename to indicate the presence of invasive alien species by pressing a button on a remote control. The CamAlien system can inadvertently record images of people or registration plates of cars. During the recording process a software installed on the system directly processes each image and blocks out detected people, bikes and cars to comply with the General Data Protection Regulation of the European Union (GDPR).

The AMI traps record images (4096 × 2160 pixels in jpg format) based on a user defined schedule. In the IAS sub-pilot, the recording period is set to 11pm to 3am each night. During this interval, the camera records a snapshot image every 10 minutes. The system also runs a motion detection software and evaluates if motion has happened in the past two seconds, and records an additional image if this was the case (Bjerge et al. 2021). The date and time of recording are stored as part of the filename of each image.

3.2 Ensuring the FAIR data principle

All image data collected as part of the IAS pilot is stored in the Electronic Research Data Archive (ERDA) at Aarhus University, Denmark and will be made publicly available through a Danish implementation of the dataverse.

All data created in the pilot study will be made openly available at the end of the pilot. During the first year, the project has worked directly with the European Alien Species Information Network (EASIN) to ensure image-based species observations will be presented through EASIN's data portal, and associated with appropriate measures of uncertainty. The intention is to submit species occurrences to the Global Biodiversity Information Facility (GBIF), and through tagging allow EASIN to present the data directly on their species occurrence portal.

4 First results

4.1 Plant module

4.1.1 Data recording

All partners were able to record images using the CamAlien system. For various reasons, not all partners recorded a similar number of images. In some cases, this was caused by damage to the equipment during transport, logistical constraints or misunderstandings with the recording procedure. A key element was whether the system recorded images continuously as planned, or only during tagging. Together, this resulted in substantial differences in the amount of data that was collected by each partner (Fig. 3).

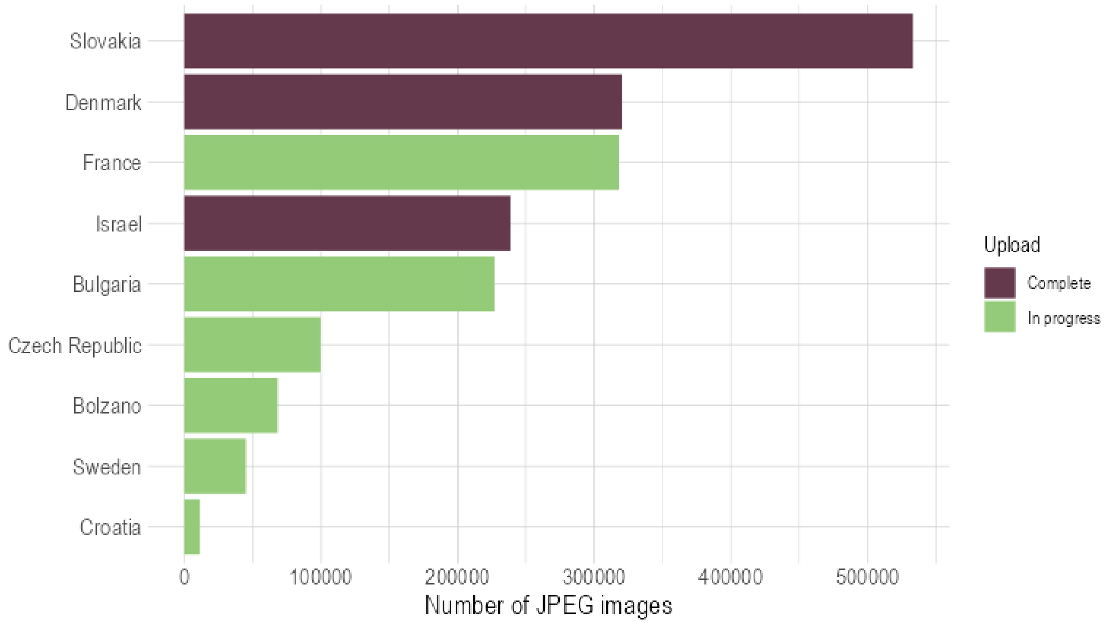

Figure 3: Number of images stored at ERDA. Not all partners were able to upload all their images due to technical difficulties with the upload. The remaining images are being uploaded by MoE_DK.

The camera was recording continuously while driving, but the operator was only tagging the images captured using the remote control when the invasive alien species were observed. The percentage of tagged images varied between partners, but was between 2% and 14% percent of the images captured except for Croatia which had 85% tagged images as they were only recording when activating the system via the remote instead of using continuous recording (Fig. 4).

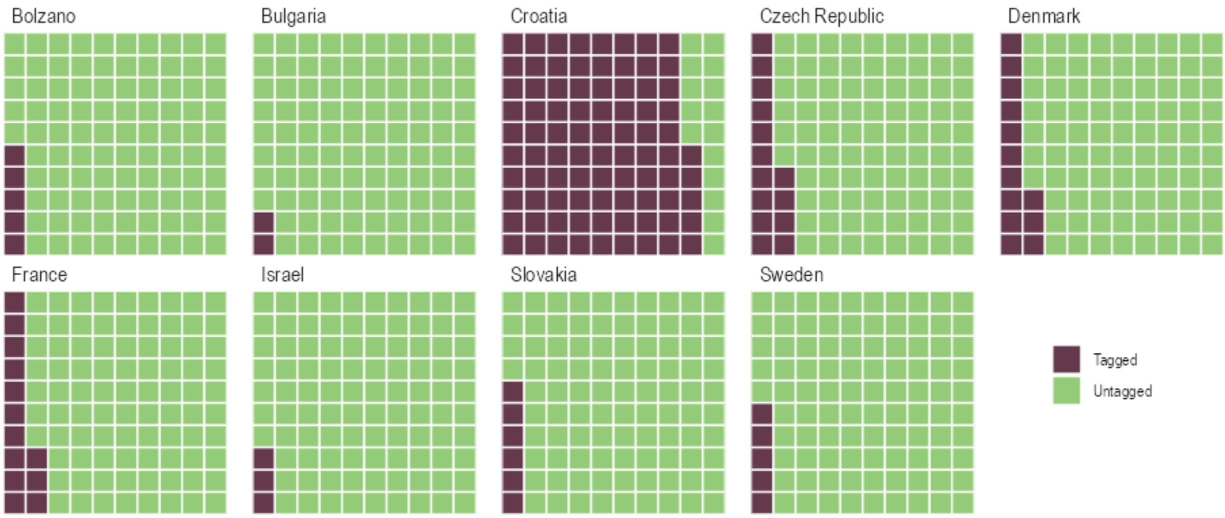

Figure 4: Proportion of tagged images. Each square represents one percent of images captured by each partner.

4.1.2 Image processing and annotation

The Pl@ntnet algorithm returns a confidence score for each detection of a plant species in the image analysed. However, it is not known in advance what the confidence score should be before it can be assumed that the plant species detected was actually present at the location where the image was taken. To estimate the confidence score at a given threshold value for each species we selected images with scores along a gradient from 0.1 to 1 (the maximum possible value) and conducted expert review of the detections. With annotations from the expert review the species-specific threshold score could be estimated (see Fig. 5 for an illustration on three species).

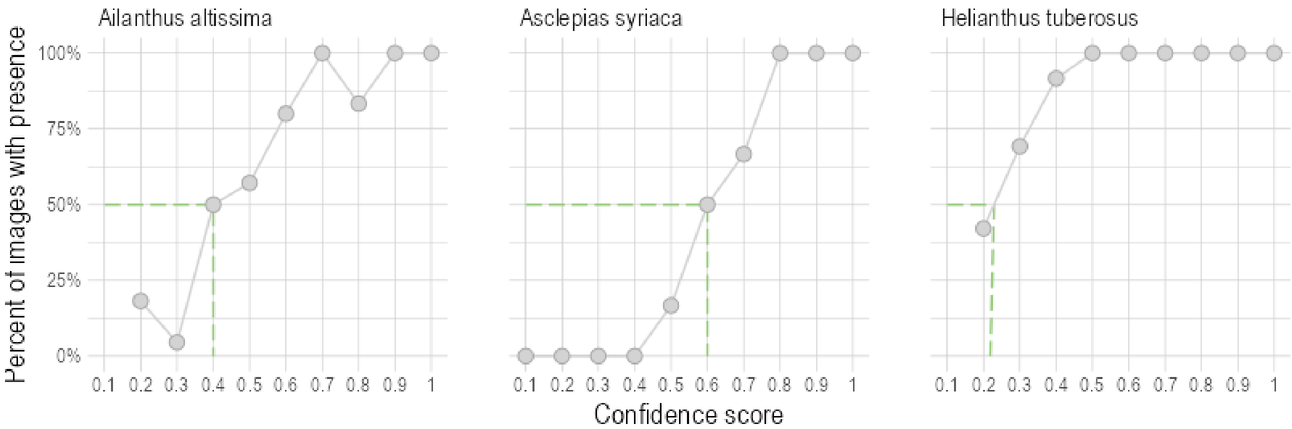

Figure 5: Examples from annotation. Relationship between the percent of the images with the focal species correctly detected by the Pl@ntNet algorithm and the confidence score from the model. Dashed green line shows the threshold confidence score for each of the species at a threshold of 50% (*A. altissima*: 0.4, *A. syriaca*: 0.6 and *H. tuberosus*: *0.22)*. Sample sizes: *A. altissima*: 90, *A. syriaca*: 44 and *H. tuberosus*: *95*.

4.1.3 Maps

Using the threshold score from the expert review process we can map the occurrences of the focal species (Fig. 6–11). Only the tagged images have been analysed and hence we can only map the positions of the species detected in those.

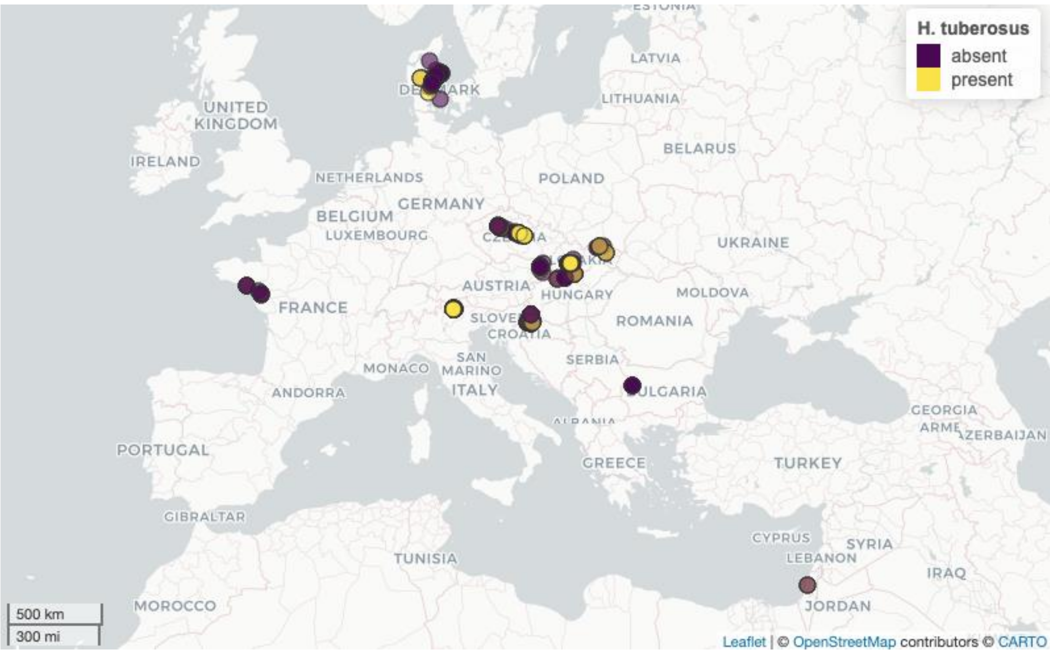

Figure 6: Map of *H. tuberosus* detections. Purple point markers are image locations where the confidence score was below the threshold for at least 50% presence. Yellow points are image locations with confidence scores above the 50% threshold.

Figure 7: Detailed map of *H. tuberosus* detections. Purple point markers are image locations where the confidence score was below the threshold for at least 50% presence. Yellow points are image locations with confidence scores above the 50% threshold.

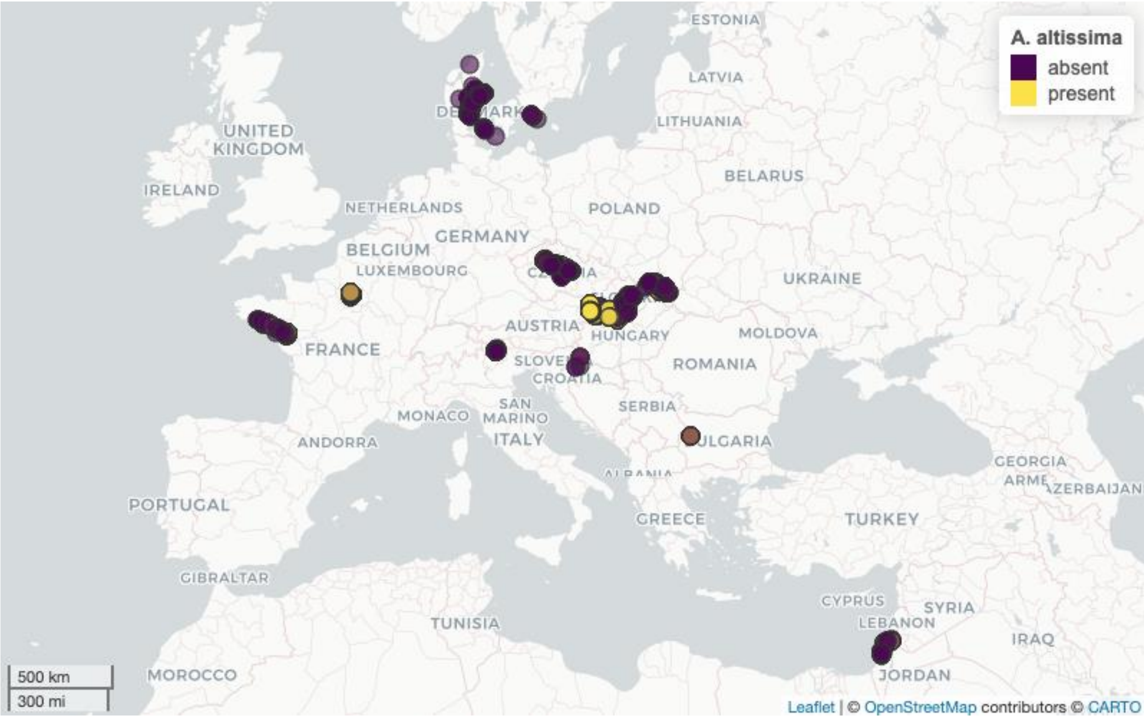

Figure 8: Map of *A. altissima* detections. Purple point markers are image locations where the confidence score was below the threshold for at least 50% presence. Yellow points are image locations with confidence scores above the 50% threshold.

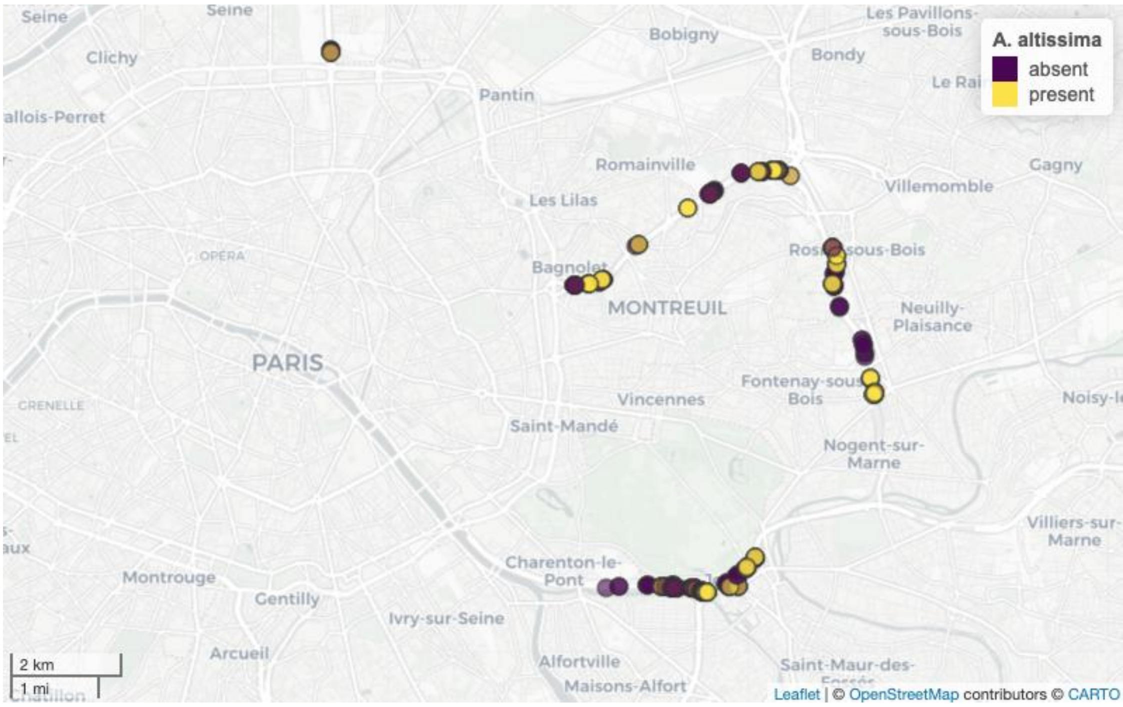

Figure 9: Detailed map of *A. altissima* detections. Purple point markers are image locations where the confidence score was below the threshold for at least 50% presence. Yellow points are image locations with confidence scores above the 50% threshold

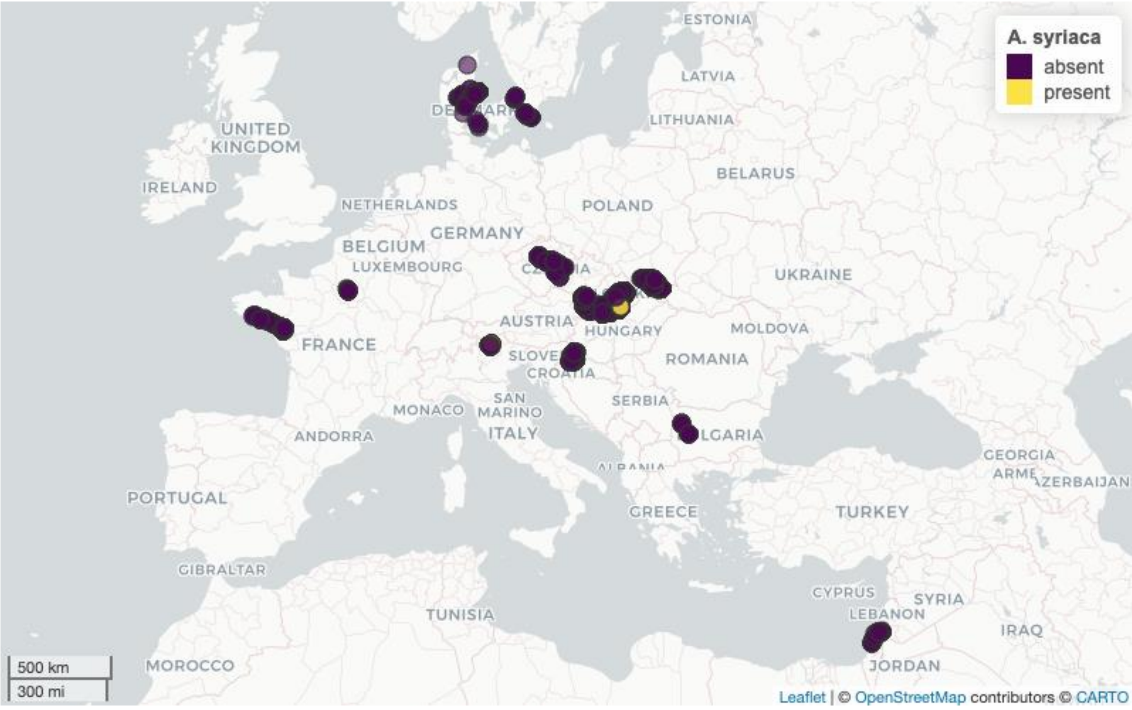

Figure 10: Map of *A. syriaca* detections. Purple point markers are image locations where the confidence score was below the threshold for at least 50% presence. Yellow points are image locations with confidence scores above the 50% threshold.

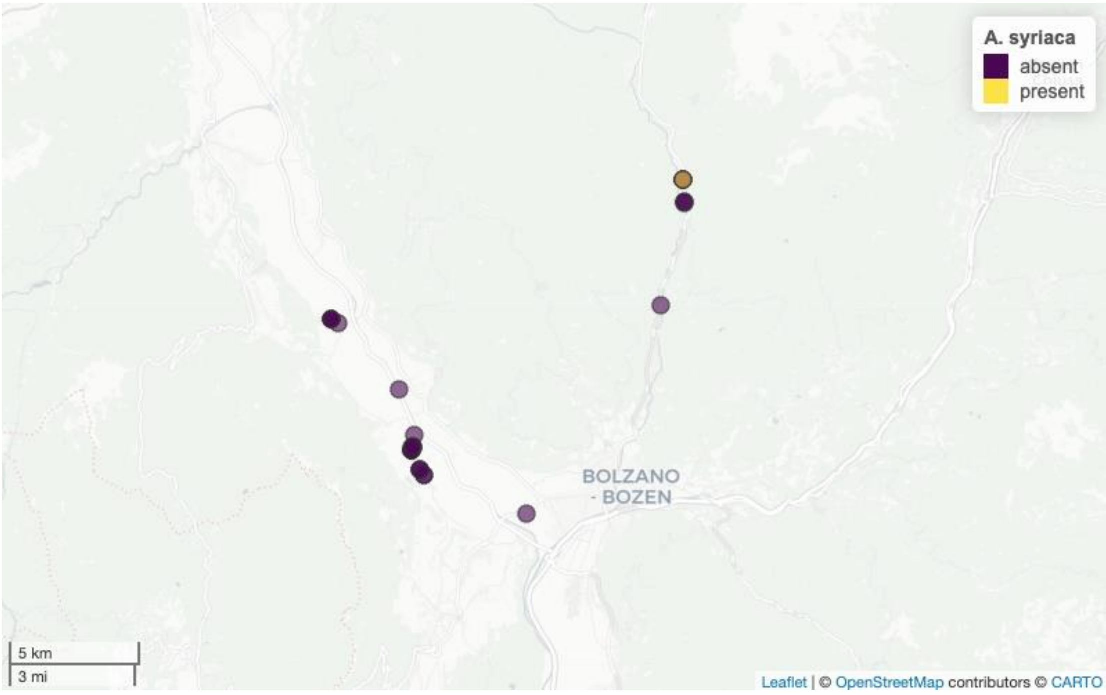

Figure 11: Detailed map of *A. syriaca* detections. Purple point markers are image locations where the confidence score was below the threshold for at least 50% presence. Yellow points are image locations with confidence scores above the 50% threshold.

4.2 Insect module

Due to challenges with the delivery of the AMI traps used in the insect module, data collection was delayed for all partners. However, most partners managed to collect data during some nights. We are using these images to trial our image analysis pipeline, which consists of localisation and classification modules as described above. Once all image data from the AMI traps has been received, the results of the image analysis pipeline will be prepared and presented in the final report of the pilot.

5. Lessons learnt

5.1 Equipment orders, delivery and functionality

With both modules, it has been an important lesson that the administrative requirements for purchase of expensive scientific equipment entails substantial work and takes time. Adding to this that each partner has their own administrative procedures, it was not really possible to learn from other partners or give common instructions. A key point of the pilot is that each partner purchases their own equipment. As such, it is neither possible to place one big order for the entire pilot, nor organise shipping, import and payment centrally. With a highly compressed time plan, such procedures can easily put a project at risk. Despite strong dedication, the pilot suffered from delays in the ordering, production and delivery of the equipment to the partners. Some partners were experiencing substantial delays in their ability to order equipment, which resulted in delayed production, particularly of the AMI traps. For the CamAlien systems, most partners received the equipment at a time of the field season, where data collection could meaningfully be carried out. However, since the provider of the AMI traps required a certain number of orders to be placed formally before production could start, this affected the entire module.

Once delivered, some challenges with both types of equipment were identified. For the CamAlien system, these related mostly to damage to the equipment during delivery. This damage during delivery transport was substantial and further delayed the deployment of AMI traps for several partners. It is anticipated that in the future this can be mostly avoided by better packaging of the traps for shipping, but some partners also reported damage to the outside of parcels indicating very rough treatment during transport. In addition, it was experienced by several partners, that the SSD hard drives in the CamAlien would stop working properly following hard shutdowns of the systems. There is an ongoing effort with the providers of the systems to remedy this issue. Finally, aspects of the software and installation manuals caused further delays. A thorough evaluation of these issues with the providers of the AMI traps has been carried out and it is expected that AMI traps will function as planned during the 2024 field season.

5.2 Data collection and data transfer

Generally, the data collection with the CamAlien system was smooth, and most challenges were solved during the recurrent project meetings with all partners every two weeks. It was challenging for partners to upload the data from the CamAlien system to the central server. For some partners, it took days or even weeks to upload and not all partners managed to upload all data. In part, local security settings and bandwidth limitations were causing these issues, but the custom built upload module also added constraints to the upload speed. The in-person meeting in November allowed the remaining data on disks to be handed over and be directly uploaded (ongoing process).

The CamAlien system produces approximately 300 GB per hour of driving with continuous recording. Uploading this amount of data to the Electronic Research Data Archive (ERDA) located at a datacentre at Aarhus University is time consuming even on fast internet connections. To aid the process of uploading and to ensure that data from each partner was stored in the correct final location we built an application which was used to do the actual upload. Nine of the ten partners uploaded at least some of their data from 2023. The uploader program had a feature which allowed for resuming an ongoing upload. This worked well during the testing phase, but we experienced multiple cases where the resume function did not work as intended. Intermittent loss of internet connection (on either end) was most likely the cause of most of these issues. Unstable internet connections is a condition we need to be able to work around, hence we plan to modify the upload process for 2024 to be more robust to loss of internet connection. Relying on existing SFTP clients like FileZilla might be the way forward. We will be exploring this solution during the winter in order to have the new module ready by spring 2024.

For 2023, we chose to work with the compressed JPEG version of the high-resolution TIFF. This reduced the amount of data to upload substantially. On the other hand, it remains to be understood if this comes at a substantial cost in species recognition capabilities. It is likely that the original TIFF images capture details of importance for species recognition that are lost during the compression of the images to jpg format. In order to store the TIFF files, we then needed to physically transfer the hard drives to Aarhus for upload to ERDA. This is not a viable long-term solution as it is time consuming and there is also the risk that disks are lost or damaged in shipment. Changing to DNG file format for 2024 reduces the volume of data in a loss-less manner and should allow us to upload all data.

For the insect module, several partners found it difficult to list relevant insect species, as knowledge about the occurrence of insect species is limited. Due to the nature of the monitoring approach, it was also a requirement that the species are attracted to light and sufficiently conspicuous to be identifiable from the images.

The AMI traps produce much less data per hour of operation compared to the CamAlien system. Due to the very short field season 2023 in most partner countries, the data transfer constituted less of a challenge. The AMI trap data was uploaded to the data storage facility ERDA using the file handling software FileZilla. This system facilitates efficient transfer of data, but does not allow users to add additional metadata related to the raw data. As such, it introduces a risk of loss of key information related to the raw data such as hardware metadata and location of the AMI trap during deployment. So far, five of the ten partners have uploaded their images from 2023.

6 Perspectives for long-term transnational monitoring

Invasive alien species are, by definition, a problem that needs to be addressed at the transnational scale (IPBES 2023). The data from each module can feed directly into a transnational database and image processing pipeline for generation of transnationally standardised and automated monitoring data.

The image recognition tools developed locally, will be of great value at the transnational scale. One important perspective is that the tools can be used to monitor species already introduced to the country. Another important perspective is that models of species found in one country can be used for early detection of species as they may become introduced in other countries. Third, the economic burden of maintaining the digital infrastructure for databases and image-recognition will be much smaller for each active partner, if shared at a Pan-European scale. Finally, the image-recognition models will be more robust, when trained on data across a larger geographical scale with different plant and insect communities.

An in-person workshop for all pilot participants was organised by the Ministry of Environment of Denmark on 17 November 2023 in Bolzano, Italy. It was held back to back with the soil biodiversity pilot to minimise travel for partners involved with both pilots. The main objective of the workshop was to collectively identify lessons learnt during the first year of the project and to discuss perspectives for the implementation of the monitoring approaches in the IAS pilot in long-term transnational monitoring.

Among the opportunities and perspectives identified are the scalability of the methods trialled in this project. In the long-term, the image-based approaches offer highly cost-effective means to map and monitor the distribution of invasive alien species and can substantially shorten the time from an observation is made until it enters databases and decision support systems on which management decisions are taken.

The pilot delivers new ways of standardised and cost-efficient monitoring of species with automated image recognition. Reliable data on the natural environment is fundamental to support science on species conservation in human-modified landscapes. The pilot results can further be used to evaluate if adapted systems could be used to monitor additional taxa of invasive alien species (e.g. mammals often observed as road kills or day active insects such as the Asian hornet, fruit flies or other day active insect species) with image-based tools.

A number of research infrastructures and international projects can support the implementation of these tools in long-term transnational monitoring. Here, we list a few:

- The EASIN information network established by the JRC provides collated information from EU member states as well as regional initiatives such as NOBANIS and ESENIAS. None of these sites present information about monitoring invasive alien species.
- Pl@ntNet is an identification system that helps with the identification of plants through images. It is a research and a citizen science project, initially supported by Agropolis Foundation, and developed since 2009. The Pl@ntNet system works by comparing visual patterns transmitted by users via photos of entire plants or plant organs (e.g. flowers, fruits,

leaves) that they seek to determine. These images are analysed and compared to an image database produced collaboratively. Pl@ntNet is governed by a consortium composed of four research organisations: CIRAD, Inria, INRAE, IRD and the Agropolis Foundation. The management of the consortium is part of the InriaSOFT program, a program to perpetuate the digital achievements of Inria and its partners.

- The AMI traps represent one of the recommendations from the STING expert group appointed by the European Commission Joint Research Centre. Among other tasks, STING develops recommendations on novel technology for pollinator monitoring, which should be tested at European scale for moth monitoring.
- The COST Action Alien CSI has already provided input to the development of this pilot and their expertise will continue to support the development of exact protocols for the spatial design and species selection for the pilot
- The Cost Action InsectAI launched in November 2023 will be engaged in the insect module regarding image analysis. In the longer term, this network can also contribute to the further development of improvements and tests of hardware and software and standards for deployments and model training as well as through the collection of additional training data.
- Horizon Europe projects such as MAMBO will provide additional technical expertise on image recognition and will specifically contribute to the training of deep learning models and biomass estimation. This project has an emphasis on monitoring of species and habitats across the EU with an emphasis on insects.

7 Next steps

The workplan for year 2 of the Biodiversa+ IAS pilot is described in a separate document. In summary, the data collection for both plant and insect modules will be continued in 2024 and the data analysis and management infrastructure will be consolidated to prepare for the possible implementation in long-term monitoring. There will be software updates to the CamAlien and the AMI trap equipment and a possible hardware upgrade on the CamAlien system. One new partner (VL O, Belgium - Flanders) will join the group of active partners. They have already been introduced to the group and will be involved with the data collection and ensuring integration with data publishing models to ensure the flow of data to GBIF.

References

Basille, M., Body, G., Eggermont, H., Mandon, C. and Vihervaara, P. (2023). Shared goals and priorities for biodiversity monitoring within Biodiversa+. *Biodiversa+ report*. 22pp. Available at: https://www.biodiversa.eu/wp-content/uploads/2023/06/D2.5-Priorities.pdf

Bjerge K, Nielsen JB, Sepstrup MV, Helsing-Nielsen F, Høye TT. (2021) An Automated Light Trap to Monitor Moths (Lepidoptera) Using Computer Vision-Based Tracking and Deep Learning. *Sensors*. 2021; 21(2):343. https://doi.org/10.3390/s21020343

Bubnicki, J. W., Norton, B., Baskauf, S. J., Bruce, T., Cagnacci, F., Casaer, J., Churski, M., Cromsigt, J. P., Farra, S. D., Fiderer, C., Forrester, T. D., Hendry, H., Heurich, M., Hofmeester, T. R., Jansen, P. A., Kays, R., Kuijper, D. P., Liefting, Y., Linnell, J. D., … Desmet, P. (2023). CAMTRAP DP: An open standard for the fair exchange and archiving of Camera Trap Data. *Remote Sensing in Ecology and Conservation*. https://doi.org/10.1002/rse2.374

Dyrmann M, Mortensen AK, Linneberg L, Høye TT, Bjerge K. (2021) Camera Assisted Roadside Monitoring for Invasive Alien Plant Species Using Deep Learning. *Sensors*. 21(18):6126. https://doi.org/10.3390/s21186126

IPBES Invasive Alien Species global assessment: IPBES (2023). Summary for Policymakers of the Thematic Assessment Report on Invasive Alien Species and their Control of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Roy, H. E., Pauchard, A., Stoett, P., Renard Truong, T., Bacher, S., Galil, B. S., Hulme, P. E., Ikeda, T., Sankaran, K. V., McGeoch, M. A., Meyerson, L. A., Nuñez, M. A., Ordonez, A., Rahlao, S. J., Schwindt, E., Seebens, H., Sheppard, A. W., and Vandvik, V. (eds.). IPBES secretariat, Bonn, Germany. https://doi.org/10.5281/zenodo.7430692

Tan, M., & Le, Q.V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *International Conference on Machine Learning*. https://doi.org/https://doi.org/10.48550/arXiv.1905.11946

Michelle Silva del Pozo, Guillaume Body, Gaby Rerig, Mathieu Basille. (2023). Guide on harmonising biodiversity monitoring protocols across scales. Biodiversa+ report. 60 pp. Available at: https://www.biodiversa.eu/wp-content/uploads/2023/10/Biodiversa_Best-practices_2023_v5_WEB.pdf