

# Unified Automatic Plant Cover and Phenology Prediction

Matthias Körschens<sup>1,2</sup>

Solveig Franziska Bucher<sup>1,2,3</sup>

Christine Römermann<sup>1,2,3</sup>

Joachim Denzler<sup>1,2,3</sup>

<sup>1</sup> Friedrich Schiller University, Jena, Germany

<sup>2</sup> German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

<sup>3</sup> Michael Stifel Center Jena, Jena, Germany

{matthias.koerschens, solveig.franziska.bucher,christine.roemermann, joachim.denzler}@uni-jena.de

## Abstract

*The composition and phenology of plant communities are paramount indicators for environmental changes, especially climate change, and are, due to this, subject to many ecological studies. While species composition and phenology are usually monitored by ecologists directly in the field, this process is slow, laborious, and prone to human error. In contrast, automated camera systems with intelligent image analysis methods can provide fast analyses with a high temporal resolution and therefore are highly advantageous for ecological research. Nowadays, methods already exist that can analyze the plant community composition from images, and others that investigate the phenology of plants. However, there are no automatic approaches that analyze the plant community composition together with the phenology of the same community, which is why we aim to close this gap by combining an existing plant cover prediction method based on convolutional neural networks with a novel phenology prediction module. The module builds on the species- and pixel-wise occurrence probabilities generated during the plant cover prediction process, and by that, significantly improves the quality of phenology predictions compared to isolated training of plant cover and phenology. We evaluate our approach by comparing the time trends of the observed and predicted phenology values on the InsectArmageddon dataset comprising cover and phenology data of eight herbaceous plant species. We find that our method significantly outperforms two dataset-statistics-based prediction baselines as well as a naive baseline that does not integrate any information from the plant cover prediction module.*

## 1. Introduction

The plant community composition as well as plant species phenology are essential indicators of environmental

change like land-use [15, 2, 4], insect abundance [20, 18], and climate change [16, 17, 14]. Phenology has even been identified as a "fingerprint of climate change" [16]; most species advance their flowering phenology in response to increasing temperatures, though the intensity with which plants respond is species-specific [7].

The plant community composition is usually monitored by ecologists directly in the field by estimating the plant cover, i.e., the percentage of area of ground covered by each plant species. Based on this value, phenology is also monitored by additionally estimating the ratios of plant individuals of a certain species in a particular phenological state, like flowering or leaf senescence. From these estimates, different aspects can be investigated, like the shift of the first flowering day (FFD) of certain plant species or the flowering duration, such that impacts of changes in the environment on plant phenology and species composition can be analyzed.

However, the manual collection of these values is highly laborious, time-consuming, and subjective, so they are surveyed in rather long intervals of at least weekly, but sometimes even longer intervals (e.g., biweekly or monthly). To enable a speedy and reliable collection of vegetation and phenology data, automated systems are required, which can collect images and analyze them with little human interaction.

While the automatic analysis of vegetation images with respect to the plant cover has already been done [9, 8, 12], the phenology of the plants in the images is ignored in these studies. Other works are concerned with the phenology of plants in different aspects [1, 23, 19, 22], usually focusing on single plant species, individual plants and monocultures. However, none analyzed entire species-rich and therewith complex plant communities simultaneously, especially not for herbaceous species.

We bridge the gap between the analysis of the plant community composition and phenology by combining these two

in our novel method introduced here. We base our approach on [9], which utilizes convolutional neural networks (CNNs) and comprises two pre-training steps and a final plant cover training phase. We extend the latter with concurrent phenology training. To enable the joint prediction of plant cover and phenology at the same time, we introduce a new phenology prediction module, which utilizes the species-wise occurrence probabilities from the plant cover prediction module as a base for the prediction of the phenology of each plant species.

Here, we will focus on two phenological stages, which are *flowering* and leaf *senescence* (i.e., the coloring of the leaf). To enable detailed analyses of phenological changes across the year including information on species abundances and species-specific differences in phenology, we will also estimate how many percent of the detected plant species are currently in each phenological stage.

Our contributions are:

- We introduce the new task of automatic phenology prediction for herbaceous plant communities on the example of the InsectArmageddon [20] dataset, and discuss its caveats.
- We show two simple calculation models, which extend a plant cover prediction method to also predict phenology in combination, and thus can lead to significant improvements of the results on phenology prediction in comparison to isolated plant cover and phenology training.
- We show a simple way to combine the two calculation models to leverage their advantages and improve prediction results even further.

## 2. Related Work

Several approaches try to automatically analyze phenology from images. Wang *et al.* [22] investigate the flowering of trees, separating the flowering process into eight different stages, which are viewed as a classification problem. Their approach is more focused on flowering phenology and does not perform species identification, as only a single species is investigated. Our method, in contrast, investigates communities of several herbaceous species and focuses on flowering and senescence.

Yalcin [23] investigates the phenology of agricultural plants, utilizing an AlexNet [11] to classify the images of different crops into several phenological stages to automatically assess the plants' development. In their investigations, they only analyze images of homogeneous communities of agricultural plants and only perform a simple classification, while we predict percentages for each stage and species in our dataset.

Triki *et al.* [19] investigate dried individuals of herbaceous species from herbaria for reproductive organs, i.e., flower buds, fruits, and flowers. To this end, they employ a modified Mask Scoring R-CNN [6] to segment the organs based on fully-supervised segmentation labels. While they also analyze herbaceous species, as in our work, they only investigate single individuals in idealized conditions, while we tackle complex images of large plant communities.

In terms of automatic analyses of plant communities, there merely exist the approaches by Körschens *et al.* [12, 8, 9], who only investigate the plant cover of the communities, but not the phenology. Our work will be based on these approaches and extend them for phenology prediction.

## 3. Approach

### 3.1. Plant Cover Prediction as Base Approach

Our base approach is the CNN-based segmentation pre-training approach from [9]. With this method, a classification network is first trained, which is then used to generate segmentations. These segmentations are then used to train a second network, utilized as initialization for the joint plant cover and phenology training with dedicated annotations.

We combine the species' phenology calculation in a novel phenology prediction module with the already existing plant cover prediction module used in [8, 9]. In this cover prediction module, pixel-wise occurrence probabilities for each plant species are calculated, which are then aggregated over the entire image to estimate the plant cover. Since the phenology predictions are based on the plants detected in the images, we utilize the predicted species-wise occurrence probabilities and predict a phenology-stage probability for each plant species in each location. Similar to the plant cover prediction module, we are then also able to aggregate these probabilities over the entire image to retrieve the final phenology values. These predicted values then represent percentages of the plant cover of each species.

### 3.2. Joint Prediction vs. Species-Wise Prediction

With modeling phenology, or, more specifically, the phenological stages within neural networks, we have two possibilities: a joint prediction of the phenological stage independent of species with weight sharing ("joint prediction") and individual predictions for each individual species without weight sharing ("species-wise"). Both approaches have certain advantages and disadvantages. As for species-wise phenology prediction, the prediction is done for each species individually, i.e., the network learns a different model for each combination of phenological stage and species. This results in a specialized network that can potentially make high-accuracy predictions for some species. With this approach, however, species with little training

data are significantly disadvantaged since there is likely no way for the model to learn a sophisticated and general phenology model for these. This issue could potentially be resolved with the joint prediction approach, where only the occurrence probabilities and a single probability for each phenological stage are calculated. These probabilities can then be combined, for example, by simple multiplication, to obtain species- and phenology-wise probabilities. The general idea behind this approach is that often particular features are shared between the phenological stages and do not differ strongly between the different plant species. For example, for identifying the *flowering* stage, the model would use blossoms, which can look very similar between several plant species. For *senescence*, the model would likely focus on the brown color of plant leaves, which is also a property shared between the different species. If features are shared like this, there is also a higher chance for species with low amounts of training data to be predicted better since the data from other species can also be used for training. However, a too general model could also perform worse since plants with large amounts of training data might not be predicted as well as with individual phenology models.

### 3.3. Phenology Prediction Models

We investigate different calculation model variations to be able to compare how the structure of the phenology prediction module affects the results.

#### 3.3.1 Baselines

We include three simple baselines in our comparison: a constant prediction, a time-dependent constant prediction, and a naive prediction.

**Constant Prediction Baseline.** This baseline simply predicts the mean *flowering* and *senescence* percentages for every species in the dataset for every image.

**Temporal Mean Constant Prediction Baseline.** This baseline predicts the mean *flowering* and *senescence* percentages for every species depending on the week the respective image was taken. Hence, it is similar to the constant prediction baseline and does not use any image information. It should be noted that the temporal aspect is not taken into account by our phenology prediction approaches, which only predict the phenology values based on images. Therefore, this baseline uses different information than our image-based prediction models, resulting in an unfair comparison. However, due to the lack of simple alternative baselines, we use this one as a point of comparison.

**Naive Phenology Prediction.** This baseline model uses the segmentation-pre-trained CNN and predicts the phenol-

ogy from images. It is trained with the phenology annotations in the dataset but, in contrast to the more complex calculation models (section 3.3.2), uses no information determined by the plant cover prediction module. I.e., this model predicts a value

$$\text{Phenology}_{p,s}(x, y) \quad (1)$$

for each phenological stage  $s$  and each plant species  $p$  at each location  $x, y$ , completely independent of the already predicted species-wise occurrence probabilities. As in the following models, the final phenology percentage is obtained by averaging the location-wise probabilities over all predicted species-wise occurrence probabilities in the entire image.

It should be noted that since this model does not utilize any occurrence probabilities calculated during plant cover prediction, it also needs to learn to distinguish the different plant species from the sparse phenology training data.

#### 3.3.2 Combined Approaches

**Species-wise Prediction.** As mentioned above, with a species-wise model, we predict a phenology probability for each stage and species, and with the inclusion of the occurrence probabilities  $P_{p,\cdot}$  from the plant cover prediction module, we calculate

$$\text{Phenology}_{p,s}^{\text{corrected}}(x, y) = P_{p,\cdot}(x, y) \cdot \text{Phenology}_{p,s}(x, y) \quad (2)$$

**Joint Prediction.** For the joint prediction, only a single value is calculated for each phenological stage, which is then combined with the species-wise occurrence probabilities to compute the species- and stage-wise phenology probabilities:

$$\text{Phenology}_{p,s}^{\text{corrected}}(x, y) = P_{p,\cdot}(x, y) \cdot \text{Phenology}_{\cdot,s}(x, y) \quad (3)$$

**Improved Joint Prediction.** To try to combine the advantages of the two methods above, we also investigate an improvement of the joint prediction module. To improve this module, we add an adaptation block, which comprises two simple  $1 \times 1$  convolutional layers and predicts an additional weight  $W_{p,s}(x, y)$  for each combination of plant and stage. This weight value is then added to the input features of the stage-wise phenology probabilities  $\text{Phenology}_{\cdot,s}(x, y)$ . This way, the network can utilize the combined advantages of the weight and information sharing aspect of the joint prediction module for the different stages but can also learn more specific features for each plant species, similar to the species-wise one.

Table 1: The number of images containing phenology data of any non-zero percentage for the respective plant species.

Species	Images containing <i>Flowering</i>	Images containing <i>Senescence</i>
<i>A. millefolium</i>	0	107
<i>C. jacea</i>	25	201
<i>L. corniculatus</i>	173	174
<i>M. lupulina</i>	174	247
<i>P. lanceolata</i>	133	205
<i>S. autumnalis</i>	158	116
<i>T. pratense</i>	303	413
<b>Total</b>	682	682

## 4. Experiments

### 4.1. Dataset

This work uses the InsectArmageddon dataset introduced in [20] and [12]. The dataset contains plant cover data of 8 herbaceous plant species in 682 images labeled by a single ecologist. In addition to the plant cover, we utilize phenology estimates not used in previous works, which is available for seven plants from the dataset. They were also estimated by the same ecologist and represent the percentage of the identified plant individuals that are currently in a particular phenological stage, i.e., *flowering* or *senescent*. Example images with plants from both phenological stages are shown in Figure 1. It should be noted that the phenological stages are not mutually exclusive. That is, a plant can be flowering and senescent at the same time.

In Figure 2, the distribution of the plant cover percentages and the phenology percentages is shown. Notably, since the phenology only represents a fraction of the number of plant individuals identified in the images, the amount of training data for phenology prediction is significantly smaller than for cover prediction. In Figure 2, we can see that *Trifolium pratense* has an average cover value of 35% over the entire dataset, i.e., approximately 35% of all pixels in the whole dataset contain *Trifolium pratense*. However, only about 10% in total are flowering parts, and less than 5% are senescent, resulting in a significant difference in training data for this plant species. This difference is even more extreme in the other plant species.

Moreover, since high outlier values possibly skew the average phenology percentages, we also take a look at the number of images containing *any* kind of plant cover percentage other than 0 for a certain plant species, as shown in Table 1. We can see that, for *Trifolium pratense*, only about half of the images contain any data on *flowering*, often significantly below 100% that would match the entire plant cover. For *Achillea millefolium*, there is not a single image with flowering individuals, making this kind of pre-

diction entirely impossible for this species. We can also see that the number of images containing phenology percentage labels is generally higher for *senescence*. Under consideration of these statistics, we would expect the prediction of *senescence* to yield better results.

### 4.2. Metrics

We evaluate the plant cover and concurrent phenology predictions using different metrics.

**Plant Cover.** For the prediction of plant cover, we utilize the so-called DCA-Procrustes-Correlation (DPC). This metric involves performing a Detrended Correspondence Analysis (DCA) [5] on the target and predicted outputs. The resulting values are then compared using a Procrustes test, yielding a correlation value. Higher correlation values indicate greater similarity between the distributions, which holds significance in ecological applications.

**Phenology.** We utilize the Pearson correlation coefficient to evaluate the phenological results. We compare the correlation of the predicted values for each species over all predicted data points and calculate the mean over all species at the end. This results in two correlation values, one for each phenological stage, reflecting the performance for the different types of phenology prediction.

Since models with several metrics are hard to compare, we also look at two compound metrics that should better reflect the entire model’s performance. The first one is  $\text{Corr}_{\text{mean}}$ , i.e., the mean of the correlation values of both phenological stages:

$$\text{Corr}_{\text{mean}} = 0.5 \text{Corr}_{\text{flowering}} + 0.5 \text{Corr}_{\text{senescence}} , \quad (4)$$

with  $t$  and  $p$  being the true and predicted values.

The second one is the geometric mean of the three aforementioned correlation metrics:

$$\text{Compound} = \sqrt[3]{DPC \cdot \text{Corr}_{\text{flowering}} \cdot \text{Corr}_{\text{senescence}}} . \quad (5)$$

As it combines the plant cover prediction metric and both phenology metrics, it should most accurately reflect the model’s performance in general without focusing too much on single aspects of the model.  $\text{Corr}_{\text{mean}}$ , in contrast, merely focuses on the model’s ability to predict the phenological stages without taking its ability for plant cover prediction into account.

Since the correlation of the constant baseline is not defined, we also compare the mean squared error of the predicted flowering values:  $\text{MSE}_{\text{flowering}}$  and  $\text{MSE}_{\text{senescence}}$ . These values demonstrate how big the difference between the predicted and target percentages is, which is not reflected in the correlation values.



Figure 1: Left: An example of an image from the dataset that contains several flowering plants. Right: An example of an image from the dataset that contains a large number of senescent plants.

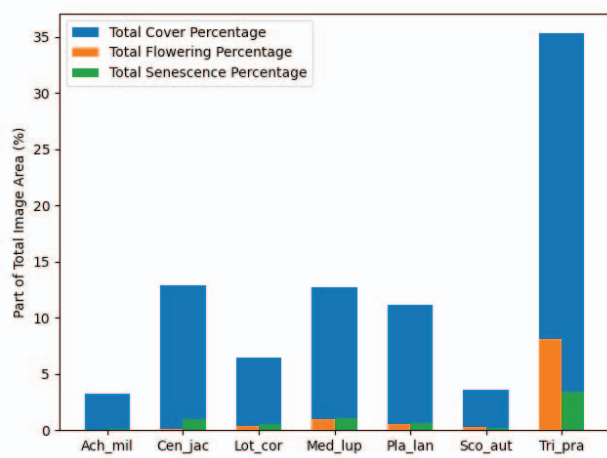


Figure 2: The distribution of plant cover, flowering and senescence percentages over the dataset.

### 4.3. Setup

For our experiments, we use the same setup as in [10]. I.e., we utilize a ResNet50 [3] with a feature pyramid network (FPN) [13]. We first pre-train the network on GBIF [21] image data using the segmentation pre-training algorithm [9], and also use the same hyperparameters as in [10]. We use an image resolution of  $1536 \times 768$  and investigate two different training setups. The first setup utilizes the weekly labels contained in the dataset, resulting in 682 images with labels. The second setup we investigate uses interpolated daily labels. Following the label interpolation approach from [10], the weekly labels are interpolated down to the daily level using linear interpolation. This is done for the plant cover and phenology labels. With this method, also the unlabeled intermediate daily images can be used for plant cover and phenology training, resulting in a about 7 times the number of images available for training. To investigate the effect of the training duration on the quality

of the results, we analyze varying epoch numbers. For the weekly labels, we train for 3, 10, 25 and 60 epochs. Since the number of iterations per epoch is significantly higher for the daily labels, in this setup we train only for 3, 10 and 25 epochs.

The plant cover training and phenology training are conducted in parallel, sharing the same base network and using a learning rate of  $10^{-5}$  for the plant cover prediction module and one of  $10^{-4}$  for the phenology prediction module, and the mean absolute error as loss for both modules. All experiments are conducted in a 12-fold cross-validation with three repetitions, as done in [10].

## 4.4. Experimental Results

### 4.4.1 Quantitative Results

In Figure 3, the experimental results for the three setups are shown. In general, we can see that the joint model and the naive model perform the worst for the phenology prediction for all setups. In contrast, the species-wise and improved joint model (“joint++”) perform significantly better. The reason for the bad performance of the naive model is likely insufficient training data. As mentioned before, since the naive model does not utilize any information from the plant cover prediction module, it also does not use any occurrence probabilities. Therefore, it not only needs to learn to recognize the phenological stages of the plants but also to recognize the species themselves. It has to learn this from significantly less training data than the combined models, which implicitly use the plant cover training data and the phenology training data. From this, we can conclude that having a model which can include knowledge from plant cover data is vastly superior to naive models that do not do this.

As mentioned, the joint model also performs poorly, in most cases even worse than the naive model. The likely explanation for this is the nature of how the plant cover prediction module predicts the plant cover values. Since it simply predicts the occurrence probabilities of each plant,

Table 2: A numerical comparison of the top results of each approach.  $\text{Corr}_{\text{flowering}}$ ,  $\text{Corr}_{\text{senescence}}$  and  $\text{Corr}_{\text{mean}}$  are the respective phenology correlation values, DPC is the DCA-Procrustes correlation metric and Compound refers to the compound metric that aggregates  $\text{Corr}_{\text{flowering}}$ ,  $\text{Corr}_{\text{senescence}}$  and DPC.  $\text{MSE}_{\text{flowering}}$  and  $\text{MSE}_{\text{senescence}}$  are the mean squared errors of the predicted phenology percentages and the reference estimates.

Method	$\text{Corr}_{\text{flowering}}$	$\text{Corr}_{\text{senescence}}$	$\text{Corr}_{\text{mean}}$	$\text{MSE}_{\text{flowering}}$	$\text{MSE}_{\text{senescence}}$	DPC	Compound
Constant	-	-	-	0.009	0.042	-	-
Temporal Mean	<b>0.544</b>	0.445	0.495	0.005	0.032	-	-
Naive	0.184	0.239	0.211	0.011	0.044	0.777	0.324
Species-wise	0.528	0.452	0.490	<b>0.003</b>	0.030	0.776	0.570
Joint	-0.072	-0.134	-0.103	0.012	0.050	0.764	0.194
Joint Improved	0.532	<b>0.461</b>	<b>0.496</b>	<b>0.003</b>	<b>0.030</b>	<b>0.780</b>	<b>0.576</b>

which are not mutually exclusive, several plants can have a high occurrence probability in the same location. If we then apply the joint prediction model, due to the simple multiplication, each plant with a high probability is assigned the same phenology probability, which leads to a potentially high *flowering* or *senescence* probability for several plants at the same location at the same time. However, usually, there are only a few or even only a single plant individual flowering in a single location, resulting in strong mispredictions for the joint model. Since the species-wise model predicts a separate probability for each species and phenological stage, it does not have this issue, resulting in a much better performance. However, since the original idea of sharing phenology information between different plant species still holds value, we developed the aforementioned improved joint calculation model, which combines the advantages of the joint and the species-wise model. Due to this, the improved joint model (shown as joint++ in Figure 3) outperforms the species-wise model in all kinds of phenology predictions. We also notice that the difference in performance of the improved joint model is generally slightly larger for *senescence*. The likely reason for this is that, while the looks of flowers can significantly differ from one plant species to the next, the features of senescent plants are usually similar: brown leaves. Therefore, information sharing between several species is more advantageous for *senescence* prediction.

Moreover, in Figure 3, we can also see that training with interpolated daily images is significantly better than training only with weekly images. We see an improvement of about 0.13 for  $\text{Corr}_{\text{flowering}}$  for the improved joint model and about 0.03 for  $\text{Corr}_{\text{senescence}}$ . We can also see that the DPC is generally better with this setup, achieving a correlation of about 0.8, in contrast to about 0.73 with weekly annotations only. In the compound metric, it is reflected that the model in the daily setup is significantly better than the one in the weekly one.

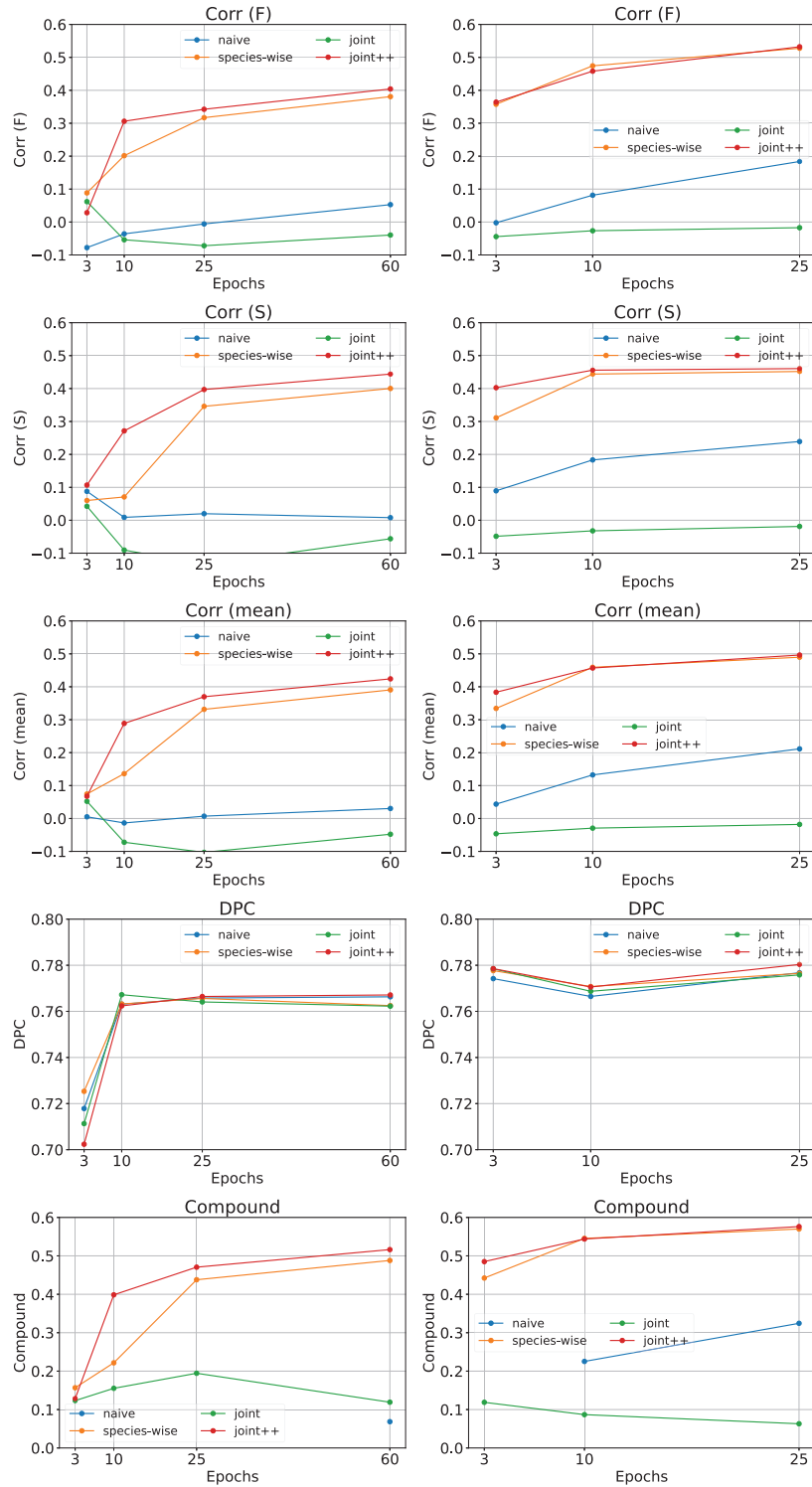
In Table 2, we see the numerical metric values of the best models selected based on our compound metric from each category, as well as the results of the baselines. Overall, it is

visible that the joint improved model outperforms all other models, most of them by a large margin. Only the temporal mean baseline yields a slightly better value in  $\text{Corr}_{\text{flowering}}$  but is strongly outperformed in terms of  $\text{MSE}_{\text{flowering}}$  and  $\text{MSE}_{\text{senescence}}$ . While the species-wise model also performs well, its performance is slightly worse than that of the improved model. Interestingly, the joint improved model also appears to utilize the phenology data for improving the plant cover prediction since the DPC value of the plant cover prediction is also slightly higher than the one of the other models. Lastly, the naive model is outperformed even by the constant and temporal mean baseline. From this, we can, again, conclude that predicting phenology without including additional information, e.g., from the cover prediction module, leads to drastically worse results.

#### 4.4.2 Qualitative Results

In Figure 4, we can see two examples of the predictions of phenology values on two sites of the InsectArmageddon dataset. For *flowering* in general, we can see that the trend over time is captured quite well by our model. While the amplitudes, e.g., the peak of flowering, often do not match, the general upward and downward trends are clearly reflected in the predictions. Moreover, the amplitudes' temporal location usually appears to match or lie only one week apart from the reference estimates. The best-performing species appears to be *Trifolium pratense* (Tri\_pra), which is also the most dominant species in the dataset and, therefore, has the most training data. Its prediction usually matches the reference estimates best. In contrast, plants like *Scorzoneroideis autumnalis* (Sco\_aut) are predicted the worst, likely also due to small amounts of training data in the dataset.

For *senescence*, there are often slight mispredictions by the model in the earlier weeks; however, in the later weeks, the time trend is usually captured quite well, similar to the *flowering* stage. However, here we can also see that, in some cases, the model predicts *senescence* significantly earlier for several species. This effect might be caused



(a) Weekly images and labels

(b) Daily images and labels

Figure 3: The experimental results for the different metrics in different setups.

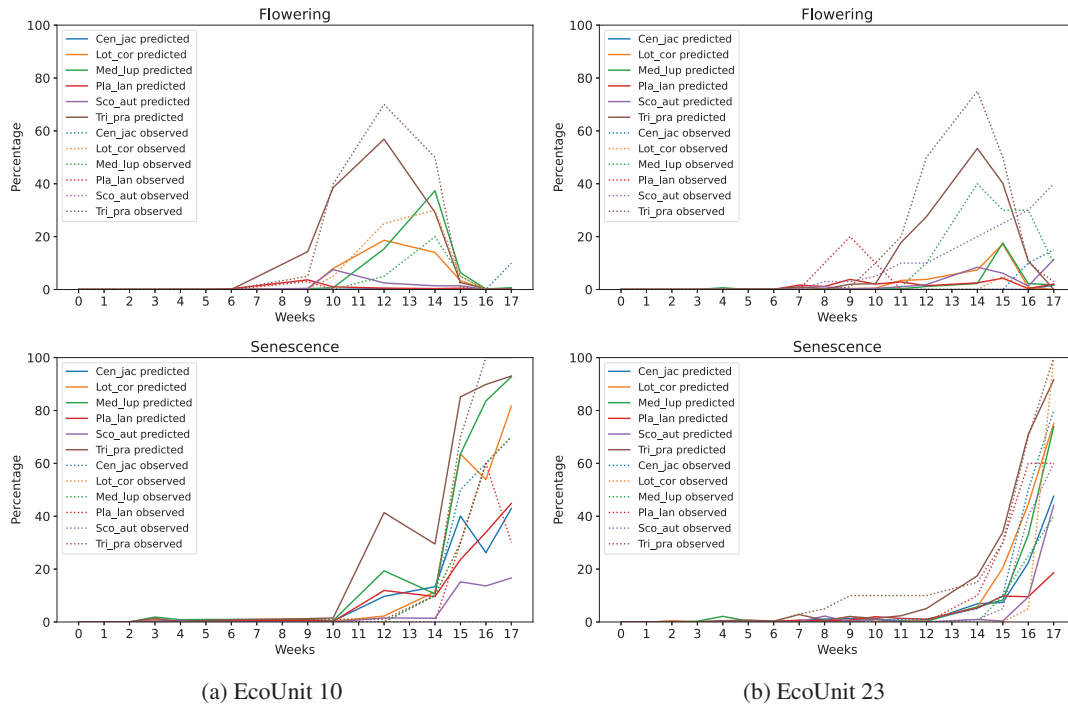


Figure 4: Qualitative prediction results for *flowering* and *senescence* over the time for two sites of the InsectArmageddon dataset.

by inconsistencies in the images of the InsectArmageddon dataset, which show human intervention in some images, affecting the looks of the ecosystem in the images.

Overall, the phenological trends are well-captured by our model and can prove helpful in future ecological studies.

## 5. Conclusion

In this work, we have presented the novel task of automatic phenology prediction for plant communities. In addition to simple prediction methods, we demonstrated that we can combine this model with plant cover prediction to significantly boost the results for phenology prediction. Especially since this kind of data is usually collected together in the field by ecologists, data for both is usually available, enabling this kind of combined approach. We have also seen that weight sharing in the phenology prediction module can improve results since it implicitly shares training data between classes, from which especially plant species benefit that are not abundant in the underlying vegetation dataset and, thus, also have only little training data available. Since this dataset merely contained a small number of plant species, this effect is likely stronger on datasets with more and diverse plant species, which is subject to future investigations.

The approach shown here can be of significant help to ecologists since it can not only extract data on the plant

community composition from images but also on their phenology, enabling high-quality analyses of environmental effects with high temporal resolution. Especially since ecological studies with a high temporal resolution have been rare to nonexistent until now, this approach can help acquire significant novel insights in future ecological studies.

For future work, we aim to include more intuitive knowledge into the network, for example, a color separation via simple thresholding in the color space to separate the green leaves from the flowers or the senescent plants, which would counter the small amounts of training data by directly integrating the knowledge. Moreover, the integration of the temporal aspect of the dataset would likely also improve the results since the phenological stages usually have a specific time they occur. Therefore, a deeper analysis of the time aspect is also part of our future work.

## Acknowledgements

Matthias Körschens thanks the Carl Zeiss Foundation for the financial support. We thank Alban Gebler for enabling the image collection process in the iDiv EcoTron and Josephine Ulrich for the data collection. We acknowledge funding from the German Research Foundation (DFG) via the German Centre for Integrative Biodiversity research (iDiv) Halle-Jena-Leipzig (FZT 118) for the support of the FlexPool project PhenEye (09159751).



## References

- [1] Mengying Cao and Qinchuan Xin. A deep learning method for detecting leaf phenology from phenocam imagery. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 6889–6892. IEEE, 2021. 1
- [2] Katharina Gerstner, Carsten F Dormann, Anke Stein, Ameer M Manceur, and Ralf Seppelt. Editor’s choice: Review: Effects of land use on plant diversity—a global meta-analysis. *Journal of Applied Ecology*, 51(6):1690–1700, 2014. 1
- [3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016. 5
- [4] Juliane Helm, Thierry Dutoit, Arne Saatkamp, Solveig F Bucher, Matthias Leiterer, and Christine Römermann. Recovery of mediterranean steppe vegetation after cultivation: Legacy effects on plant composition, soil properties and functional traits. *Applied Vegetation Science*, 22(1):71–84, 2019. 1
- [5] Mark O Hill and Hugh G Gauch. Detrended correspondence analysis: an improved ordination technique. In *Classification and ordination*, pages 47–58. Springer, 1980. 4
- [6] Zhaojin Huang, Lichao Huang, Yongchao Gong, Chang Huang, and Xinggong Wang. Mask scoring r-cnn. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 6409–6418, 2019. 2
- [7] Patrizia König, Susanne Tautenhahn, J Hans C Cornelissen, Jens Kattge, Gerhard Bönisch, and Christine Römermann. Advances in flowering phenology across the northern hemisphere are explained by functional traits. *Global Ecology and Biogeography*, 27(3):310–321, 2018. 1
- [8] Matthias Körschens, Paul Bodesheim, Christine Römermann, Solveig Franziska Bucher, Mirco Migliavacca, Josephine Ulrich, and Joachim Denzler. Automatic plant cover estimation with convolutional neural networks. In *Computer Science for Biodiversity Workshop (CS4Biodiversity), INFORMATIK 2021*, pages 499–516, 2021. 1, 2
- [9] Matthias Körschens, Paul Bodesheim, Christine Römermann, Solveig Franziska Bucher, Mirco Migliavacca, Josephine Ulrich, and Joachim Denzler. Weakly supervised segmentation pretraining for plant cover prediction. In *DAGM German Conference on Pattern Recognition*, pages 589–603. Springer, 2021. 1, 2, 5
- [10] Matthias Körschens, Solveig Franziska Bucher, Christine Römermann, and Joachim Denzler. Improving data efficiency for plant cover prediction with label interpolation and monte-carlo cropping, 2023. 5
- [11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012. 2
- [12] Matthias Körschens, Paul Bodesheim, Christine Römermann, Solveig Franziska Bucher, Josephine Ulrich, and Joachim Denzler. Towards confirmable automated plant cover determination. In *ECCV Workshop on Computer Vision Problems in Plant Phenotyping (CVPPP)*, 2020. 1, 2, 4
- [13] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2117–2125, 2017. 5
- [14] Huiying Liu, Zhaorong Mi, LI Lin, Yonghui Wang, Zhenhua Zhang, Fawei Zhang, Hao Wang, Lingli Liu, Biao Zhu, Guangmin Cao, et al. Shifting plant species composition in response to climate change stabilizes grassland primary production. *Proceedings of the National Academy of Sciences*, 115(16):4051–4056, 2018. 1
- [15] Jose Raul Romo-Leon, Willem JD van Leeuwen, and Alejandro Castellanos-Villegas. Land use and environmental variability impacts on the phenology of arid agro-ecosystems. *Environmental management*, 57:283–297, 2016. 1
- [16] Terry L Root, Jeff T Price, Kimberly R Hall, Stephen H Schneider, Cynthia Rosenzweig, and J Alan Pounds. Fingerprints of global warming on wild animals and plants. *Nature*, 421(6918):57–60, 2003. 1
- [17] Cynthia Rosenzweig, Gino Casassa, David J Karoly, et al. Assessment of observed changes and responses in natural and managed systems. *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 79–131, 2007. 1
- [18] Lara Souza, Tamara Jane Zelikova, and Nathan J Sanders. Bottom-up and top-down effects on plant communities: nutrients limit productivity, but insects determine diversity and composition. *Oikos*, 125(4):566–575, 2016. 1
- [19] Abdelaziz Triki, Bassem Bouaziz, Jitendra Gaikwad, and Walid Mahdi. A deep learning-based approach for segmenting and counting reproductive organs from digitized herbarium specimen images using refined mask scoring r-cnn. In *TACC*, 2021. 1, 2
- [20] Josephine Ulrich, Solveig Franziska Bucher, Nico Eisenhauer, Anja Schmidt, Manfred Türke, Alban Gebler, Kathryn Barry, Markus Lange, and Christine Römermann. Invertebrate decline leads to shifts in plant species abundance and phenology. *Frontiers in plant science*, 11:1410, 2020. 1, 2, 4
- [21] [GBIF.org](https://www.gbif.org). Gbif home page, 2022. <https://www.gbif.org>, Accessed 9 June 2022. 5
- [22] Xu Annie Wang, Julie Tang, and Mark Whitty. Deepphenology: Estimation of apple flower phenology distributions based on deep learning. *Computers and Electronics in Agriculture*, 185:106123, 2021. 1, 2
- [23] Hulya Yalcin. Phenology recognition using deep learning: Deeppheno. In *2018 26th Signal Processing and Communications Applications Conference (SIU)*, pages 1–4. IEEE, 2018. 1, 2