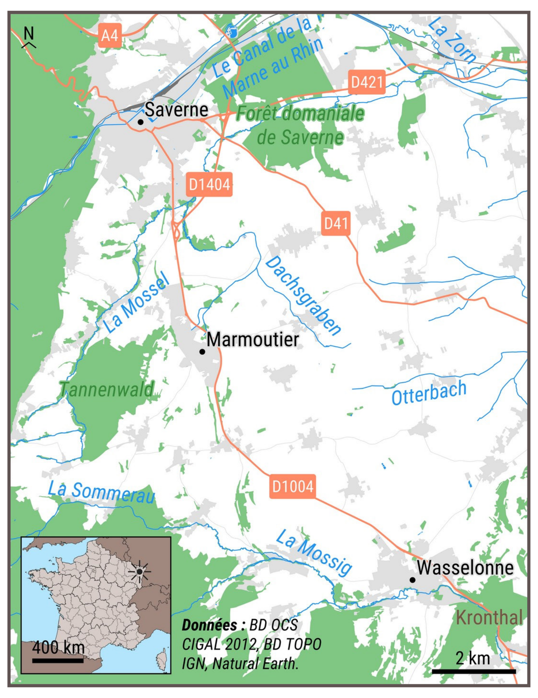
Unprecedented and ancient spatial data for assessing the effect of past changes on current floristic biodiversity

Q. Poterek[1], P.-A. Herrault[1], D. Schwartz[1], G. Forestier[2]

1 LIVE, CNRS UMR 7362, University of Strasbourg, 67083 Strasbourg, France
2 IRIMAS, University of Haute-Alsace, 68093 Mulhouse, France

Laboratoire image, ville, environnement | LIVE
Université de Strasbourg

Background



Within the last decades, changes in climate and land cover have been the driving forces behind a global decline in biodiversity. Understanding these phenomena is crucial for implementing proper conservation measures (Kleijn *et al.*, 2011). However, various studies have shown the importance of environment history for assessing current biodiversity. Indeed, many organisms experience time lags when responding to specific changes (Kuussaari *et al.*, 2009). Thus, studying past states of a landscape, at proper spatial and temporal scales, is valuable in the understanding of bioversity patterns.

In this context, century-old archive aerial photographs appear to be of incredible value, due to their scale and reasonable temporal resolution. However, with a limited spectral resolution, deterioration and irregular specifications, their processing has proved to be rather troublesome. In this thesis, we first propose techniques for restoring and assembling these photographs to produce time series of proper quality. Secondly, we explore deep learning algorithms for extracting environmental indices and land cover from this archive data set. Finally, we create species-habitat model, based on these spatial products and floristic data that were collected during field missions carried out in the study area (Figure 1).

Figure 1: Proposed study area, located in the French Bas-Rhin department.

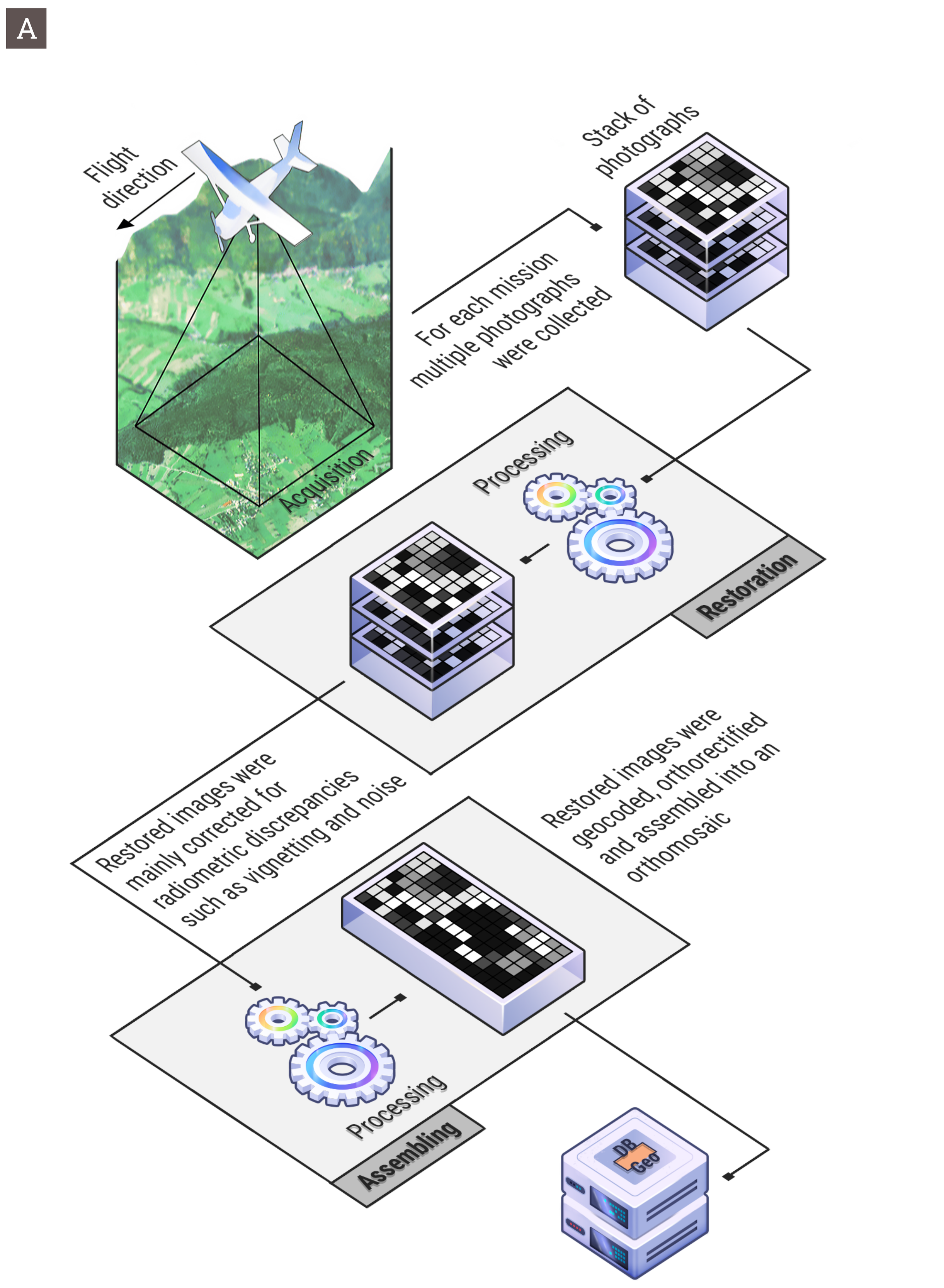
Associated publications

Peer-reviewed journals

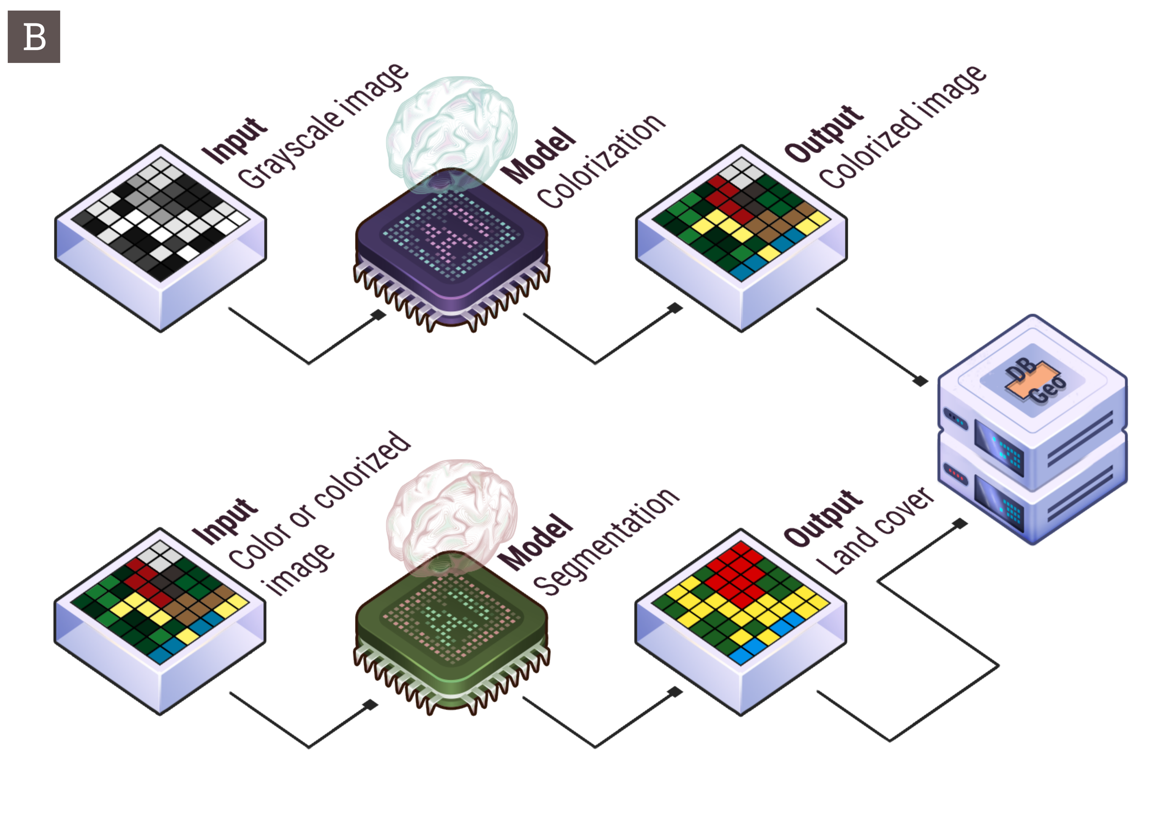
[1] Q. Poterek, P.-A. Herrault, G. Skupinski, and D. Sheeren (2020). **Deep Learning for Automatic Colorization of Legacy Grayscale Aerial Photographs**, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, doi: 10.1109/jstars.2020.2992082.

[2] Q. Poterek, P.-A. Herrault, G. Forestier, and D. Schwartz (2020). **Revealing Long-term Physiological Trajectories of grasslands from Legacy B&W Aerial Photographs**, *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, V-3-2020, 549–557, doi: 10.5194/isprs-annals-V-3-2020-549-2020.

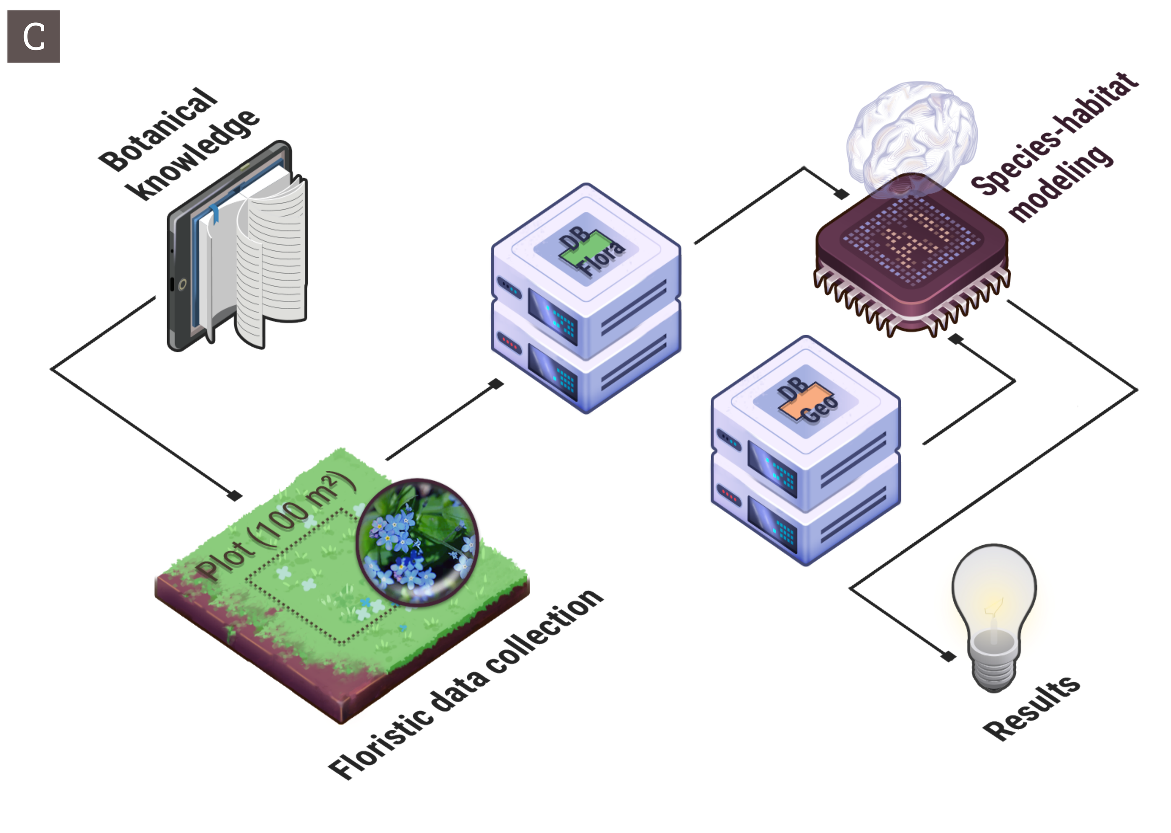
A



B



C



Spatiotemporal archive data preparation

In preparation for this retrospective analysis, we first started by building a substantial geographic database for the study area. It is composed of land cover layers, elevation data, topographic data, etc. Other data were also digitized, such as old maps dating back to 1760, in order to provide better spatial and temporal hindsight. Finally, this database was futher complemented with archive aerial photographs, provided online by the French National Institute of Geographic and Forest Information. The aerial missions, during which photos were taken, are presented in Figure 2.

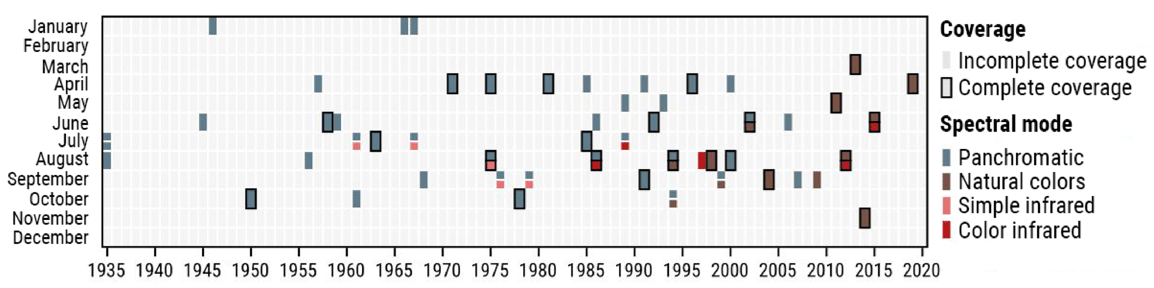


Figure 2 : Temporal distribution of the available aerial missions for the study area. Coverage and spectral mode are also specified, as they were used for selecting proper missions.

Among these aerial missions, those with a complete coverage of the study area were kept. Only 12 missions required further processing, from 1950 to 1996, as they were either panchromatic or unassembled. After compiling metadata for each mission, photographs were first cleaned and underwent radiometric correction for various deteriorations, such as vignetting and noise (Figure 3). To that end, correction maps were modelled with a parametric polynomial curve. Finally, all photos were geocoded, orthorectified and assembled into a panchromatic orthomosaic for each mission.

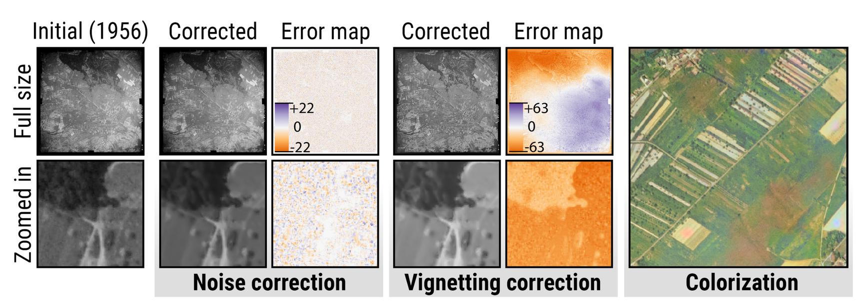


Figure 3 : Example of a single photograph for the proposed pipeline for radiometric correction and enhancement. *Colorization is explained in the next section.*

Spatiotemporal data enhancement

After preparing a time series of orthophotographs dating back to 1950, a major challenge remained before assessing the relationship between environment history and current floristic biodiversity in the study area. Indeed, panchromatic photos lack radiometric information, required for extracting land cover. Moreover, these data cannot be used to compute environmental indices. However, deep learning has helped bypass such limitation. Indeed, it can be used to solve difficult or ill-posed problems. In this case, the enhancement of panchromatic photographs by predicting radiometric information.

To that end, we started by developing a colorization model, based on the conditional generative adversarial network (cGAN) framework (Mirza & Osindero, 2014). We trained it with current color aerial photos and their grayscale counterparts, corrupted in order to mimick the look of legacy data. The model learned the mapping between grayscale and color by searching for similar spatial semantics between input and output. During inference, we then only had to pass an archive orthophoto, and the model would predict the corresponding colors channels, namely blue, green, red and near infrared.

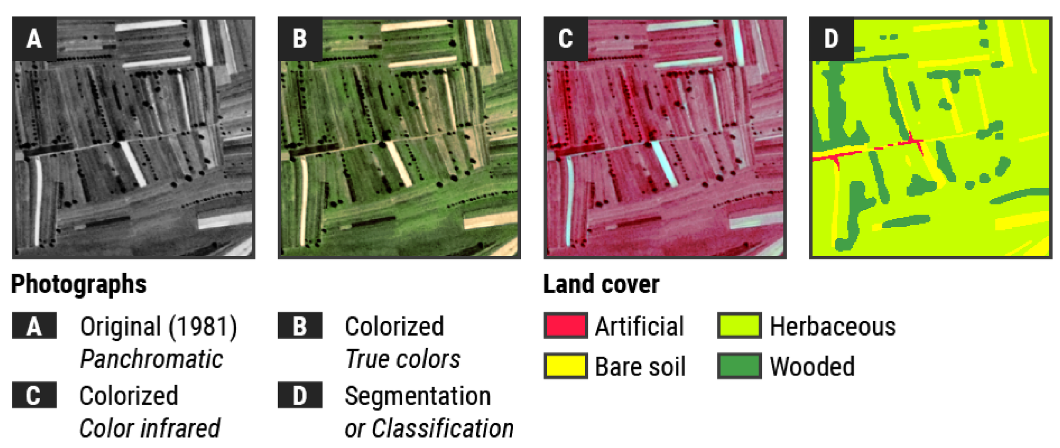


Figure 4 : Example of colorization and semantic segmentation results.

A convolutional neural network (CNN) was then trained to perform semantic segmentation on color and colorized photographs (Kemker *et al.*, 2018). A total of 5 land cover classes were predicted, namely artificial and water surfaces, bare soils, herbaceous and wooded vegetation. Examples of results for both colorization and segmentation are presented in Figure 4.

Land use history and biodiversity

The last step of this work involves the development of species-habitat models, based on local floristic data and the spatiotemporal series.

Floristic data have yet to be collected. Field missions will be organized in Spring 2021, during which the specific richness of 60 grasslands in the study area will be measured. Data collection will take place in a 10 × 10 m plot for each grassland, representative of its flora. Sample sites have been selected at random, following a stratified sampling scheme. Various features were used to ensure diversity in the grasslands, namely land cover composition and configuration measured at different scales, along with parcel history (e.g. land cover or land use change, fragmentation). After extensive testing, composition and configuration were summarized by two landscape metrics, MSIDI and mean CAI respectively (Figure 5).

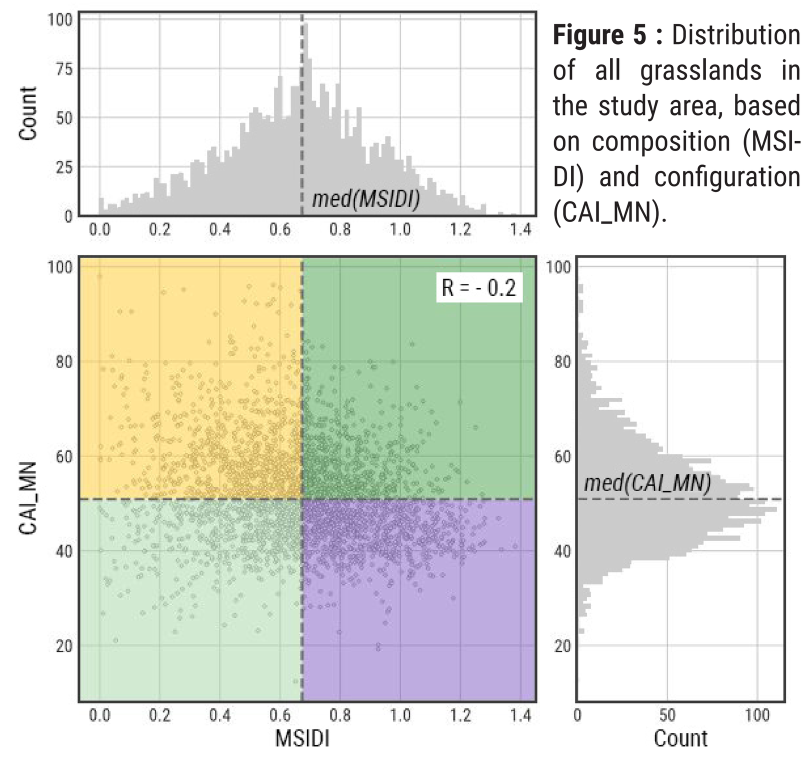


Figure 5 : Distribution of all grasslands in the study area, based on composition (MSIDI) and configuration (CAI_MN).

After floristic data collection, the last step will be to create said species-habitat models, most likely with generalized linear models (GLMs) based on a negative binomial or Poisson distribution.

References

Kemker, R., Salvaggio, C., & Kanan, C. (2018). Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. ISPRS journal of photogrammetry and remote sensing, 145, 60-77.

Kleijn, D., Rundlöf, M., Scheper, J., Smith, H. G., & Tscharntke, T. (2011). Does conservation on farmland contribute to halting the biodiversity decline?. Trends in ecology & evolution, 26(9), 474-481.

Kuussaari, M., Bommarco, R., Heikkinen, R. K., Helm, A., Krauss, J., Lindborg, R., ... & Steffan-Dewenter, I. (2009). Extinction debt: a challenge for biodiversity conservation. Trends in ecology & evolution, 24(10), 564-571.

Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784.

Preliminary discussion

So far, the proposed methodology has shown varying levels of success, as highlighted in the associated publications. Indeed, we are able to reconstruct temporal trajectories for different kinds of surfaces, compute environmentals indices and extract past land covers with better results than with a panchromatic mosaic and ancillary features.

However, due to the nature of GANs, colorization results may vary substantially from one photograph to another, due to differences in spatial, semantics, season, scale, etc. The effets of such variation have yet to be quantified.

Moreover, due to the retrospective nature of these works, no proper data are available for validating some of our results, namely the quality of colorized orthophotos and environmental indices. Thus, special precautions will have to be taken when interpreting and discussing modeling results.