



**HAL**  
open science

## Super-resolution by fusing multi-spectral and terrain models: application to water level mapping

Emilien Alvarez-Vanhard, Fernandez-Garcia Guglielmo, Thomas Corpetti

### ► To cite this version:

Emilien Alvarez-Vanhard, Fernandez-Garcia Guglielmo, Thomas Corpetti. Super-resolution by fusing multi-spectral and terrain models: application to water level mapping. *IEEE Geoscience and Remote Sensing Letters*, 2023, 20, pp.1-5. 10.1109/LGRS.2023.3319548 . hal-04310773

**HAL Id: hal-04310773**

**<https://hal.science/hal-04310773v1>**

Submitted on 27 Nov 2023

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

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

# Super-resolution by fusing multi-spectral and terrain models: application to water level mapping

Alvarez-Vanhard Emilien, Fernandez-Garcia Guglielmo, Corpetti Thomas

**Abstract**—Recently, deep convolutional networks have made great progress on the task of super resolution, i.e. reconstructing images with finer spatial resolution. However, although the reconstructions are visually impressive, they may lack physical consistency. This aspect is sought in remote sensing, where the resolution of satellite imagery (e.g. Sentinel-2) may be too coarse to characterize the physical structure and dynamics of certain landscapes. Through the study of flooding dynamics in wet grasslands, we propose a super resolution approach that allows deriving fine resolution patterns that are visually realistic and physically exploitable. This approach is based on an architecture, Fusion-UNet, allowing the fusion of multispectral data with a digital terrain model (DTM) associated with a loss function combining content, structure and segmentation losses. Our results show that this model can precisely predict water levels while restituting the fine structure of the landscape. This approach allows to refine the production of hydrological and ecological indicators to define the state of the ecosystem.

**Index Terms**—Super resolution, fusion networks, Sentinel-2, Digital Terrain Models, flood dynamics, wet grasslands

## I. INTRODUCTION

Monitoring the water levels in wet grasslands is of prime importance to recover important ecosystem functions and to help managers to better manage their resources. Indeed, water dynamics of these areas lead ecological, hydrological, physical, biogeochemical, climatic and agronomic functions [1]–[3]. However, due to their spatiotemporal variability, flood measurements (water extent, level and residence time) remain a challenge. To this end, Sentinel-2 (S2) data can be an appealing solution.

The arrival of S2 data has indeed caused significant changes in the monitoring of land use/land cover [4]. Thanks to its good compromise between resolution on spatial (from 60 m to 10 m depending on spectral bands), temporal (one revisit every  $5 \approx 6$  days) and spectral (12 bands) axes, many new use cases have been explored in a wide range of applications related either to agriculture monitoring [5], imperviousness estimation [6], vegetation [7], or water quality for example [8]. S2 has also revealed to be an appealing complementary source to other satellite (Sentinel-1 radar, higher spatial resolution as WorldView-3 satellite with 1.24m) for data fusion [9], [10] as for example in [11] for soil moisture analysis.

Sentinel-2 features make such data widely used, in particular for the monitoring of dynamics, as floods, through the production of temporal series revealing fine patterns as seasonal and monthly intra-variability. However, in some applications, the 10 m spatial resolution may be limited. This is more generally true in the field of ecology where important variables generally belong to higher resolutions [12]

such as the biological effects of irregular flooding due to small scale topographic features (e.g. ditches). The particular example of flood dynamics monitoring in wet grasslands is relevant because these areas are composed of fine spatial patterns (below 10 m spatial resolution) that reflect habitat structure and ecosystem fragmentation [1], [13]. Moreover, other physical fine-scaled variables play important roles in the characterization of wetlands, such as water level or volume, that have to be evaluated together with the spatial patterns.

The recent progress of machine learning methods, especially with convolutional neural networks (CNN), associated with the availability of high-resolution data make now possible the reconstruction of high-resolution images using S2 data. This task, called *super-resolution* (SR), consists in learning an upscaling model to improve the spatial resolution of S2 data. In practice, from pairs of S2 images and corresponding high-resolution, the idea is to learn a model able to improve the initial S2 resolution and authorize an analysis of spatial patterns at a finer scale.

Today, some techniques give visually impressive results using adversarial losses [14]. However, despite being visually plausible, super-resolved images remain not exploitable in terms of physical consistency. The use of Generative Adversarial Networks (GANs) indeed generates patterns which, although realistic, do not correspond to any reality on the field, preventing practical applications related, for example, to vegetation monitoring, urban growth analysis, estimation of flooded areas, etc [15].

In this paper, we tackle the issue of estimating high resolution water level (WL) images from S2 images. To this end, we exploit the fact that flooded areas are connected to local topology and we propose an original neural network that fuses S2 images with Digital Terrain Models (DTM) to estimate water level images with high spatial resolution. Furthermore, to eliminate the use of adversarial losses likely to generate realistic but physically inconsistent patterns, we suggest relying on structural/content losses to derive visually realistic and physically exploitable patterns.

## II. RELATED WORKS

Based on the fact that the topography largely controls the spatial distribution of flooded areas, most of the super-resolution of water level mapping are today performed by mixing a coarse resolution water surface information (e.g. satellite observations) with finer resolution data, usually issued from DTM. For example, the studies in [16]–[18] propose optimal fittings of the position of flood mask profiles according to

the underlying terrain information to extract higher resolution flooded information. Such approaches are straightforward but suffer from some drawbacks : 1) sensitivity to classification errors, particularly due to the coarse resolution ; and 2) difficulty to detect a reliable boundary between flooded and non-flooded areas.

On the computer vision side, with the democratization of CNN architectures in the last decade, several techniques have emerged to design relevant reconstruction of high-resolution data from low-resolution ones (see for example [19], [20]) or to improve the resolution of some spectral bands [21], [22]. By definition, the super-resolution task is subject to two major problems: i) ill-posedness: several high-resolution images can give the same low-resolution one and ii) accuracy analysis: the definition of an objective criterion evaluating the quality of an estimated super-resolution image is tricky (usual image-comparison metrics do not discriminate details and high frequencies). Among the various CNN architectures, GANs have brought interesting solutions to these two issues [23]. In particular, they have partly solved the problem of finding an objective criterion to optimize: usual loss functions based on image comparisons are replaced by an adversarial task where a generative network is trained to generate fake data with realistic features and a discriminative network is trained to recognize whether the input is real or fake data. The first network tries to fool the second network and an increase of the error rate of the latter means that it becomes increasingly difficult to differentiate false data from real data, until the generated data is sufficiently realistic. GAN based super-resolution approaches are the current state of the art for SR tasks by directly learning an end-to-end mapping between low-resolution and high-resolution images.

Although for everyday life image analysis applications, super-resolution images produced by GAN-based networks are of excellent quality, when focusing on details related to the physics of the reconstructed objects, one observes errors or a lack of physical consistency. For example, in the case of water level mapping, GANs do not ensure that topographical features (e.g. ditches, riverbanks, depressions, etc.) and their geographical location are respected. This is due to the adversarial losses that are likely to introduce visually realistic images, but physical consistency is not ensured. This is the reason why, in addition to adversarial losses, some studies have proposed to use perceptual losses based for examples of features from Visual Geometry Group (VGG) neural network [24], [25] or on contours [26].

In this study, we suggest to super-resolve S2 images to recover flooded areas by i) efficiently combining the DTM (to rely as much as possible on the physics of wet grasslands) to S2 and ii) exploiting various loss functions to reconstruct realistic patterns, without the use of GANs.

### III. MATERIALS AND METHODS

#### A. Data Description

We address the super resolution’s problem for water-level estimation in the Sougéal marsh (175 ha), part of the river basin of the Couesnon valley (France, 48.52 N, 1.53 W,

Brittany region). The flooding dynamics of wet grasslands are characterised by high spatiotemporal variability, due to the heterogeneity of the topography, hydraulic conductivity of soils and hydrological discontinuities [2]. Therefore marshes are usually divided in hydrogeomorphological units, i.e. portions of space whose hydrological response is homogeneous with respect to external stresses. This work used three piezometers for measuring WL in three different units through the study site.

Atmospherically corrected [27] cloud-free S2 images were downloaded for 74 dates between 01-01-2017 and 23-04-2021, chosen to take seasonal dynamics into account: in order to avoid an imbalance between flooded and non-flooded areas the acquisitions were mainly selected during the wet period. Only bands with ground resolution between 10 m and 20 m were used.

DTM with 5 cm ground resolution was generated from photogrammetric processing of Unmanned Aerial Vehicle (UAV) images [28], [29]. The flight was carried out during the summer of 2020, when the grazing is well advanced in order to have a low vegetation cover. High vegetation patches and features that stand out from the scene surface (e.g. cows) have been removed and elevation has been approximated by Inverse Distance Weighting (IDW) interpolation. The effect of vegetation is not fully corrected as vegetation after grazing is around 10 cm high which corresponds to the vertical error of the DTM : 9 cm (+/- 6 cm), estimated via 50 positioning surveys by differential Global Navigation Satellite System with Post-Processed Kinematic (centimetric vertical precision). This vertical error identified in the production of ground truth data represents a decimeter bias in assessing the accuracy of predicted water levels. The DTM was resampled at 1 m to reduce the scale factor with the S2 data. We assume that topographic changes are negligible over the studied period.

The data has been, first, divided into tiles of 112x112 pixels, resulting in 4620 tiles. Then, this dataset was divided into three sets, with the train set comprising data from the years 2017-2019 (75% of total tiles), the validation set 2020 (23%) and the test set 2021 (2%). Finally, data mirroring and random deformation techniques were used for data augmentation.

#### B. Fusion network for super resolution

Our approach, called Fusion-UNet (Fig. 1), consists of merging two sources of information useful for estimating surface water heights: coarse-resolution optical data (S2) and fine-resolution topographic data (DTM). The data fusion is performed through a semantic segmentation model (U-Net [30]) adapted to super-resolution and regression.

Audebert *et al.* [31] have proposed different fusion strategies that go beyond the simple concatenation of input data, as treating each data source in its own encoding branch. Inspired by this work, we applied an early fusion using a third “virtual” branch that allows fusion during encoding in a symmetrical way, i.e. by treating the two data sources independently. This approach, which encodes the two data sources separately and in parallel, avoids imbalances arising from the different nature of the data. The “virtual” branch corresponds to the sum at

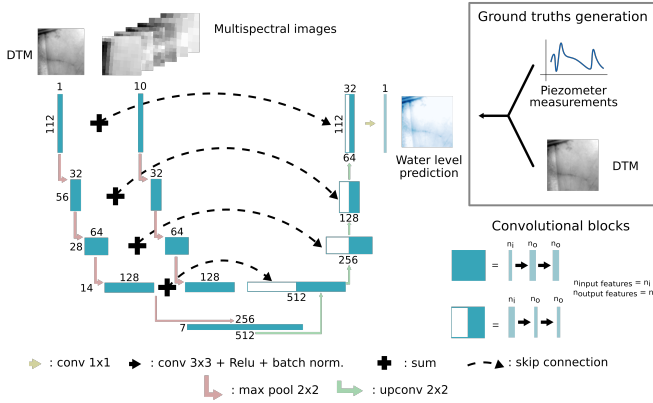


Fig. 1. Fusion-UNet workflow.

each level of the two encoding branches, which is then used for explicit connections to the decoding branch (and also at the lowest encoding level). In addition, a regularisation step with batch normalisation was added to the original blocks [32].

Due to the task at hand, that is pixel-wise regression from fused data, the structure of the loss function plays a critical role. Given an input  $y \in \mathbb{R}^{d^1 \times d^2 \times c}$ , with  $d^1$  and  $d^2$  its sizes and  $c$  the number of channels, and the predicted output  $\hat{y} \in \mathbb{R}^{d^1 \times d^2}$ , the loss function  $\mathcal{L}(\hat{y}, y)$  can be written as a sum of three parts:

$$\mathcal{L}(\hat{y}, y) = \lambda_p \ell_{percep.}^{\phi, j}(y, \hat{y}) + \lambda_c \ell_{cont.}(y, \hat{y}) + \lambda_s \ell_{seg.}(y, \hat{y}) \quad (1)$$

where  $\ell_{percep.}^{\phi, j}(y, \hat{y})$  is a perception function [25],  $\ell_{cont.}(y, \hat{y})$  a content function and  $\ell_{seg.}(y, \hat{y})$  a segmentation function. All three terms are preceded by a  $\lambda$  hyperparameter to adjust its importance with respect to the others. For the perception term, we used a pre-trained network (VGG-16 [33]) as a feature extractor, with  $\phi_j$  the features at the  $j$ th convolutional layer :

$$\ell_{percep.}^{\phi, j}(y, \hat{y}) = \frac{1}{c_j d_j^1 d_j^2} |\phi_j(\hat{y}) - \phi_j(y)| \quad (2)$$

with  $c_j, d_j^1$  and  $d_j^2$  depending on the layer. This term encourages the network to generate similar features for  $\hat{y}$  and  $y$ . On the other hand,  $\ell_{cont.}(y, \hat{y})$  is the simple  $L_1$  distance between input  $y$  and output  $\hat{y}$ :

$$\ell_{cont.}(y, \hat{y}) = \frac{1}{d^1 d^2} |\hat{y} - y| \quad (3)$$

Finally,  $\ell_{seg.}(y, \hat{y})$  forces to consider non-wet surfaces as 0 m water heights, with reference data classified as "water" and "non-water". This allows us to frame this part as a semantic segmentation term and to use the IoU (Intersection-over-Union) as a loss function.

The hyperparameters have been chosen empirically and kept fixed for all experiments, except for the  $\lambda$  parameters: an SGD optimiser was used with a momentum of 0.9, gamma of 0.2, step of 15. Due to computing limitations, all experiments ran for 50 epochs with a batch of 10 images. The model algorithm is written with Pytorch and is available here : <https://gitlab.com/EmilienAlvarez/spim-nn>.

## IV. EXPERIMENTS AND DISCUSSION

In this section, we aim to evaluate the contribution of the fusion strategy and the loss function for reconstructing high-resolution images of flooded areas. For the first task, we compared Fusion-UNet with a standard UNet with two configuration: with (UNet-MNT-S2) and without (UNet-S2) DTM data as auxiliary data in input. To evaluate the performance of the models on  $N$  estimated data  $\{\hat{y}_1, \dots, \hat{y}_N\}$ , the mean absolute error (MAE), which compares the  $\hat{y}$  estimates with the  $y$  reference data, is often used :

$$MAE = \frac{1}{N \times d^1 \times d^2} \sum_k \sum_i \sum_j |\hat{y}_k(i, j) - y_k(i, j)| \quad (4)$$

where,  $i$  and  $j$  correspond to the pixel's indices.

Given the strong disparity in the distribution of water levels in the reference data - non-flooded and lightly flooded areas are over-represented in the dataset - the use of the MAE, alone, may then appear biased. Therefore, we adjust this metric by rebalancing the contribution of WLS in its calculation. The sample is then divided into  $G$  ranges of 10 cm WL, i.e.:  $[0; (0, 10]; \dots (G-10, G]$ . The adjusted MAE ( $MAE_{adj}$ ) is then the average of the MAEs of each  $G$  ranges, i.e. :

$$MAE_{adj} = \frac{1}{G} \sum_{k=0}^G MAE_k \quad (5)$$

with  $k$  an index running on the total  $G$  ranges.

### A. Contribution of the fusion strategy

The performance evaluation of each model is presented in Table I. The model UNet-S2 shows the worst performance with an  $MAE_{adj}$  of 0.16 m on the test dataset. On the same dataset, the models UNet-MNT-S2 and Fusion-UNet both show better performances with an  $MAE_{adj}$  of 0.09 m. However, on the training and validation datasets, *Fusion-UNet* slightly outperforms UNet-MNT-S2 by +0.01 m. Finally, whatever the model, the results on the validation and test dataset are better than on the training. This point accounts for an imbalance between the learning, validation and test sets due to the inter-annual variability of the flooding dynamics. Indeed, the high ranges of WL, i.e.  $>140$  cm, are rare and only present in the training set.

Visually, we can see differences between the predictions of the different models (Figure 2). It can be seen that the predictions of the model UNet-S2 are correct for coarse spatial patterns but do not recover fine patterns. The predictions of the models UNet-MNT-S2 and Fusion-UNet, which incorporate the fine resolution DTM, are closer to the reference with very similar spatial patterns. However, the UNet-MNT-S2 model restores in its predictions the coarse resolution pixel raster of the Sentinel-2 data, unlike the Fusion-UNet model.

These results show, visually and via the evaluation metric ( $MAE_{adj}$ ), that the Fusion-UNet model performs better than the UNet-S2 and UNet-MNT-S2 models. Therefore, the rest of the performance and sensitivity evaluation focuses on this model and validation dataset.

TABLE I  
PERFORMANCE COMPARISON ( $MAE_{adj}$  IN METER) BETWEEN THE MODELS UNET-S2, UNET-MNT-S2 AND FUSION-UNET.

Models	Training	Validation	Test
UNet-S2	0.27	0.24	0.16
UNet-MNT-S2	0.17	0.14	0.09
<b>Fusion-UNet</b>	0.16	0.13	0.09

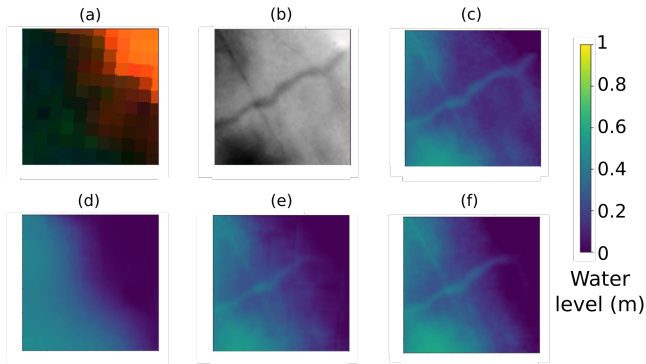


Fig. 2. Visual comparison of the water level predictions of the models UNet-S2, UNet-MNT-S2 and Fusion-UNet. a) Sentinel-2 false colour composition (R: B8 - G:B4 - B:B3); b) UAV digital terrain model; c) reference water level; d) UNet prediction (S2); e) UNet prediction (S2 + DTM); and f) Fusion-UNet prediction

### B. Contribution of the loss function

We first evaluated the contribution of the perception loss in the computation of the compound loss. We evaluated the perception loss computed for various positions of the VGG network from low to high abstraction levels. The first layers (low) indeed concentrate fine-scale structures, while deeper layers (high) extract more global scales. We observed that each layer used separately could not significantly reduce alone the  $MAE_{adj}$  value contrary to the use of the mix of each layer (Table IV-B).

TABLE II  
PERFORMANCE COMPARISON ( $MAE_{adj}$  IN METER) BETWEEN THE USED OF DIFFERENT VGG LAYERS FOR THE COMPUTATION OF THE PERCEPTION LOSS.

	Low	Mid	Mid-High	High	Mix
$MAE_{adj}$	0.13	0.13	0.13	0.13	<b>0.11</b>

Then we analyzed the efficiency of *content* vs *perception* losses, i.e.  $\ell_{cont.}$  and  $\ell_{percep.}$  proposed in relations (3) and (2). To this end, we train a network with  $\lambda_c = 1 - \lambda_p$  and  $\lambda_s = 0$  for a set of values of  $\lambda_p \in [0, 1]$  (Fig. 3). This experiment shows that the perception loss is better optimized with  $\lambda_p$  tending to 1, i.e. higher weighting for this component in the compound loss, contrary to the content loss and the  $MAE_{adj}$  metric which are optimal for around 0.5-0.6, i.e. with an equal mix. The segmentation loss shows no trend in the variation of the  $\lambda_p$  coefficient and therefore does not seem to be influenced by the mixture of these two components.

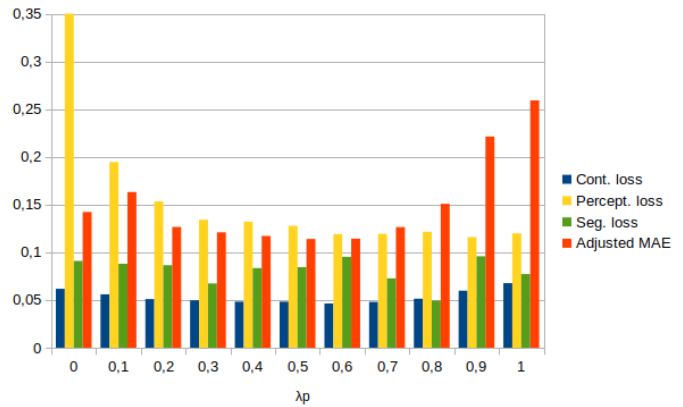


Fig. 3. Losses and metrics comparison for various weighting of the content and perception loss.

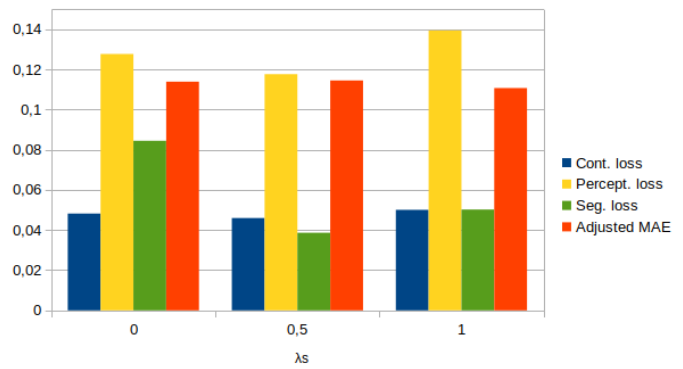


Fig. 4. Losses and metrics comparison for various weighting of the segmentation loss.

Finally, we evaluated the contribution of the segmentation loss  $\ell_{seg.}$  with  $\lambda_c = \lambda_p = 1$  and a set of values of  $\lambda_s \in [0, 1]$  (Fig. 4). We observed that this component in the compound loss has no significant influence on the  $MAE_{adj}$  metric but enhances the distinction between water and no-water surface as the segmentation loss is minimized with  $\lambda_s$  equal to 0.5, moreover it slightly reduces the perception and content losses.

## V. CONCLUSION

The proposed approach exploits the super resolution based on CNNs to map water levels at fine resolution from S2 satellite data. This approach allowed to improve water level mapping in wet grasslands and to derive indicators of ecosystem status for better management.

The Fusion-UNet and UNet-MNT-S2 models make restoring physically exploitable patterns possible, but the former provided more visually realistic results. A sensitivity study on the different components of the compound loss function has allowed to improve the results and to show the interest in crossing content, perception and segmentation losses.

Future research should focus on exploiting SAR data as well, in order to avoid the cloud constraint and to increase the acquisition time series. Domain adaptation methods should be used to jointly process these heterogeneous data through a single model.

## ACKNOWLEDGEMENTS

This work has been partly supported by ANR Projet-ANR-18-CE23-0022 MULTISCALE.

## REFERENCES

- [1] I. Rodríguez-Iturbe, "Hydrologic dynamics and ecosystem structure," *Water Science and Technology: A Journal of the International Association on Water Pollution Research*, vol. 47, no. 6, pp. 18–24, 2003. [Online]. Available: <https://doi.org/10.2166/wst.2003.0347>
- [2] M. Hayashi, G. van der Kamp, and D. O. Rosenberry, "Hydrology of Prairie Wetlands: Understanding the Integrated Surface-Water and Groundwater Processes," *Wetlands*, vol. 36, no. 2, pp. 237–254, Dec. 2016. [Online]. Available: <https://doi.org/10.1007/s13157-016-0797-9>
- [3] C. Damgaard, A. Merlin, and A. Bonis, "Plant colonization and survival along a hydrological gradient: demography and niche dynamics," *Oecologia*, vol. 183, no. 1, pp. 201–210, Jan. 2017. [Online]. Available: <https://doi.org/10.1007/s00442-016-3760-9>
- [4] D. Phiri, M. Simwanda, S. Salekin, V. R. Nyirenda, Y. Murayama, and M. Ranagalage, "Sentinel-2 data for land cover/use mapping: A review," *Remote Sensing*, vol. 12, no. 1, p. 2291, 2020. [Online]. Available: <https://doi.org/10.3390/rs12142291>
- [5] F. Vuolo, M. Neuwirth, M. Immitzer, C. Atzberger, and W.-T. Ng, "How much does multi-temporal sentinel-2 data improve crop type classification?" *International journal of applied earth observation and geoinformation*, vol. 72, pp. 122–130, 2018. [Online]. Available: <https://doi.org/10.1016/j.jag.2018.06.007>
- [6] A. Lefebvre, C. Sannier, and T. Corpetti, "Monitoring urban areas with sentinel-2a data: Application to the update of the copernicus high resolution layer imperviousness degree," *Remote Sensing*, vol. 8, no. 7, p. 606, 2016. [Online]. Available: <https://doi.org/10.3390/rs8070606>
- [7] M. M. F. Wong, J. C. H. Fung, and P. P. S. Yeung, "High-resolution calculation of the urban vegetation fraction in the pearl river delta from the sentinel-2 ndvi for urban climate model parameterization," *Geoscience Letters*, vol. 6, no. 1, pp. 1–10, 2019. [Online]. Available: <https://doi.org/10.1186/s40562-019-0132-4>
- [8] K. Toming, T. Kutser, A. Laas, M. Sepp, B. Paavel, and T. Nöges, "First experiences in mapping lake water quality parameters with sentinel-2 msi imagery," *Remote Sensing*, vol. 8, no. 8, p. 640, 2016. [Online]. Available: <https://doi.org/10.3390/rs8080640>
- [9] Q. Wang, W. Shi, Z. Li, and P. M. Atkinson, "Fusion of sentinel-2 images," *Remote sensing of environment*, vol. 187, pp. 241–252, 2016. [Online]. Available: <https://doi.org/10.1016/j.rse.2016.10.030>
- [10] Q. Wang and P. M. Atkinson, "Spatio-temporal fusion for daily sentinel-2 images," *Remote Sensing of Environment*, vol. 204, pp. 31–42, 2018. [Online]. Available: <https://doi.org/10.1016/j.rse.2017.10.046>
- [11] M. El Hajj, N. Baghdadi, M. Zribi, and H. Bazzi, "Synergic use of sentinel-1 and sentinel-2 images for operational soil moisture mapping at high spatial resolution over agricultural areas," *Remote Sensing*, vol. 9, no. 12, p. 1292, 2017. [Online]. Available: <https://doi.org/10.3390/rs9121292>
- [12] A. K. Skidmore, N. C. Coops, E. Neinavaz, A. Ali, M. E. Schaepman, M. Paganini, W. D. Kissling, P. Vihervaara, R. Darvishzadeh, H. Feilhauer, M. Fernandez, N. Fernández, N. Gorelick, I. Geijzendorffer, U. Heiden, M. Heurich, D. Hobern, S. Holzwarth, F. E. Muller-Karger, R. Van De Kerchove, A. Lausch, P. J. Leitão, M. C. Lock, C. A. Múcher, B. O'Connor, D. Rocchini, C. Roesli, W. Turner, J. K. Vis, T. Wang, M. Wegmann, and V. Wingate, "Priority list of biodiversity metrics to observe from space," *Nature Ecology & Evolution*, vol. 5, no. 7, pp. 896–906, Jul 2021. [Online]. Available: <https://doi.org/10.1038/s41559-021-01451-x>
- [13] B. Krause, H. Culmsee, K. Wesche, and C. Leuschner, "Historical and recent fragmentation of temperate floodplain grasslands: Do patch size and distance affect the richness of characteristic wet meadow plant species?" *Folia Geobotanica*, vol. 50, no. 3, pp. 253–266, Sep. 2015. [Online]. Available: <https://doi.org/10.1007/s12224-015-9220-1>
- [14] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," in *Proceedings of the European conference on computer vision (ECCV) workshops*, 2018. [Online]. Available: <https://doi.org/10.48550/arXiv.1809.00219>
- [15] M. Yates, G. Hart, R. Houghton, M. Torres Torres, and M. Pound, "Evaluation of synthetic aerial imagery using unconditional generative adversarial networks," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 190, pp. 231–251, Aug. 2022. [Online]. Available: <https://doi.org/10.1016/j.isprsjprs.2022.06.010>
- [16] F. Ling, F. Xiao, Y. Du, H. P. Xue, and X. Y. Ren, "Waterline mapping at the subpixel scale from remote sensing imagery with high-resolution digital elevation models," *International Journal of Remote Sensing*, vol. 29, no. 6, pp. 1809–1815, 2008. [Online]. Available: <https://doi.org/10.1080/01431160701802489>
- [17] D. Schaffer-Smith, J. J. Swenson, M. E. Reiter, and J. E. Isola, "Quantifying shorebird habitat in managed wetlands by modeling shallow water depth dynamics," *Ecological Applications*, vol. 28, no. 6, pp. 1534–1545, 2018. [Online]. Available: <https://doi.org/10.1002/eap.1732>
- [18] H. Zwenzner and S. Voigt, "Improved estimation of flood parameters by combining space based SAR data with very high resolution digital elevation data," *Hydrology and Earth System Sciences*, vol. 13, no. 5, pp. 567–576, 2009. [Online]. Available: <https://doi.org/10.5194/hess-13-567-2009>
- [19] P. Wang, B. Bayram, and E. Sertel, "A comprehensive review on deep learning based remote sensing image super-resolution methods," *Earth-Science Reviews*, p. 104110, 2022. [Online]. Available: <https://doi.org/10.1016/j.earscirev.2022.104110>
- [20] L. Yue, H. Shen, J. Li, Q. Yuan, H. Zhang, and L. Zhang, "Image super-resolution: The techniques, applications, and future," *Signal Processing*, vol. 128, pp. 389–408, 2016. [Online]. Available: <https://doi.org/10.1016/j.sigpro.2016.05.002>
- [21] M. Gargiulo, A. Mazza, R. Gaetano, G. Ruello, and G. Scarpa, "Fast super-resolution of 20 m sentinel-2 bands using convolutional neural networks," *Remote Sensing*, vol. 11, no. 22, p. 2635, 2019. [Online]. Available: <https://doi.org/10.3390/rs11222635>
- [22] C. Lanaras, J. Bioucas-Dias, S. Galliani, E. Baltasvias, and K. Schindler, "Super-resolution of sentinel-2 images: Learning a globally applicable deep neural network," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 146, pp. 305–319, 2018. [Online]. Available: <https://doi.org/10.1016/j.isprsjprs.2018.09.018>
- [23] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017. [Online]. Available: <https://doi.org/10.48550/arXiv.1609.04802>
- [24] L. Salgueiro Romero, J. Marcello, and V. Vilaplana, "Super-resolution of sentinel-2 imagery using generative adversarial networks," *Remote Sensing*, vol. 12, no. 15, p. 2424, 2020. [Online]. Available: <https://doi.org/10.3390/rs12152424>
- [25] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual Losses for Real-Time Style Transfer and Super-Resolution," *arXiv:1603.08155 [cs]*, Mar. 2016. [Online]. Available: <https://doi.org/10.48550/arXiv.1603.08155>
- [26] K. M. Masoud, C. Persello, and V. A. Tolpekin, "Delineation of agricultural field boundaries from sentinel-2 images using a novel super-resolution contour detector based on fully convolutional networks," *Remote sensing*, vol. 12, no. 1, p. 59, 2020. [Online]. Available: <https://doi.org/10.3390/rs12010059>
- [27] O. Hagolle, M. Huc, C. Desjardins, S. Auer, and R. Richter, "MAJA Algorithm Theoretical Basis Document," Dec. 2017. [Online]. Available: <https://zenodo.org/record/1209633>
- [28] A. Annis, F. Nardi, A. Petroselli, C. Apollonio, E. Arcangeletti, F. Tauro, C. Belli, R. Bianconi, and S. Grimaldi, "UAV-DEMs for Small-Scale Flood Hazard Mapping," *Water*, vol. 12, no. 6, p. 1717, Jun. 2020. [Online]. Available: <https://doi.org/10.3390/w12061717>
- [29] G. J.-P. Schumann, J. Muhlhausen, and K. M. Andreadis, "Rapid Mapping of Small-Scale River-Floodplain Environments Using UAV SfM Supports Classical Theory," *Remote Sensing*, vol. 11, no. 8, p. 982, Jan. 2019. [Online]. Available: <https://doi.org/10.3390/rs11080982>
- [30] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *arXiv:1505.04597 [cs]*, May 2015. [Online]. Available: <https://doi.org/10.48550/arXiv.1505.04597>
- [31] N. Audebert, A. Boulch, B. Le Saux, and S. Lefèvre, "Distance transform regression for spatially-aware deep semantic segmentation," *Computer Vision and Image Understanding*, vol. 189, p. 102809, Dec. 2019. [Online]. Available: <https://doi.org/10.1016/j.cviu.2019.102809>
- [32] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *arXiv:1502.03167 [cs]*, Mar. 2015. [Online]. Available: <https://doi.org/10.48550/arXiv.1502.03167>
- [33] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv:1409.1556 [cs]*, Apr. 2015. [Online]. Available: <https://doi.org/10.48550/arXiv.1409.1556>