Scalable Wetland Condition Monitoring With Harmonized Sentinel-2-Based Inundation Regime Indicators Across Europe

Sebastiaan Verbesselt[1,], Stien Heremans[1], Tytti Jussila[2], Vít Ježek[3], Maria Šibíková[4], Risto K. Heikkinen[2]*

[1] Research Institute Nature and Forest (INBO), Brussels, Belgium, sebastiaan.verbesselt@inbo.be, stien.heremans@inbo.be

[2] The Finnish Environment Institute (SYKE), Helsinki, Finland, tytti.jussila@syke.fi, risto.heikkinen@syke.fi

[3] Nature Conservation Agency of the Czech Republic (AOPK ČR), Praha, Czech Republic, vit.jezek@aopk.gov.cz

[4] Slovak Academy of Sciences (SAS), Bratislava, Slovakia, maria.sibikova@savba.sk

[*] corresponding author

doi: 10.5281/zenodo.20396223

Abstract: Reliable biodiversity reporting under the EU Habitats Directive and the Nature Restoration Law requires harmonized monitoring of hydrological dynamics across Europe's (wet) grasslands and wetlands. Within the Biodiversa+ Habitat Pilot, we evaluated the scalability of Sentinel-2-based inundation models and the robustness of derived inundation regime indicators. We validated two inundation models across diverse European regions, achieving medium to high accuracy in open habitats (Macro-F1 between 0.64 and 0.93) but encountering limitations in densely vegetated areas and small, fragmented water bodies. To assess the reliability of long-term inundation regime monitoring for wetland habitat condition assessment, we conducted a sensitivity analysis. We quantified how different time series preprocessing decisions such as cloud filtering, temporal aggregation, and interpolation methods impact the indicator Mean Annual Inundated Fraction (MAIF). Our results indicate that this indicator is remarkably stable across various processing workflows variants, with site-specific ecological characteristics accounting for most of the variance. This suggests that simple, Sentinel-2-based inundation workflows can be sufficiently robust for large-scale reporting, provided they are adapted to local ecological contexts. For future research, we recommend exploring multi-class inundation classification or continuous soil moisture modeling to capture the finer ecological gradients in even more detail.

Keywords: wetland condition; inundation regime; Sentinel-2; satellite time series.

1 Introduction

Monitoring inundation regimes - the spatio-temporal dynamics of surface water extent, frequency and duration - is fundamental for assessing the ecological quality and integrity of Europe's Natura 2000 wetland and wet grassland habitats (Jussila et al. 2024). As important hydrological drivers, these regimes affect nutrient cycling, carbon sequestration potential and the maintenance of specialized biodiversity within specific Annex I habitats, such as transition mires (7140) and lowland hay meadows (6510). While traditional field surveys can offer high-quality data, they lack the temporal continuity required to capture ecosystem's hydrological dynamics. Advances in remote sensing, particularly the application of inundation detection models to satellite time series, can provide a harmonized framework for large-scale inundation regime monitoring. This is proven in large-scale products like the Global Surface Water Explorer (GSWE) and the CLMS layer Water Cover Duration (Copernicus Land Monitoring Service 2025, Pekel et al. 2016).

However, optical sensors (like the multispectral instrument of Sentinel-2) introduce some challenges, including irregular data density due to cloud cover. The reliability of the derived hydrological indicators - such as annual average inundation levels - depends heavily on how these irregular time series are aggregated, interpolated, and smoothed. As far as we know, the effect of these variations in time series (pre)processing on deriving S2 based hydro-ecological indicators is not really evaluated. Moreover, ensuring that remote sensing products align with ecological requirements is essential for providing the reliable evidence base needed for policy implementation.

Building on the work conducted in the Biodiversa+ Habitat Pilot (Biodiversa 2024), a trans-European collaboration project, this study aims to evaluate the scalability of Sentinel-2-based inundation monitoring by investigating the technical and ecological challenges associated with large-scale time series analysis. Specifically, we aim to (i) validate the performance of simple inundation models across diverse European wetland types and (ii) quantify the impact of various time-series preprocessing options on the stability of yearly inundation regime indicators through a multi-site sensitivity analysis. While the methodological approach is relatively simple, its application offers considerable innovation from a policy-making perspective, demonstrating how accessible tools can support operational wetland monitoring and decision-making.

2 Materials and methods

2.1 Data acquisition and inundation modeling

Sentinel-2 L2A time series (Bands B02–B12 and SCL) were acquired via the OpenEO API (openEO 2026) for the period 2020-2024. To track surface water dynamics, we used two existing inundation segmentation models: Jussila model (Jussila et al. 2024) and Water-in-Wetlands (WiW) model (Lefebvre et al., 2019). These models were selected for their transparency and simplicity, especially compared to more complex and black-box deep-learning approaches (Vali et al., 2020). Validation was performed for 24 sites in seven European regions covering several Annex I habitats. Reference data were generated using (semi-automated) segmentation of orthophotos, drone-based thresholding, and in-situ water delineation; and performance quantified as macro-F1 scores. More information can be found in section 3.3.2 of the final report of the Biodiversa+ Habitat Pilot.

2.2 Hydro-ecological indicator and sensitivity analysis

As part of the Biodiversa+ Habitat Pilot, ecological experts from the participating European countries and regions proposed a set of inundation-based habitat condition indicators for wetlands and wet grasslands, focusing on inundation frequency, duration, and spatial extent. To understand how data processing influences these metrics, a sensitivity analysis was conducted on five selected habitat sites, squared regions of 50 km by 50 km, in Flanders, Finland, Czech Republic and Slovakia, focusing on two habitat types: 7140 (Transition mires) and 6510 (Lowland hay meadows). More information about the study sites and the number of plots per country or region can be found in Table 1.

From the set of proposed indicators, the Mean Annual Inundated Fraction (MAIF) was selected as an indicator for this study. This is the fractional inundated area of a habitat (Tellman et al. 2022), calculated at each time step and averaged over the time steps within a year. It's a simple metric that describes the general conditions of our target habitats throughout the year. Due to time constraints, it was not possible to test further habitat-specific indicators within the project. The reference pipeline included standard settings. We tested five sensitivity analysis variants: (i) with stricter cloud threshold (< 10 %), (ii) without spatial interpolation (iii) using twice per month median for temporal aggregation, (iv) using Gaussian Process Regression for temporal interpolation, and (v) without time series smoothing.

Figure 1 illustrates the conceptual flowchart. Trees were always masked with the CLMS Tree Cover Density dataset (CLMS 2021) and inundation was segmented using the Jussila model.

2.3 Statistical evaluation

The impact of these processing scenarios was evaluated using linear mixed-effects models (LMMs) in R-software (Lenth and Piaskowski, 2017; R Core Team, 2026). The five analysis variants were treated as a fixed effect, while year and habitat type (nested within site) were included as random effects. Post-hoc pairwise comparisons were performed using the emmeans package with Benjamini-Hochberg p-value adjustments (Lenth and Piaskowski, 2017) to identify significant methodological biases in the resulting ecological metrics.

Table 1. Summary of the five sites used for the sensitivity analysis. Central coordinates are provided in EPSG:4326.

Site	X-coord	Y-coord	# 7140 patches	Area (ha)	# 6510 patches	Area (ha)
Finland (North)	26.78	65.27	2	720	/	/
Finland (South)	23.27	60.43	/	/	32	12
Flanders	5.08	51.00	4	4	8	7
Czech Republic	15.48	50.74	60	288	78	2450
Slovakia	18.08	47.76	/	/	6	51

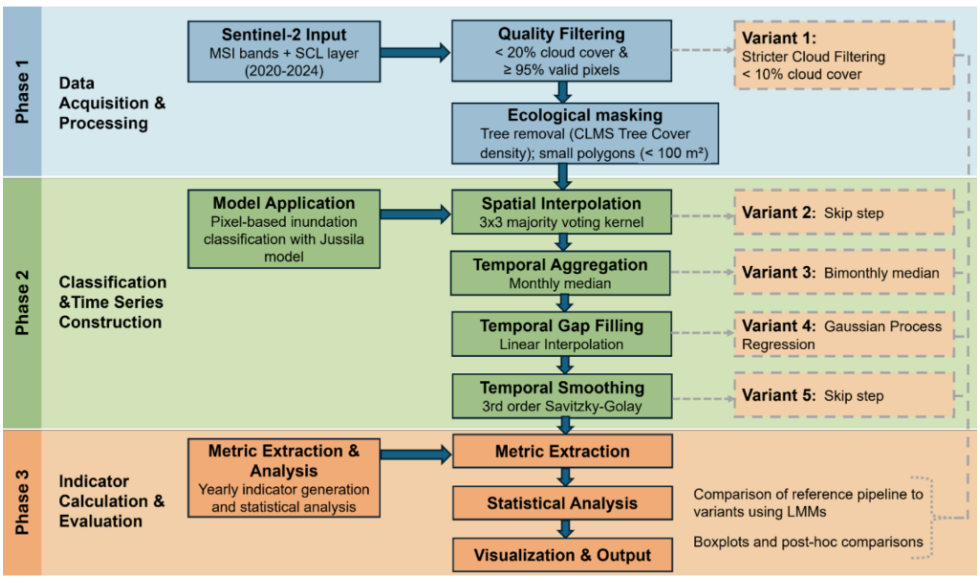

Figure 1. Conceptual flowchart of the analysis workflow including the reference pipeline and the five analysis variants.

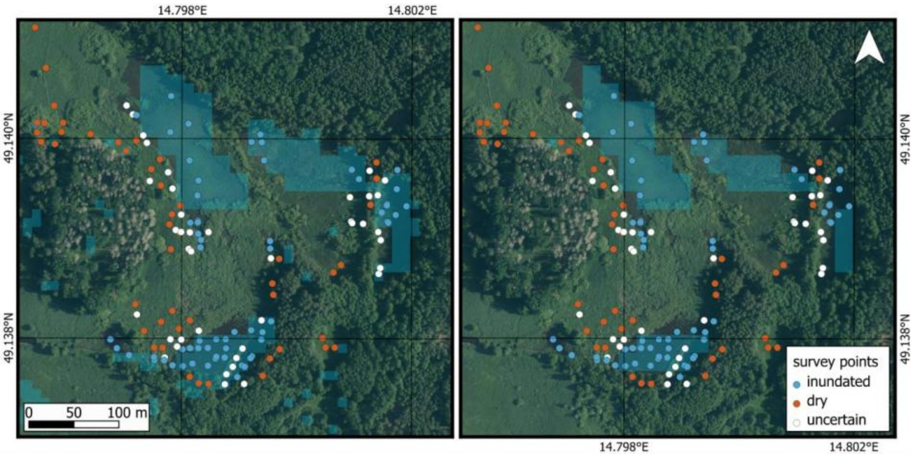

Figure 2. Jussila model output (left) vs. Lefebvre model output (right) for the Kramářka study site in Czechia (inundated = blue). At each survey point, the validator evaluated if the water table depth was on average above (inundated) or below (dry) 5 cm within the surrounding 10 meter. Predictions on the S2 image were made on the closest available date. The background ortho image is from a different date compared to the S2 predicted output.

3 Results

3.1 Inundation model performance

The validation of the Jussila and WiW models across seven European regions showed macro-F1 scores ranging from 0.74 to 0.92 for Jussila and 0.64 to 0.93 for WiW. While both models performed well in open water, and give quite similar outputs (Figure 2), the Jussila model generally detected water in mixed pixels more effectively in Finland and Czechia, while the WiW model performed better in Flanders. Two general limitations were identified: (i) dense tree cover, with coniferous trees and shadows misclassified as water; (ii) dense reedbeds, which obscured surface water detection. Small water patches, such as Mediterranean temporary ponds in Croatia, proved undetectable. More details can be found in section 3.3.2 of the Habitat Pilot final report.

3.2 Ecological inundation indicators and expert needs

Expert surveys from the Biodiversa+ Habitat Pilot indicated that weekly or two-weekly temporal resolutions are optimal for capturing critical hydrological patterns, though monthly data remains useful for most habitat types. Minimum seasonal wetness was identified as a key indicator of habitat stress for most wetlands and wet grasslands, particularly for sensitive systems like aapa mires (7310) and alkaline fens (7230). While inundation monitoring is universally relevant for wetlands, its utility for grasslands varies by region and type, with high-frequency monitoring specifically required for alluvial meadows (6440 and 6510) subject to short-term flooding.

3.3 Time series sensitivity analysis

The availability of valid S2 scenes varied significantly by region: Slovakia and Finland provided the highest number of cloud-free observations, while Belgium and Czechia were more constrained by persistent cloud cover. Stricter cloud filtering (10% vs. 20%) and valid-pixel thresholding (95% SCL) reduced the available data by up to 75% in some regions. The LMM revealed that varying the time series processing pipeline (e.g., interpolation method or smoothing) did not have a statistically significant effect on MAIF ($p > 0.07$ in all cases). Stricter cloud filtering resulted in the lowest estimated marginal means effect (~ average effect), likely because excluding cloudy scenes increases the risk of missing some flooding events. Conversely, coarser temporal aggregation led to higher estimates due to the use of median values. Crucially, our random effects analysis showed that the specific site and habitat type accounted for far more variance ($\sigma = 0.037$) than the processing methodology or temporal fluctuations. This suggests that while preprocessing choices influence data density, the inherent (hydro-)ecological differences between sites and habitats are the primary driver of the variation in calculated MAIF values.

Figure 3 displays the contrast estimates in marginal mean effects, with standard error bars, between the reference pipeline and the five processing variants. Positive estimates indicated an on average lower predicted MAIF compared to the reference processing pipeline and vice-versa, but no significant differences were observed.

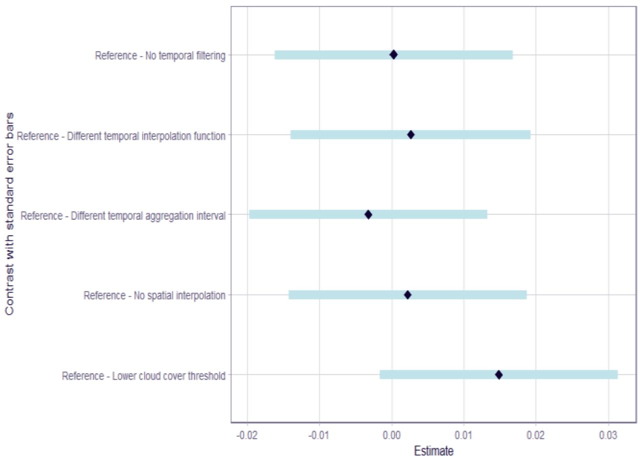

Figure 3. Contrasts estimates in marginal means, with standard error bars, between the reference processing pipeline and the five processing variants.

4 Discussion

The Jussila and WiW inundation segmentation models demonstrated strong potential for harmonized EU-wide monitoring due to their high accuracy combined with transparency and ease of implementation. However, dense vegetation remains a primary constraint. While tree cover can be partially removed via masking, detecting inundation beneath shrubs, thick moss layers, reedbeds, and other aquatic vegetation requires further methodological refinement.

S2 based inundation classification over time can be translated into relevant ecological indicators. We propose the MAIF indicator as a suitable indicator for general condition monitoring of wetland and wet grassland habitats. More habitat specific indicators were proposed within the Pilot by combining expert knowledge of ecologists and remote sensing scientists.

While lacking good quality multitemporal reference data, we evaluated the influence of different preprocessing options for creating regular S2 time series on the effect of calculating MAIF of 66 Annex I 7140 and 124 Annex I 6510 habitats, in seven regions across Europe, through a linear mixture model.

Our sensitivity analysis indicates that despite regional and habitat disparities in data availability, that no significant differences were found in the calculated annual inundation indicator MAIF between the different S2 time series (pre)processing choices.

5 Conclusions

The Biodiversa+ Habitat Pilot demonstrated the potential of S2-based inundation segmentation models as a transparent, scalable and harmonizable option for monitoring inundation regimes across Europe's diverse open habitats. Dense biomass remains a technical hurdle for inundation mapping, which could be (partially) overcome with the integration of digital terrain models (DTMs) (Landuyt et al. 2021, Munasinghe et al. 2018).

The Pilot did however identify several core challenges in the validation process itself. A key conceptual challenge is that inundation is often a continuous spatial gradient, not a discrete "water/no-water" boundary. This ambiguity directly conflicts with a binary classification approach. This raises challenges in obtaining unbiased accuracy metrics. To address the limitations of binary classification in capturing continuous spatial gradients, future monitoring should explore multi-class models or regression modelling to better capture variations in soil moisture levels (Maertens 2025).

Reference data (orthophotos, drone data and in-situ) are heterogeneous and might introduce inconsistencies over different sites. We recommend a hybrid validation strategy that combines the objectivity of field-verified points or polygons with the upscaling potential of drone imagery, ideally timed to maximally coincide with satellite acquisitions. Furthermore, as these models are extrapolated across time and space, quantifying epistemic uncertainty or total uncertainty (e.g., through Shannon Entropy) becomes essential to evaluate model reliability at the European scale.

Besides the MAIF indicator, several other interesting habitat specific indicators were formulated during the project. Some of these could be already derived from large scale open data products like the Global Surface Water Explorer (GSWE) and the CLMS layer Water Cover Duration product. While the former is limited a medium spatial resolution (30 meter/pixel) and observations between 1984-2021 (Pekel et al. 2016), the latter provides an up-to-date product of higher spatial resolution (10 meter/pixel), but only for a very limited set of indicators (Copernicus Land Monitoring Service 2025).

The sensitivity analysis shows that annual hydrological indicators are remarkably stable across various (pre)processing scenarios. However, failing to detect statistical difference does not mean that there are no differences ("fail to reject the null hypothesis" does not mean "accept the null hypothesis"). Further research should focus on (i) equivalence hypothesis testing, but also (ii) on more robust validation of these dynamic systems with larger multi-temporal reference datasets.

Optical data will always be affected by atmospheric conditions. Future improvement could therefore focus on data fusion between active radar and optical sensors (Landuyt et al. 2024, Rahmati et al. 2026 Yin et al. 2024), alongside advanced spatiotemporal interpolation (Consoli et al. 2024, Zhang et al. 2026). Ultimately, establishing common European standards for indicator definitions, open-source processing chains, and sharing multi-temporal reference datasets via collaboration platforms like Eionet will be critical for achieving the remote sensing based harmonized monitoring requirements in the framework of the EU's Nature Restoration Law and Habitats Directive (European Commission 2026, European Union 2024).

Acknowledgments: This research was funded by Biodiversa+, as part of the Habitat Pilot under WP2.

6 References

Biodiversa, 2024. Remote sensing for habitat monitoring, https://www.biodiversa.eu/biodiversity-monitoring/pilots/habitat-pilot/. (Accessed 22 May, 2026).

CLMS, 2021. European Tree Cover Density 2021, https://land.copernicus.eu/en/products/high-resolution-layer-forests-and-tree-cover/tree-cover-density-2021-raster-10-m-100-m-europe-yearly. (Accessed 22 May, 2026).

Consoli, D., Parente, L., Simoes, R., Şahin, M., Tian, X., Witjes, M., Sloat, L., Hengl, T., 2024. A computational framework for processing time-series of earth observation data based on discrete convolution: global-scale historical Landsat cloud-free aggregates at 30 m spatial resolution. PeerJ 12, e18585.

Copernicus Land Monitoring Service, 2025. Water Cover Duration 2016-present (raster 10 m and 100 m), Europe, yearly [WWW Document], https://land.copernicus.eu/en/products/water-

bodies/high-resolution-water-cover-duration. (Accessed 10 April, 2026).
European Commission, 2026. The Habitats Directive - Environment - European Commission, https://environment.ec.europa.eu/topics/nature-and-biodiversity/habitats-directive_en. (Accessed 25 February, 2026).
European Union, 2024. REGULATION (EU) 2024/1991 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 24 June 2024 on nature restoration and amending Regulation (EU) 2022/869, https://eur-lex.europa.eu/eli/reg/2024/1991/oj/eng. (Accessed 25 February, 2026).
Jussila, T., Heikkinen, R.K., Anttila, S., Aapala, K., Kervinen, M., Aalto, J., Vihervaara, P., 2024. Quantifying wetness variability in aapa mires with Sentinel-2: towards improved monitoring of an EU priority habitat. Remote Sensing in Ecology and Conservation 10, 172–187.
Landuyt, L., Ivashkovych, X., Van Achteren, T., 2024. Consistent Surface Water Monitoring by Fusing Sentinel-1 and -2 Through Convolutional Neural Networks. In: Proceedings of the IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, pp. 4701–4705.
Landuyt, L., Van Coillie, F.M., Vogels, B., Dewelde, J., Verhoest, N.E., 2021. Towards operational flood monitoring in Flanders using Sentinel-1. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 11004-11018.
Lefebvre, G., Davranche, A., Willm, L., Campagna, J., Redmond, L., Merle, C., Guelmami, A., Poulin, B., 2019. Introducing WIW for Detecting the Presence of Water in Wetlands with Landsat and Sentinel Satellites. Remote Sensing 11, 2210.
Lenth, R.V., Piaskowski, J., 2017. emmeans: Estimated Marginal Means, aka Least-Squares Means.
Maertens, M., 2025. Remote sensing of soil moisture: from complex algorithms to commercial applications, https://digitalwater.vito.be/en/remote-sensing-soil-moisture-complex-algorithms-commercial-applications. (Accessed 26 March, 2026).
Munasinghe, D., Cohen, S., Huang, Y., Tsang, Y., Zhang, J., Fang, Z., 2018. Intercomparison of Satellite Remote Sensing-Based Flood Inundation Mapping Techniques. Journal of the American Water Resources Association, 54(4), 834-846.
openEO, 2026. EO data available from Copernicus Data Space Ecosystem, https://documentation.dataspace.copernicus.eu/APIs/openEO/openEO.html. (Accessed 22 May, 2026).
Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418–422.
R Core Team, 2026. R: The R Project for Statistical Computing, https://www.r-project.org/. (Accessed 9 April, 2026).
Rahmati, M., Balenzano, A., Bechtold, M., Brocca, L., Fluhrer, A., Jagdhuber, T., Karamvasis, K., Mengen, D., Reichle, R.H., Kim, S., Taghizadeh-Mehrjardi, R., Walker, J., Zhu, L., Montzka, C., 2026. Soil moisture retrieval from Sentinel-1: Lessons learned after more than a decade in orbit. Remote Sensing of Environment 333, 115146.
Tellman, B., Lall, U., Islam, A.K.M.S., Bhuyan, M.A., 2022. Regional Index Insurance Using Satellite-Based Fractional Flooded Area. Earth's Future 10, e2021EF002418.
Vali, A., Comai, S., Matteucci, M., 2020. Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review. Remote Sensing 12, 2495.
Yin, Z., Wu, P., Li, X., Hao, Z., Ma, X., Fan, R., Liu, C., Ling, F., 2024. Super-resolution water body mapping with a feature collaborative CNN model by fusing Sentinel-1 and Sentinel-2 images. International Journal of Applied Earth Observation and Geoinformation 134, 104176.
Zhang, Y., Liang, S., Li, W., Ma, H., Xu, J., Ma, Y., ..., Xia, X.G., 2026. UniTS: Unified Spatio-Temporal Generative Model for Remote Sensing. arXiv preprint arXiv:2512.04461.