Satellite-observed flood indicators are related to riparian vegetation communities

Miriam Herrmann\textsuperscript{a,*,}  Ephraim Schmidt-Riese\textsuperscript{b}, Daria Alison Bäte\textsuperscript{c}, Fabian Kempfer\textsuperscript{c}, Fabian Ewald Fassnacht\textsuperscript{d}, Gregory Egger\textsuperscript{c,d}

\textsuperscript{a} Department of Remote Sensing and Geoinformation, Institute of Geographic Sciences, Freie Universität Berlin, 12249 Berlin, Germany
\textsuperscript{b} Geoinformation in Environmental Planning Lab, Technical University Berlin, 10623 Berlin, Germany
\textsuperscript{c} Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany
\textsuperscript{d} Wetland Ecology, Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany

A R T I C L E  I N F O

Keyword:
Time series
Flood regime
Disturbance
Remote sensing
Braided river
Biogeomorphic succession

A B S T R A C T

In riverine ecosystems, flood disturbances govern the distribution and assembly of vegetation communities. However, anticipated connections between disturbance regimes and habitats often cannot be numerically described as the quantification of disturbances remains challenging. This work presents a novel approach to numerically characterize disturbance patterns in a spatially and temporally explicit way. For this, a dense time series of flood extents was created applying a semi-automated water detection approach to multispectral optical satellite data (Landsat 4, 5, 7, and 8; Sentinel-2) to an 8-km stretch of the alpine Lech River in Tirol, Austria. Twelve multitemporal metrics referred to as Flood Indicators (FIs) were derived and compared to habitat classes formed by vegetation communities as observed in field campaigns in July 2020 and 2021.

The FI values indicated high disturbance intensities for areas with early successional habitat classes, whereas little to no disturbance were related to late successional classes. This is in accordance with the presumed successional pathway of riverine vegetation communities. We observed numerous significant differences in FI values among early successional classes, supporting the assumption that flood disturbance is a dominant factor shaping the corresponding habitats. In late successional stages, the FI values differed less among the vegetation classes, pointing that other processes are more important in shaping the vegetation community in these stages.

The observed relationship between vegetation communities and the satellite-derived disturbance measures underlines the potential of the presented approach to characterize riverine disturbance regimes. Given the global availability of satellite data, the approach should be easily transferable to other study areas.

1. Introduction

Riverine ecosystems are strongly shaped by disturbances and a clear causal link among flood disturbances and riparian vegetation communities has been identified in earlier studies (Resh et al., 1988; Tochner and Stanford, 2002; Tiegs et al., 2005). Pickett and White (2005) define a disturbance as a “discrete event in time that disrupts the ecosystem, community, or population structure and changes resources, substrate availability or the physical environment”. Biomass is destroyed, organisms are removed, and space is cleared for recolonization (Townsend, 1989; Pickett et al., 1999; Grime, 2006). A disturbance can be described using three dimensions: magnitude, frequency, and size (Pickett and White, 2005). Although disturbances often occur during a short time period compared to the life span of affected species, they have lasting effects (White, 1979).

Driven by the pulsing of the river discharge, riparian ecosystems are highly heterogeneous and dynamic (Tochner et al., 2000; Naiman et al., 2005; Camporeale et al., 2013). The river’s hydro-regime with its flow extremes (floodings) defines and shapes river ecosystems with its riparian habitats and plant communities (Poff et al., 1997; Pettit et al., 2001; Stromberg, 2001; Lyle and Merritt, 2004; Perona et al., 2009; Gurnell et al., 2012; Egger et al., 2013; Gurnell et al., 2016). Particularly in braided river ecosystems, flood disturbance initiates erosion, sedimentation and relocation processes which determines different lateral vegetation compositions (Gurnell et al., 2016) and interferes with plants on a physical and a physiochemical level (Townsend, 1989; Bendix,
1997). On a physical level, floods can exert shear stress and sediment erosion, leading to physical damage and uprooting; sediment deposition can result in plant burial. On a physicochemical level, flooding can saturate soils and inundate plants, leading to anoxia and, consequently, impeded photosynthesis (Bendix, 1997; Bendix and Hupp, 2000; Garsen et al., 2015). Vegetation, however, also influences its environment. Roots anchor sediments, aboveground vegetation parts decrease water flow speed which leads to sediment deposition, and vegetation branches and leaves physically trap sediment particles (Abbe and Montgomery, 2003; Fetherston et al., 1995; Gurnell, 1997; Han et al., 2020). Vegetation is hence acting as a riparian engineer (Gurnell, 2014; Surian et al., 2015; Vesipa et al., 2015).

The interplay between vegetation and disturbances leads to the forming of distinct spatial patterns (Gurnell and Grabowski, 2016) with disturbances providing the physical template for stream communities (Poff and Ward, 1989). In riverine ecosystems, sites with different disturbance regimes are in close proximity to each other. On a small scale, habitats range from newly disturbed barren shorelines to mature riverine forests (Townsend, 1989; Tockner et al., 2006). Areas close to the river are ecotones between aquatic and terrestrial ecosystems (Tockner and Stanford, 2002; Naiman et al., 2005; Tockner et al., 2010) and can be characterized according to their disturbance regime (Richter et al., 1996; Tockner and Stanford, 2002; Rusnak et al., 2022).

To better numerically describe the influence of the disturbance regime on local vegetation communities, we require data that are a) spatially explicit, b) available at fine grain to depict local effects, and c) have a high temporal resolution (Pickett et al., 1999; Tockner et al., 2000). Although long-term field sampling is an approach to obtain such data (Pollock et al., 1998), this is associated with a high workload and high costs for continuous resampling (Rusnak et al., 2022). Moreover, a long time period is needed to obtain an ecologically meaningful time series (Rusnak et al., 2022).

An alternative approach to create dense time series of spatially explicit flood disturbance data is the use of hydrological models based on digital elevation models (DEM) and measured gauge data to calculate flood extents (Egger et al., 2017). Although compared to field sampling, the workload decreases, the use of a static DEM is problematic as changes in height due to sediment erosion and deposition are not considered (Abbe and Montgomery, 2003; Vesipa et al., 2015; Vesipa et al., 2017).

A third approach is the use of remotely sensed passive optical data (Betz et al., 2023). Remote sensing is increasingly being used in habitat classification and mapping (Borre et al., 2011; Corbane et al., 2015; McMahon et al., 2024). Time series of airborne RGB aerial images have already been used since the 1990 s to detect flood extents and analyze the influence of flooding on vegetation. In previous studies, flood extents were often digitized manually, which provides high spatial accuracy but limited temporal resolution (e.g., Parsons and Gilvear, 2002; Tegs et al., 2017; Pelizzo et al., 2017). This increases the likelihood that most critical high- or low-water situations are missed (Tegs et al., 2005; Muro Martin et al., 2020; Rusnak et al., 2022).

Higher temporal frequencies (at the cost of lower spatial resolution) can be achieved with the use of satellite images (Ozesmi and Bauer, 2002; Guerschman et al., 2011), which also allows for automated water detection as more spectral bands are available (Ozesmi and Bauer, 2002). Nevertheless, the number of used images and, thus, the temporal resolution are limited (e.g., Bertoldi et al., 2011; Thomas et al., 2011; Marchetti et al., 2020; Fu et al., 2020).

To use the full potential of available satellite time series and create dense flood disturbance history data sets, all available satellite images should be used. This became possible during the last years as cloud computing power increased and satellite images have become freely available (Tulbure and Broich, 2013; Mahdianpari et al., 2018). To our knowledge, up to now, most studies that took advantage of full time series or at least numerous time steps used satellites with low spatial resolution (e.g., Islam et al., 2010; Deng et al., 2014; Huang et al., 2014 all used MODIS with a 500-m pixel size). To analyze local disturbance effects instead of large-scale patterns, satellite data with a higher resolution are preferable (Mueller et al., 2016). Such higher resolution is, for example, provided by the Landsat and Sentinel-2 missions. Their combination of comparatively high spatial and temporal resolution may enable new pathways to numerically describe fluvial disturbance patterns and regimes and their effect on the vegetation communities (Betz et al., 2023).

In our study, we combine Landsat (longest possible time series of more than 35 years) and Sentinel-2 (high spatial and temporal resolution) images to derive proxies for the hydrogeomorphological disturbance regime of an alpine river. The spatially explicit dense time series of flood disturbance events created from satellite images allows us to relate currently observed vegetation communities to the historical, satellite-observed flooding regime (Poff and Ward, 1989; Fausch et al., 2002). We address the following research questions:

(I) Which dimensions (time, space, magnitude) of hydrogeomorphological disturbance can be described using satellite data?

(II) How are the proposed indicators of flood disturbance related to vegetation communities?

2. Methods

2.1. Study area

This study was conducted at an 8-km-long stretch of the Lech River located at approximately 900 m a.s.l. near the village Forchach on the northern edge of the European Alps in Tyrol, Austria (Fig. 1). Here, the river flows in a braided channel network with extensive and constantly changing gravel banks, maintaining an extraordinary diversity of habitats (Müller, 1996). The main sediments are limestone and dolomite, whose coarse structure promotes the development of nutrient-poor soils with a low water storage capacity (Mueller, 1988; Scheurmann and Karl, 1990). The local climate is humid-continental (Dfb sensu Koppen and Geiger, 1930; Kottek et al., 2006), with a mean annual temperature of 8.1 °C and an annual precipitation of 1418 mm (reference station: Reutte, Auer et al., 2021). The river regime is nivo-pluvial, with peak discharge mainly caused by snowmelt in mid-summer and increased discharge caused by precipitation in autumn (Mader et al., 1996). As the study area is relatively small, we assume the climate and weather patterns to be relatively constant over all samples.

We focused our analysis on a core area which included all areas that were detected as at least once flooded using the hereafter described approach. Areas where water detections were implausible (steep terrain, settlements) were excluded. This core area was buffered by a 150-m zone to show the discriminatory power of the flood indicators against areas outside the river basin (Fig. 1).

2.2. Remote sensing analysis

We acquired optical imagery from the Landsat 4, 5, 7, and 8 and Sentinel-2 missions using Google Earth Engine (GEE; Gorelick et al., 2017) (Table 1). The following steps were applied to each image to obtain a continuous time series of water and land observations for each pixel:
First, pixels for which no reliable water detection was possible were masked. This included pixels with clouds, cloud shadows, and terrain shadows. Second, the remaining image was binarized into water and land pixels using a threshold applied to the Normalized Difference Water Index (NDWI). Third, Landsat and Sentinel-2 pictures were merged into one continuous time series (Fig. 2).

2.2.1. Cloud mask
For cloud masking of Landsat 4, 5, 7, and 8 images we used the pre-calculated QA PIXEL band based on the CFmask algorithm (Foga et al., 2017). As the QA PIXEL-based cloud mask is prone to overestimate cloud cover over bright surfaces (Zhu and Woodcock, 2012; Foga et al., 2017; USGS, 2020), its output was refined using a gravel mask to avoid an unintended masking of gravel banks. Here, gravel was detected using the ratio of the near infrared and the surface temperature band which proved to excellently discern gravel from clouds (Supplementary material A). For Sentinel-2 images, cloud masking was performed using an algorithm proposed by Schmitt et al. (2019) (Supplementary material B).

2.2.2. Masking terrain shadows
Terrain shadow masks were created using solar zenith and azimuth angles and a DEM (2019, 5 m pixel length, Land Tirol – data.tirol.gv.at; R function ray_shade, package rayshader, Morgan-Wall, 2021). To conservatively exclude cloud and shadow fringes, all masks were buffered by 30 m.

2.2.3. Water detection
To detect water occurrence, the Normalized Difference Water Index (NDWI; McFeeters, 1996) was calculated. Subsequently, a threshold was applied to binarize images into the classes water and non-water.

\[
\text{NDWI} = \frac{\text{Green} - \text{NIR}}{\text{Green} + \text{NIR}}
\]

Throughout the water detection procedure, several thresholds had to be applied, which were optimized using training points collected on a sample of 69 satellite images. On each of those images 130 training points were placed (100 randomly and 30 in anticipated difficult areas) and labeled as cloud, shadow, land, water, or inconclusive (Supplementary material D).

Another 40 images were chosen to enhance comparability between Sentinel-2 and Landsat water detections. For that 20 dates were identified on which Sentinel-2 as well as Landsat images were taken and corresponding images were downloaded. For each date, 50 points were
randomly sampled for each image pair in areas identified as water and land by Sentinel-2, respectively. These were used as additional training points for the Landsat sensors (Supplementary material D).

For each thresholding operation, a range of thresholds was tested, and that with the overall highest value for Cohen’s Kappa was chosen for the analysis (Supplementary material E).

The calculations of cloud and cloud shadow masks, NDWI, and thresholding thereof were performed in GEE (Gorelick et al., 2017). Training data points were collected using QGIS (Version 3.10, QGIS Development Team, 2022). Threshold optimization, the calculation of terrain shadow masks, and the masking of the binarized NDWI images was performed in R (R Core Team, 2021).

2.3. Post processing

To improve the quality of investigated flood history observations, several quality filters were applied: First, duplicated images present in the GEE archives were automatically removed. Second, images which contained only masked pixels (e.g., completely cloudy scenes) and images without water observations (clouds or shadows obscuring the...
whole riverbed) were automatically removed. Third, each remaining image was rated visually for the plausibility of water observations. Images were claimed unsuitable and excluded if they had an unreasonable low or high proportion of water detections (Supplementary material F).

Landsat images were resampled to the resolution of Sentinel-2 using nearest-neighbor interpolation. Landsat imagery was not used for the time period when Sentinel imagery was available to take full and consistent advantage of the higher spatial resolution of Sentinel-2 scenes.

### 2.4. Quality assessment of the time series

The extent of flooded areas is expected to vary in response to the discharge (Tockner et al., 2006). To validate the plausibility of the remote sensing water detection approach, gauge station measurements of the station Lechaschau were employed (BMLRT, 2020; Hydrographische Dienst Tirol, 2022). The station is located 12 km downstream from the study area. To evaluate the relationship between flooded areas and measured runoff, the Pearson correlation coefficient was calculated for the relative proportion of flooded pixels in the unmasked study area (%) and the daily water runoff (m$^3$/s) of the same date (Pearson, 1895; Supplementary material G).

The final water detection raster images derived from all Landsat and Sentinel-2 sensors contained 594 images, spanning a time period of 36 years (1985 to 2021; Table 1).

### 2.5. Flood indicators

Using the time-series of binary flooded/non-flooded images described above, six Flood Indicators (FIs) as summarized in Table 2 were calculated.

To detect similarities in trends shown by the FIs, correlations between FIs were calculated using the Spearman correlation coefficient (Spearman, 1904). Additionally, the influence of the available imagery frequency on FIs was tested. For this, FIs were calculated from random subsets of all images (Supplementary material N).

### 2.6. Vegetation field sampling

To characterize vegetation composition and habitat structure, field data were collected in summer 2020 and 2021. In total, 281 plots were established using stratified random sampling, using FIs and the Normalized Difference Vegetation Index (NDVI) to define strata (Supplementary material H).

At the location of each chosen point, a 2 × 2-m plant relevé plot was established. In each plot, we recorded ground cover, vegetation structure, and plant species. The recorded ground cover classes were “vegetated soil”, “water”, “silt” (<0.063-mm particle size), “sand” (0.063—2 mm), “gravel” (2—63 mm), “stones” (63—200 mm), and “blocks” (>200 mm; DIN-4022). To assess the vegetation structure, total plant cover was estimated for the vegetation layers moss (on the surface), herb (up to 0.5 m), shrub (0.5—5 m), and tree (over 5 m). All species were recorded. For each species, the cover was estimated separately for each height layer. Covers were estimated on a stepwise scale (1 %, 5 %, 10 % to 100 % in 10 % steps; Londo, 1975) (Supplementary material I).

### 2.7. Habitat classification

Plots were grouped by vegetation and environmental similarity. To obtain an unbiased classification, a purely data-driven, hierarchical classification method was applied (R function isopam, package isopam; Schmidtlein, 2012; Supplementary material J). As input for the classification cover estimates for species, vegetation layers and ground cover classes were used. Combining vegetation and ground cover in the classification allows a meaningful classification of unvegetated plots. Subclasses with less than 10 plots were compiled into their higher-level class to avoid classes with small sample sizes and thus low statistical validity.

### 2.8. Relationship between vegetation and flood indicators

To determine the effects of flood disturbances on the vegetation composition, the FIs were compared to the habitat classes. Analyses of Variance (ANOVA) (Fisher, 1925) was performed to test for differences in mean FI values among classes.

In the ANOVA, only field plots with at least one water detection were used as FI values cannot be expected to distinguish among vegetation classes.

### Table 2

Overview of the flood indicators (FIs).

<table>
<thead>
<tr>
<th>Name</th>
<th>Formula</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flood Frequency (FF)</td>
<td>$\frac{n_{water}}{n_{total}}$ = number of water observations</td>
<td>Thomas et al. (2011), Huang et al. (2014) and Deng et al. (2014), Mueller et al. (2016), Hopkinson et al. (2020), Chen et al. (2021)</td>
</tr>
<tr>
<td>Weighted Flood Frequency (WFF)</td>
<td>$\frac{1}{n} \sum_{i=1}^{n} f(x_i) - f(x_i - 1)$ where $x_i$ denotes the time difference between observation $i$ and $i+1$ in days and $n$ denotes the total number of days.</td>
<td>Tiegs et al. (2005), Bätz et al. (2016), Tiegs et al. (2005)</td>
</tr>
<tr>
<td>Days since Last Flood (DLF)</td>
<td>$\text{DLF} = \text{date(last observation)} - \text{date(last water observation)}$</td>
<td>Vesipa et al. (2017), Townsend (2001)</td>
</tr>
<tr>
<td>Flood Duration (FD) mean, maximum and standard deviation</td>
<td>$\text{FD} = \text{date(start water period)} - \text{date(start land period)}$</td>
<td>Vesipa et al. (2017)</td>
</tr>
<tr>
<td>Dry Period (DP) mean, maximum and standard deviation</td>
<td>$\text{DP} = \text{date(start land period)} - \text{date(start water period)}$</td>
<td>Vesipa et al. (2017)</td>
</tr>
<tr>
<td>Return Interval (RI) mean, maximum and standard deviation</td>
<td>$\text{RI} = \text{date(start water period i+1)} - \text{date(start water period i)}$</td>
<td>Egger et al. (2007), Formann et al. (2014), Vesipa et al. (2017)</td>
</tr>
</tbody>
</table>
differences for completely undisturbed plots. Homogeneity of variance was tested for the FI values at field plots using Levene’s test (Levene, 1961; R function leveneTest, package car, Fox and Weisberg, 2011). If variance homogeneity was not given, data were transformed using the Box-Cox transformation or, if variances were not homogeneous after that, a logarithmic transformation was applied (Levene, 1961; Box and Cox, 1964). For the DLF and the DP standard deviation, no homogeneity of variances was achieved. These indicators were not tested for significant group differences (Supplementary material K).

If significant group differences were found in the ANOVA, the post-hoc test Tukey’s Honest Significant Difference-Test (Tukey’s HSD) was applied (Tukey, 1949; R function TukeyHSD). To test whether the results were consistent when unclassified data were used, we additionally performed Non-Metric Multidimensional Scaling (NMDS) (Kruskal, 1964; Supplementary material M).

3. Results

3.1. Binary water rasters

On average, each pixel in the study area was observed 423 times (Fig. 3a). Pixels were masked due to clouds or shadows approximately 30% of the time. The mean time distance between two observations was 31 days (Fig. 3b) and markedly longer in the period when Landsat data were used (47 days) than in the time period when Sentinel-2 data were used (11 days).

3.2. Flood indicators

All FIs showed a pronounced spatial pattern with clear differences between the active river channel, adjacent gravel banks, and higher, seldomly to never flooded river terraces (Fig. 4, Supplementary material N). The FIs showed values indicating high disturbance intensities in bottlenecks where the river was artificially constrained (at bridges, in areas with groynes). In areas where the river was flowing unconstrained, a variety of disturbance regimes can be observed.

To evaluate the plausibility of the water detection approach, the flooded area was compared to the water volume as measured by a gauge station on the day of the satellite image acquisition. The correlation between the percentage of flooded pixels and the gauge measurements was high (Pearson’s \( r = 0.72, p < 0.001 \), Supplementary material G).

For the calculation of all FIs, the same data were used, and some FI formulas were highly similar. Thus, numerous FIs were strongly correlated (e.g., FF and WFF or FD and DP). Correlations were strongest between indicators that measure similar variables or different statistical measures of the same variable (Table 3; FF and WFF, all DP measures, all FD measures and all RI measures, respectively).

3.3. Habitat classification

Overall, 161 plant species were recorded in 281 plots during the field samplings. Using the isopam algorithm, the field plots were classified on three hierarchical levels (Fig. 5). The classes showed pronounced differences in vegetation layers and ground cover, with an increasing
vegetation cover from the class Water to the class Forest (Fig. 6 b, c; Supplementary material J).

3.4. Relationship between flood indicators and vegetation

All FIs showed pronounced class differences when the values among the plot classes were compared. As a showcase example, FI values per class are shown for WFF, DLF, mean DP, and mean FD (Fig. 7).

The WFF resulted in the highest number of significant class differences (12 out of 15 possible class combinations, Fig. 7a). It showed a continuous decrease from the class Water to the class Forest. The difference between Water and Barren Gravel Bank was pronounced, whereas the other classes were in a similar value range.

The DLF and the DP showed an inverse pattern; the values increased from Water to Forest (Fig. 7 b, c). Here, two groups of classes emerged: The classes Water to Herb and Willow Shrub showed a low mean DP (group median up to 2332 days), whereas the classes Heath and Pine Shrub to Forest had higher values (both median 12841 days).

The class differentiation by mean FD was less pronounced: classes showed a bigger overlap in FD value ranges, but an overall decrease in the FD from Water to Forest was still clearly visible (Fig. 7, d).

The ANOVA results for all 12 FIs showed significant differences

---

Fig. 4. Maps of the Flood Indicators. Shown are: Weighted Flood Frequency (a), Days since Last Flood (b), mean Dry Period (c) and mean Flood Duration (d). Pink areas: No FI value could be calculated, e.g., because the area was never flooded. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
between classes in at least one FI for 13 out of 15 possible class combinations (Supplementary material M). Classes that were often distinguished from other classes were Water and Barren Gravel Bank, and classes that were rarely distinguished from other classes were Herb and Willow Shrub and Forest. The high number of significant class differences generally suggests a high robustness of class differentiation by FIs. The NMDS results were complementary to the findings of the ANOVA (Supplementary material M). Different plot classes were well differentiated in the ordination space. Secondary fitting of the FIs onto the ordination axes determined by the vegetation and environmental data (Supplementary material M).

Both ANOVA and NMDS results confirm the strong relationship between FIs and the observed vegetation patterns.

4. Discussion

4.1. Flood Indicators

Comparison with the gauge data indicated that the combination of satellite data with different spatial resolution is a sound approach to characterize flooding patterns as the relationship between flooded areas and gauge data was similar for all sensors. This is in line with the findings of Guerschman et al. (2011), who reported no significant differences between the flood extents detected by two sensors of different ground sampling distances (MODIS with 250 and 500 m resolution).

Despite the promising gauge measurement validation results, the conducted approach to characterize flood disturbance patterns has limitations. We identified four conceptual challenges which need to be addressed depending on the flood regime, location, biome, and size of the river system.

First, time gaps between satellite observations are often longer than average flooding durations. The corresponding flood omissions are particularly critical if severe floods disturb areas otherwise detected as never flooded (Hamilton et al., 2007; Huang et al., 2014; Mueller et al., 2016). Furthermore, even if a cloud-free satellite image is acquired during a flood event, the likelihood that it represents the peak of a flood event is low. We tried to minimize this problem by using all Landsat 5 to 8 images up to the year 2015 and Sentinel-2 thereafter. This offers the best possible return rate for freely available satellite data while still providing a good spatial resolution. However, even when using all available images, the mean return interval averaged over the whole study period and area is 31 days, and the maximum return interval averaged over the study area is even 506 days (Fig. 3). Both the low frequency of observations and the long return intervals are partly caused by terrain shadows during the winter months, which hamper the sound detection of water-bearing areas. In summary, the mean period between observations might be too short to consistently catch the peak aerial extent of the floodings, but this period notably decreased with the switch from Landsat to Sentinel-2. In the future, this issue might be overcome with the launch of further satellites, for example Sentinel-2C and Sentinel-2 D, which will additionally reduce the overall return interval.

Second, the detection of peak flood extents can be impeded by clouds, particularly when clouds cause precipitation. This is a potential issue here as the given study area has a nivo-pluvial flood regime, and some floods are caused directly by precipitation (Mader et al., 1996). In purely nival systems, where most floods are caused by snow melt, clouds obscuring peak flows might be less problematic.

Third, in densely vegetated stands, a closed canopy can obscure floodings underneath the canopy (Hamilton et al., 2007), potentially leading to a systematic underestimation of floodings in such areas. In our study, this might be the case for plots with dense Salix vegetation close
Fig. 6. Distribution and characteristics of the habitat classes. The results of the classification using isopam are shown in form of (a) a map of the distribution of the different plot classes in the study area, (b) mean vegetation layer covers, and (c) mean ground covers in the different habitat classes.
to the river. The problem of clouds or vegetation obscuring the water might be overcome in future research by the use of SAR instead of optical images (Mahdianpari et al., 2018; Mohammadimanesh et al., 2018; Muro Martín et al., 2020). With the use of SAR, however, mountainous terrain might pose a major challenge.

Fourth, the spatial resolution may, in some cases, be too coarse to capture disturbances to the level of detail needed to establish sound links to the studied vegetation patterns (Pickett et al., 1999; Ozesmi and Bauer, 2002). The used spatial resolution of 10 to 30 m results in mixed water-land pixels along the river boundary, which are always partly misclassified (Tulbure and Broich, 2013). For plots situated close to the river edge, floodings might thus be systematically over- or underestimated. However, the vegetation patches observed in this study are mostly larger than the resolution of the satellite imagery.

Despite these technical challenges, the calculated FIs show a pronounced spatial pattern over the study area, which agreed well with the observed habitat classes and the expected disturbance regimes. We observed high disturbance intensity indicator values in main channels and little to no disturbance farther away from the current river course, which was particularly exemplified by the Weighted Flood Frequency (WFF) (Fig. 4a). This is in line with the finding of higher disturbance close to the main river channel in other studies (Corenblit et al., 2007).

Fig. 7. Boxplots for the comparison of Flood Indicator values in the different habitat classes. Weighted Flood Frequency is dimensionless, all other FIs are given in days. Numbers below boxplots show the number of observations for each FI per class. Brackets above Boxes show significant group differences (Significance level: *: p < 0.05, **: p < 0.01, ***: p < 0.001). Significances are calculated from flooded plots only. For DLF (b) no significances can be calculated as the prerequisites for the ANOVA are violated. Black Boxes show FI results for all field plots. If not all field plots were flooded in the study period, grey boxes are added that show results for pixels that were flooded at least once.

In summary, the satellite-based FIs seem capable of describing the spatial patterns of the disturbance regime. However, it should be noted that the FIs should not be interpreted as “absolute” values but rather as sampling-based indicators which should work fairly well when integrated over several years of observations but would probably not be suitable to, e.g., capture differences in flooding regimes over shorter time periods such as 1–2 years.

4.2. Flood indicators and riparian vegetation communities

When comparing the patterns of the FIs with those of the vegetation communities, some particularly interesting relationships became apparent. The FF and WFF showed a clear trend, with highest values in unvegetated, highly disturbed plots and decreasing values towards zones with higher vegetation covers. This matches the expectation as in zones with high flood frequencies, destructive forces of flooding are dominant (Corenblit et al., 2007). Most of the examined habitat classes showed significant differences in their FF and WFF values (WFF: 12 significant class differences; Fig. 7), indicating that these indicators are well suited to show ecologically relevant differences in the flood regime among classes.

The Days since the Last Flood (DLF) depict the time the terrestrial vegetation has had to develop since the last hydrogeomorphological disturbance. If the disturbance was intense, it depicts the time the vegetation community had to proceed along the successional pathway (Poff and Ward, 1989). We found the highest proportions of herbaceous species and vegetation at highest DLFs. Moreover, we found an increase
in tree cover with increasing DLF (Fig. 6b, Fig. 7b). This supports the assumption that the DLF is a good indicator of habitat development along the successional pathway. However, the indicator is not highly robust as only one flooding (the last in the time series) is considered. Single, and in some cases false, water detections have a large impact on the observed patterns. In our study area, the stripy DLF pattern orthogonal to the river is most likely due to the rim of mountain shadows that were not correctly masked (Fig. 4b). A single false water detection late in the time series can thus result in a critically inadequate DLF evaluation. Moreover, the force of the flooding is not considered, but only disturbances with a high magnitude completely reset ecosystems to primary succession (Pickett et al., 1999). Furthermore, established vegetation is more stable than newly sprouting pioneer vegetation (Surian et al., 2015). If vegetation is already established in a location, short or shallow floodings might not lead to a clearing of the plot, which would be needed for the establishment of a new vegetation community (Surian et al., 2015; Vesipa et al., 2017). Hence, the local vegetation community might be older than suggested by the DLF.

The Dry Period (DP) indicator reflects the time a pixel is consecutively not flooded. During this time, terrestrial vegetation can uninterrupted develop, and the successional pathway proceeds (Vesipa et al., 2017). In our data, the maximum DP values were observed for plots with upcoming woody vegetation (Herb and Willow Shrub, Heath and Pine Shrub and Forest). However, it has to be considered that the satellite-based DP values are not actual measurements and there is a chance that a flooding event has occurred in-between satellite image acquisitions. Hence, the reported values should rather be interpreted as relative differences within the study area as compared to real measurements of time.

The Flood Duration (FD) is the time period a location is continuously flooded. Similar to the FF, it is considered one of the main factors determining plant survival (Hupp and Osterkamp, 1996; Townsend, 2001; Deng et al., 2014; Garsen et al., 2015). The longer the FD, the lower the chance of plant survival (Garsen et al., 2015), mostly due to root anoxia (Vesipa et al., 2017). Townsend (2001) found the FD to be the dominant factor controlling woody vegetation. Considering our root anoxia (Vesipa et al., 2017), this result may as well point out that the full potential of the time series is not yet exploited (s. below).

4.3. Further research

To better represent the magnitude dimension of hydro-geomorphological disturbances, further research could include the flooded area or the water depth to obtain additional information. Bendix (1997) found the flooded area to be related to the flood magnitude as the energy in the system increases with discharge. Flooded areas could directly be calculated using the water detection rasters, which implies the combination of the direct pixel-based approach we followed with additional spatial raster processing steps. Vesipa et al. (2017) describes the flood depth as a good substitute for the flood magnitude. Although the flood depth could be calculated using a DEM as an additional input, flood depth estimations based on DEMs are limited by the usually low temporal resolution of terrain measurement: A mono-temporal DEM will not be able to adequately represent sedimentation and erosion processes of dynamics river systems. To refine satellite-based indices, more applicable approaches for water depth estimation are needed. However, the scale of corresponding satellite missions is currently limiting these approaches for the provided example of a medium-sized mountainous river.

In addition, the movement of sediments is a central determinant of disturbance, and the presented indicators represent this only indirectly. Here, the visibility of moving gravel banks might complement the presented disturbance characterization. Further research could therefore include the detection and monitoring of gravel banks, with an approach similar to that used for water in this study.

The presented FIs already capture habitat differences. However, a perfect differentiation of classes by FIs was not expectable as FIs aggregate over a time span in which one or several habitat shifts may occur. Thus, only the more recent FI data might be predictive of habitat classes while earlier FI data may hamper a clear relationship. The DLF is the only indicator accounting for this in only using the last water detection instead of the whole time series. This is also reflected in its relative importance for class separation. However, the DLF is sensitive to isolated mis-classifications suggesting the need for a more robust design of FIs to account for the higher relevance of recent dynamics (Seger et al., 2017). This is a challenging task because time weights for optimal prediction might differ among classes. However, this issue might be overcome by more effectively exploiting the time series data available. For instance, FIs could be calculated continuously in time, ideally resources in anchoring root systems, and have flexible roots and stems (Karrenberg et al., 2002). The presence of these perfectly adapted species in areas with regular flood regimes supports the hypothesis that flood-adapted species benefit from a regular flood regime.

Regarding the three dimensions of disturbances, namely magnitude, frequency, and size, the suggested FIs clearly focus on frequency (time dimension by satellite time series) and size (spatial dimension by continuous calculation of the FIs over the study site), whereas the magnitude is not fully represented. For example, the shear stress induced by flood events may be notably higher for areas closer to the main channel and during flood events with large amounts of water.
resulting in a probabilistic prediction of the current habitat state for every point in time. This description of the current habitat state may then support the determination of suitable time weights, and, consequently, a more informed prediction of the subsequent state (Gurnell, 2014). In addition, FI values in the close to far surrounding could be considered here to refine and increase the validity of the state estimates. Finally, such a theory-driven framework could be compared to a purely empirical approach, such as Long-Short-Term-Memory models (van Houdt et al., 2020).

Further improvements to the FI analysis might be achieved by considering seasonality in the calculations. For instance, flood events during the growing season likely have a higher impact on the vegetation than those during winter (Garssen et al., 2015). For young seedlings, floods are most harmful in late summer as the plants have already sprouted, but anchoring by roots is not yet fully developed (Vesipa et al., 2017). For grown trees, spring is considered the most fatal season for floods as most root and leaf growth happens during this time (Townsend, 2001). For Salix species, however, spring floods are also favorable as they are used for seed dispersal (Karrasberg et al., 2002; Pickett and White, 2005). Thus, a proper incorporation of the ecological meaning of the timing of flood is challenging but might significantly alter the effectiveness of tracing ecological dynamics.

Finally, imagery filtering algorithms need further refinement. We accounted for clouds, shadows and gravel banks as potential sources of mis-classification. Still, a semi-objective manual revision of resulting water masks was necessary to verify their plausibility. To allow for a more general application of the presented approach, we suggest to increase the robustness of water detection by either improving imagery filters or applying more sophisticated water detection algorithms, e.g. machine-learning based image classification techniques (Bijee and Narasimhamurthy, 2020).

5. Conclusions

In this work, we employed a novel approach to describe flood disturbances over longer time periods using satellite time series-based flood indicators. With this approach, we spatially characterized the hydro-geomorphological disturbance regime of the study area and compared it to the emerging local vegetation communities. Our findings are as follows:

(I) Satellite-derived disturbance indicators are well suited to describe the time and space dimension of disturbances. The time dimension is considered as time series are used and the space dimension as each indicator is specifically calculated for each pixel in the satellite data. The magnitude dimension of disturbances can, however, only indirectly be described as the satellite images provide no direct measure of physical forces.

(II) There was a remarkable agreement between the habitat classification of our field data and the disturbance regime as depicted by our FIs. The FIs showed highest disturbances and the greatest discriminatory power in early successional stages, supporting the established knowledge that disturbance is a main driver of early successional vegetation types. In late successional stages, the FIs indicated generally low disturbance values, and the differences in FI values among habitat classes were often insignificant. Hence, late successional stages are apparently no longer reigned by flood disturbances in the study area.

Our satellite-based approach to measure disturbances is in agreement with long-standing hypotheses and theories in riverine research. As we used globally available data, our approach is easily transferable to other study regions. Its straightforward applicability makes it promising for further use to answer an array of research questions as well as for the monitoring of human-induced changes in disturbance regimes.

CRediT authorship contribution statement

Miriam Herrmann: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. Ephraim Schmidt-Riese: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. Daria Alison Bate: Writing – original draft, Visualization, Data curation. Fabian Kemper: Writing – original draft, Visualization, Data curation. Fabian Ewald Fassnacht: Writing – review & editing, Supervision. Gregory Egger: Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding statement and Acknowledgements

This research did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors. Open Access Funding was provided by Freie Universität Berlin. We thank two anonymous reviewers for their comments. Special thanks go to Miriam Paul and Isabell Becker for their valuable contributions in revising the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2024.112313.

References


M. Herrmann et al.