



Large-scale windthrow on Iturup Island (Kurils, Russia) following the December 2014 storm event revealed by very-high-resolution multispectral imagery and semantic segmentation

Kirill A. Korznikov^{1*} & Tatyana Ya. Petrenko^{2,3}

Kirill A. Korznikov^{1*}
e-mail: korzki@gmail.com

Tatyana Ya. Petrenko^{2,3}
e-mail: petrenkotya@gmail.com

¹ Institute of Botany CAS, Třebon, Czech Republic

² Botanical Garden-Institute FEB RAS, Vladivostok, Russia

³ Sakhalin State University, MSU-SakhalinTech Biodiversity Research Laboratory, Yuzhno-Sakhalinsk, Russia

* corresponding author

Manuscript received: 15.07.2026
Review completed: 10.10.2026
Accepted for publication: 20.10.2026
Published online: 23.10.2026

ABSTRACT

Cyclone disturbances can produce heterogeneous forest loss that is difficult to quantify with medium-resolution satellite products. We mapped windthrow caused by a severe winter extratropical cyclone on southeastern Iturup Island using very-high-resolution (VHR) multispectral imagery and compared it with Global Forest Change data. We trained a compact U-Net on 19 manually labelled 1280×1280-pixel tiles using grouped four-fold spatial cross-validation. The final map contained 10 224 physical windthrow patches covering 798.56 ha, approximately 7 % of forest within the VHR extent. The median patch area was 29.3 m², whereas the largest patch covered 180.45 ha. The ten largest patches contained 54.2 % of mapped disturbance. Windthrow was concentrated at low elevations (~45.5 m) compared with 98.4 m for all forests. Global Forest Change Forest loss layer showed limited spatial agreement with VHR mapping (precision 0.531, recall 0.594, IoU 0.389), despite broadly similar total loss area, partly due to whole-pixel activation from partial disturbance.

Keywords: forest disturbance, Global Forest Change, Kuril Islands, U-Net, VHR imagery, windthrow, winter storm

РЕЗЮМЕ

Корзников К.А., Петренко Т.Я. Крупномасштабные ветровалы на острове Итуруп (Курильские острова, Россия), вызванные циклоном в декабре 2014 года: детекция методом семантической сегментации по мультиспектральным снимкам сверхвысокого разрешения. Воздействие циклонов формирует неоднородные по размеру участки нарушения лесного покрова, которые трудно оценивать по спутниковым данным среднего пространственного разрешения. Мы создали карту ветровалов, вызванных сильным зимним внетропическим циклоном года на востоке центральной части острова Итуруп, используя мультиспектральные снимки сверхвысокого разрешения (VHR), и сравнили результаты с данными среднего разрешения Global Forest Change (GFC). Компактную сеть U-Net обучили на 19 вручную размеченных изображениях размером 1280×1280 пикселей с использованием пространственно сгруппированной четырехкратной кросс-валидации. Итоговая карта содержала 10 224 физических ветровальных участка общей площадью 798,56 га, что соответствует примерно 7 % площади лесов в пределах VHR-покрытия. Медианная площадь ветровального участка составляла 29,3 м², максимальная – 180,45 га. Десять крупнейших участков включали 54,2 % общей площади нарушений. Ветровалы были приурочены к меньшим высотам н.у.м. (~45,5 м против 98,4 м для всех лесов). GFC показал низкое пространственное соответствие VHR-сегментации (precision 0,531; recall 0,594; IoU 0,389), несмотря на сходную общую площадь, что объясняется классификацией целого GFC пикселя 30×30 м даже при частичном нарушении лесного покрова.

Ключевые слова: нарушение лесного покрова, глобальное изменение лесов, Курильские острова, U-сеть, изображения с очень высоким разрешением, вывал, зимний шторм

Wind disturbance is a major component of forest dynamics and can rapidly alter canopy structure, biomass, carbon cycling, regeneration pathways, and habitat conditions (Lindroth et al. 2009, Mitchell 2013). Climate change is also altering disturbance regimes, making the spatially explicit quantification of severe events increasingly important for understanding forest vulnerability and recovery (Seidl et al. 2017). Extratropical cyclones are among the principal sources of extreme winds in the mid-latitudes, and climate-model

projections indicate changes in storm tracks, cyclone intensity, and the spatial extent of damaging winds during the twenty-first century (Priestley & Catto 2022, Little et al. 2023).

The Northwest Pacific is a particularly important region for studying wind disturbance at the southern boreal forests and temperate-boreal transition zone. Previous studies from the Russian Far East and the Insular Sector of the Far East documented severe forest disturbance caused by both tropical and extratropical cyclones, with evergreen coniferous forests

often among the most affected vegetation types (Korznikov et al. 2022, 2023, 2025). On Kunashir Island, immediately south of Iturup, a severe winter extratropical cyclone on 17 December 2014 produced widespread windthrow under a combination of extreme winds and wet snow. The event was followed by extensive secondary spruce mortality, making it one of the best-documented compound forest disturbances in the southern Kurils (Korznikov et al. 2025a). Very-high-resolution (VHR) imagery from Iturup shows a concentrated disturbance pattern with the same regional timing and characteristic windthrow morphology, suggesting that the ecological footprint of the December 2014 storm event extended northward across more than one island.

Reliable quantification of such disturbance depends strongly on spatial resolution. Landsat-based products such as Global Forest Change (GFC) provide globally consistent annual information on forest-cover loss and are indispensable for broad-scale monitoring (Hansen et al. 2013). However, medium-resolution data aggregate heterogeneous canopy conditions within individual pixels and can miss small gaps, partial canopy disturbance, and disturbances embedded within structurally complex forests. These limitations were demonstrated in the southern Kurils, where medium-resolution approaches underestimated or spatially displaced forest disturbance mapped from VHR imagery (Korznikov & Altman 2024, Korznikov et al. 2025b). VHR imagery, by contrast, resolves individual canopy structures and small disturbance patches, and deep learning techniques can exploit both spectral and spatial texture to delineate irregular forest-damage boundaries (Kislov & Korznikov 2020, Kislov et al. 2021).

Two methodological issues remain particularly important for event-scale VHR RGB (red-green-blue channels) +NIR (near-infrared channel) mapping. First, multispectral information may improve discrimination between intact vegetation and mechanically disturbed canopy, but the gain from adding NIR information should be demonstrated rather than assumed. Second, accuracy estimates can be strongly inflated when neighbouring image tiles are divided randomly between training and validation sets. Spatially grouped cross-validation is therefore essential when the aim is to estimate transfer to geographically distinct forest contexts (Roberts et al. 2017). In addition, a model that maximizes pixel overlaps does not necessarily provide an unbiased estimate of disturbed areas, which is often the principal ecological quantity of interest.

In this paper we combine VHR RGB+NIR multispectral imagery, spatially grouped semantic-segmentation validation, elevation analysis, and comparison with GFC forest loss product to quantify the December 2014 windthrow on Iturup Island. Our objectives were to (i) map windthrow with a compact U-Net and test whether RGB+NIR input improves spatial generalization relative to RGB alone; (ii) select an inference and threshold configuration that balances segmentation accuracy with unbiased area estimation; (iii) quantify the area, connectivity, and patch-size structure of the mapped disturbance; (iv) describe its elevational distribution relative to the available forest area; and (v) determine how closely the VHR map is reproduced by medium-resolution annual forest-loss GFC data.

MATERIAL AND METHODS

Study site and forest disturbance event

The study area is located in the south part of Iturup Island on its eastern coast and includes the coastal landscapes around Dobroye Nachalo and Odessky bays, Atsonupuri Volcano, and the vicinity of Lakes Lesozavodskoye and Dobroye. The analysed VHR coverage was approximately 152.6 km² and was selected to encompass the principal concentration of post-cyclone windthrow visible in this part of the island. Unlike Kunashir, where windthrow from the same regional event was distributed across much of the island, the conspicuous Iturup disturbance was strongly concentrated within this southeastern sector (Fig. 1).

The terrain is predominantly low-lying, with steeper relief associated mainly with the slopes of Atsonupuri Volcano. Lowland forests are dominated by conifers *Abies sachalinensis* and *Picea jezoensis*, locally with broadleaf admixture (mainly *Betula ermanii*). Wet and swampy sites include stands of *Alnus hirsuta* and *Salix* spp. Forest dominated by *Betula ermanii* forms a distinct vertical vegetation belt on the volcanic slopes, while shrubland-like communities including *Alnus fruticosa* and *Pinus pumila* occur toward the upper part of the volcano. Broadleaf forests occupy a smaller proportion of the lowland landscape than spruce-fir forest. This vegetation pattern is consistent with the regional forest zonation of the southern Kuril Islands (Krestov 2003, Korznikov et al. 2024).

The mapped disturbance was interpreted as part of the same severe winter extratropical cyclone that affected Kunashir on 17 December 2014 (Korznikov et al. 2025a). We use the well-documented Kunashir observations to characterize the regional event, not as direct meteorological measurements from Iturup. On Kunashir, the cyclone reached a central pressure of 949 hPa and combined extreme easterly to northeasterly winds with precipitation dominated by wet snow; sustained winds exceeded 43 m s⁻¹ and gusts reached 51 m s⁻¹ (Korznikov et al. 2025a). The timing, regional storm trajectory, and morphology of the Iturup windthrow are consistent with the same regional wind-wet-snow disturbance.

Very-high-resolution imagery and reference data

We mapped windthrow using two VHR multispectral WorldView-3 satellite images covering the analysed part of Iturup Island. Both images were acquired on 29 September 2017 and supplied as pansharpened Standard Level 2A products comprising blue, green, red and near-infrared bands at a pixel spacing of 0.30 m. The imagery was radiometrically corrected and atmospherically compensated and projected in WGS 84 / UTM zone 55N (EPSG:32655). Contrast-stretched RGB images were generated only for visual inspection and were not used for model training.

Windthrow was manually delineated in 19 spatially distributed image tiles of 1280 × 1280 pixels, corresponding to approximately 384 × 384 m at the nominal image resolution. Eleven tiles contained windthrow, and 8 were true negative examples without windthrow. Manual delineations were converted to binary masks, with windthrow pixels assigned a value of 1 and all other valid pixels assigned a value of 0.

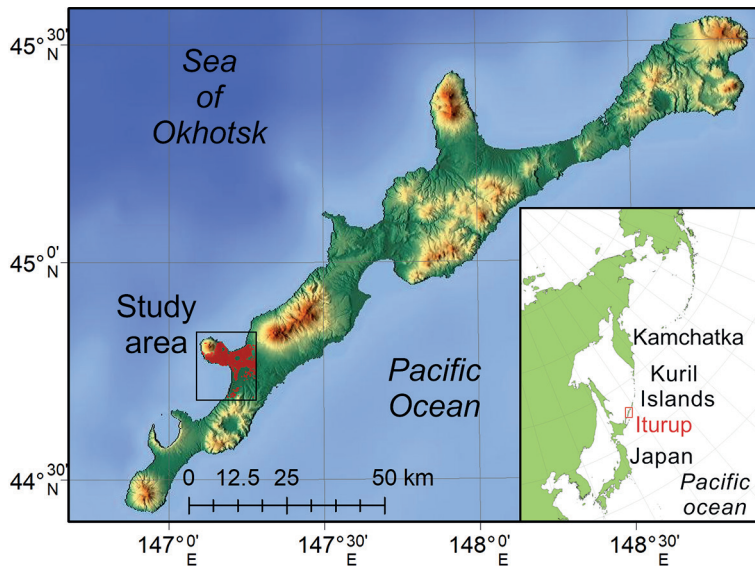


Figure 1 Geographic location of the study area on Iturup Island, Kuril Islands, Northwest Pacific. The inset shows the position of Iturup within the Kuril Archipelago. Red polygons indicate mapped windthrow patches and are not shown to scale

These masks were treated as the sole ground-truth reference for model development and validation.

Spatial cross-validation

We used grouped spatial cross-validation to reduce optimistic accuracy caused by spatial autocorrelation between neighboring image tiles (Roberts et al. 2017). Within each source image, annotated tiles touching along an edge or corner were assigned to the same spatial group and could not be divided between training and held-out data. The resulting groups were distributed among four outer cross-validation folds while approximately balancing source images and positive and negative examples.

For each outer fold, held-out spatial groups were excluded from all stages of model fitting and selection. The remaining groups were further divided into inner training and validation subsets. Channel normalization, model fitting, checkpoint selection, learning-rate adjustment, and early stopping were based only on the inner training and validation data. The outer fold was used only for final evaluation of the selected model. This nested design therefore estimated model transfer to spatially distinct parts of the imagery rather than performance on randomly separated but spatially adjacent samples.

Semantic segmentation model

Windthrow was detected using a compact U-Net semantic-segmentation network based on the encoder-decoder architecture of Ronneberger et al. (2015). The network accepted either three input channels (RGB) or four channels (RGB+NIR). The encoder started with 24 feature channels and increased through 48, 96, 192, and 384 channels over four downsampling levels. Each convolutional block contained two 3×3 convolutions followed by Group Normalization (Wu & He 2018) and rectified linear unit activation. Downsampling used max pooling. Decoder features were bilinearly upsampled, concatenated with corresponding encoder fea-

tures through skip connections, and processed by convolutional blocks. A final 1×1 convolution produced a single windthrow logit layer. Group Normalization was used because GPU memory constraints required a batch size of one.

Models were trained using 512×512 -pixel image patches. For positive tiles, 16 patches were sampled per realization; 60 % were targeted to foreground pixels and 40 % were positioned randomly. Eight random patches were sampled from each negative tile. Data augmentation consisted of rotations by 0, 90, 180, or 270 degrees and independent horizontal and vertical reflections.

Class imbalance was addressed with a weighted binary cross-entropy term. The positive-class weight was calculated from the ratio between background and windthrow pixels in the inner training subset and constrained to the range 1–20. The final loss combined weighted binary cross-entropy and soft Dice loss. Optimization used AdamW (Loshchilov & Hutter 2019) with an initial learning rate of 1×10^{-3} , weight decay of 1×10^{-4} , gradient clipping at 5.0, and FP32 precision. Training continued for a maximum of 60 epochs. The learning rate was halved after four epochs without improvement in validation Dice, and early stopping was applied after a minimum of 15 epochs and 10 consecutive epochs without improvement. The checkpoint with the highest inner-validation Dice was retained for evaluation on the corresponding outer fold. Training and inference were implemented in Python with PyTorch.

RGB versus RGB+NIR comparison

To quantify the contribution of the additional spectral information, we performed paired experiments using either RGB or RGB+NIR input. Both experiments used identical spatial folds, network architecture, training patches, augmentation procedures, optimizer settings, and stopping criteria. The random seed was reset before each paired experiment so that patch selection and stochastic augmentation were directly comparable between spectral configurations. Channel means and standard deviations were calculated exclusively from the inner training data.

Performance was quantified on the spatially independent outer folds using precision, recall, Dice coefficient, and intersection-over-union (IoU). Because estimation of total disturbed area was a principal ecological objective, we also calculated the ratio between predicted and observed windthrow area and the relative area bias.

Model calibration and mapping

After selecting RGB+NIR as the preferred spectral configuration, we optimized the segmentation operating point using out-of-fold probability maps. Probability thresholds from 0.20 to 0.80 in increments of 0.02 and minimum 8-connected-component sizes of 0, 32, 64, 128, 256, 512, and 1024 pixels were evaluated. Calibration was nested: for each target outer fold, we selected the operating point using predictions from the other three folds and then applied it to the target fold.

We used an area-aware selection rule. Among configurations producing a Dice value within 0.01 of the maximum calibration Dice, we selected the configuration with the smallest absolute relative area bias. We also tested 8-transform test-time augmentation and evaluated whether increasing spatial context from 512 to 768 pixels during inference improved transfer. Fine-tuning with 768-pixel patches did not improve any fold-specific checkpoint; the best checkpoint in every fold remained the original 512-trained model. We therefore retained 512-pixel training while using a 768×768 -pixel production inference window with 25 % overlap and no test-time augmentation.

The final production model was trained on all 19 labelled tiles only after the model family, spectral input, inference context, and operating-point strategy had been fixed. Production inference used RGB+NIR input, a probability threshold of 0.66, and removal of connected components smaller than 32 pixels (approximately 2.88 m²). Probabilities from overlapping inference windows were averaged before thresholding.

The binary scene maps were vectorized and geometries from overlapping source scenes were dissolved so that overlap was counted once. Unambiguous false-positive objects were then removed by visual inspection against the original VHR imagery. Because vectorization occasionally divided a single spatially continuous windthrow into several polygon parts, a final connectivity correction merged polygon parts belonging to the same physical disturbance before patch-level statistics were calculated. This step changed the patch number but not the total mapped area.

Elevation analysis

We evaluated the elevational distribution of windthrow using the ASTER Global Digital Elevation Model V.3 (GDEM; approximately 30 m spatial resolution; Abrams et al. 2020). The analysis was restricted to the VHR mapping extent. Forest was defined using the GFC tree-cover layer for the pre-disturbance year 2000, retaining pixels with canopy cover $\geq 20\%$.

The final VHR windthrow polygons were overlaid with the DEM grid, and disturbed area was calculated within each DEM cell. We summarized elevation for (i) all forest within the VHR coverage, (ii) all mapped windthrow, and (iii) windthrow restricted to the $\geq 20\%$ forest mask. Elevation statistics were weighted by represented area and included the mean, median, and 25th and 75th percentiles. We also calculated the proportion of forest area within the VHR coverage affected by mapped windthrow. Rasterization converted windthrow area to within approximately 0.02 % of the corresponding vector area in the elevation-analysis workflow.

Comparison with medium-resolution forest-loss data

We compared the VHR windthrow map with the GFC annual forest-loss product of Hansen et al. (2013). The comparison was restricted to the VHR mapping extent. The primary analysis domain included pixels with tree cover $\geq 20\%$ in the year 2000 and excluded locations already recorded by GFC as forest loss before 2015. The resulting domain covered approximately 11 401 ha.

Because the cyclone occurred in December 2014 and annual forest-loss attribution can be shifted by image acquisition and the timing of detectable spectral change, we evaluated three GFC windows: loss recorded in 2015 alone, cumulative loss in 2015–2016, and cumulative loss in 2015–2017. The VHR polygons were overlaid with the GFC grid, and the amount of VHR windthrow within each GFC cell was calculated as both disturbed area and fraction of cell area.

At the landscape scale, we quantified true-positive, false-positive, and false-negative areas relative to the VHR reference and calculated precision, recall, Dice coefficient, IoU, omission error, commission error, and the ratio of total GFC loss area to VHR windthrow area. We also evaluated the scale dependence of GFC detection among cells intersecting VHR-mapped windthrow. Binomial logistic regression related GFC loss presence to the proportion of each cell occupied by VHR windthrow and, separately, to VHR windthrow area within the cell. Fitted curves were used to estimate the within-cell disturbance associated with a 50 % probability of GFC detection.

Windthrow area and patch-size structure

For the connectivity-corrected VHR vector map, we calculated total mapped area, physical patch number, mean, median, and maximum patch area, and the contribution of the largest patches to total disturbed area. Patch-size classes were summarized as < 10 , 10–100, 100–1 000 m² (0.1 ha), 0.1–1, 1–10, and ≥ 10 ha. All patch areas were recalculated directly from the final edited geometries in EPSG:32655.

RESULTS

Contribution of the NIR band

Adding the NIR band substantially improved spatial generalization. Across the four held-out spatial folds, the RGB model achieved a mean Dice coefficient of 0.607 ± 0.158 and a mean IoU of 0.451. Mean precision and recall were 0.602 and 0.731, respectively. The corresponding RGB+NIR model increased mean Dice to 0.722 ± 0.160 and mean IoU to 0.584, with precision of 0.696 and recall of 0.781. Thus, the additional spectral band increased mean Dice by 0.115 and mean IoU by 0.133.

The improvement occurred in three of the four independent spatial folds. Dice increased from 0.635 to 0.896 in fold 1, from 0.485 to 0.706 in fold 3, and from 0.489 to 0.513 in fold 4. Fold 2 was the only subset in which RGB alone performed better, with Dice decreasing from 0.820 to 0.775 after adding NIR. The substantial among-fold variation indicates that forest structure, disturbance morphology, illumination, and local image context remained important sources of prediction uncertainty despite the overall spectral advantage of RGB+NIR.

Production model performance

Area-aware calibration of the 512-pixel inference configuration produced a nested out-of-fold Dice of 0.790, IoU of 0.653, precision of 0.786, and recall of 0.794. The aggregate predicted area differed from the reference by +1.10 %. Test-time augmentation did not improve performance: Dice decreased to 0.781, and area bias increased to +2.71 %.

Increasing inference context to 768 pixels improved both overlap and area estimation. The final nested evaluation yielded Dice = 0.800, IoU = 0.667, precision = 0.797, and recall = 0.804. The predicted-to-reference area ratio was 1.009, corresponding to an aggregate area bias of +0.92%. Because all four folds retained the original 512-trained checkpoint during the large-context experiment, the improvement was attributable to additional context during inference rather than to successful additional fine-tuning.

Final mapped windthrow and patch structure

After whole-scene inference, vectorization, manual removal of unambiguous false positives (e.g. a few coastal scree slopes and volcanic barren lands) the final map contained 10 224 physical windthrow patches. Total mapped windthrow area was 798.56 ha (7.99 km²) within the analyzed VHR extent. Of this area, 797.17 ha occurred within the GFC-defined forest mask, corresponding to 6.99 % of the 11 398.46 ha of forest within the VHR image extent.

The patch-size distribution was strongly right-skewed. Mean physical patch area was approximately 78.1 m², whereas the median was only 29.25 m². The largest connected patch covered 180.45 ha, equivalent to 22.6 % of total mapped windthrow area. The ten largest physical patches together accounted for 54.2 % of the total area (Fig. 2). Thus, the mapped disturbance combined very numerous small canopy openings with a small number of very large contiguous windthrow patches that dominated landscape-scale area (Fig. 3).

Elevational distribution of windthrow

The elevational distribution of windthrow was shifted toward the lower part of the available forest distribution. Mean elevation of all mapped forest was 98.4 m a.s.l., compared with 45.5 m a.s.l. for windthrow within the same forest mask. The 25th, 50th, and 75th elevation percentiles for all forest were 17, 35, and 87 m, respectively, whereas corresponding values for windthrow were 19, 31, and 49 m. The lower quartiles and medians were therefore similar, but the upper tail differed strongly: 75% of windthrow area occurred below 49 m, while 75% of the available forest area occurred below 87 m. High-elevation forest was consequently underrepresented in the mapped disturbance (Fig. 4).

Comparison with medium-resolution forest loss

The GFC forest loss map showed a pronounced temporal offset relative to the VHR reference. No corresponding loss was recorded in the 2015 GFC layer. When the temporal window was extended through 2016, GFC identified 842.64 ha of forest loss within the analysis domain. Extending the window through 2017 increased detected loss to 886.63 ha, compared to 792.10 ha mapped using VHR imagery (Fig. 5).

However, agreement in total area did not correspond to close spatial agreement. For the 2015-2016 window, 456.12 ha of VHR windthrow overlapped GFC loss, 335.96 ha of VHR disturbance remained undetected, and 386.52 ha were classified as GFC loss outside the VHR windthrow polygons. Precision was 0.541, recall 0.576, Dice 0.558, and IoU 0.387. Although total GFC loss exceeded the VHR

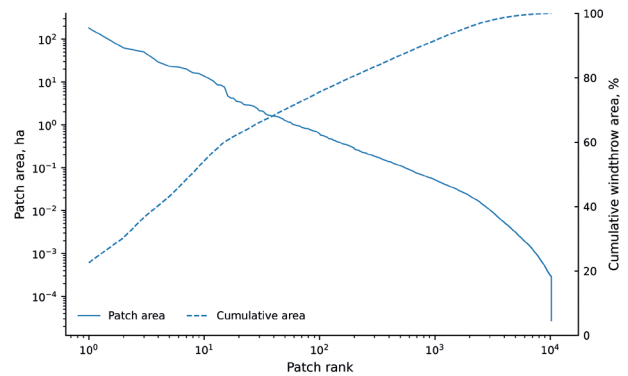


Figure 2 Patch-size classes distribution and cumulative area contribution of mapped windthrow patches. Physical windthrow patches were ranked from largest to smallest. The solid line shows individual patch area as a function of patch rank; both X and Y axes are shown on logarithmic scales. The dashed line shows the cumulative proportion of total mapped windthrow area accounted for by successively smaller patches, referenced to the right Y axis

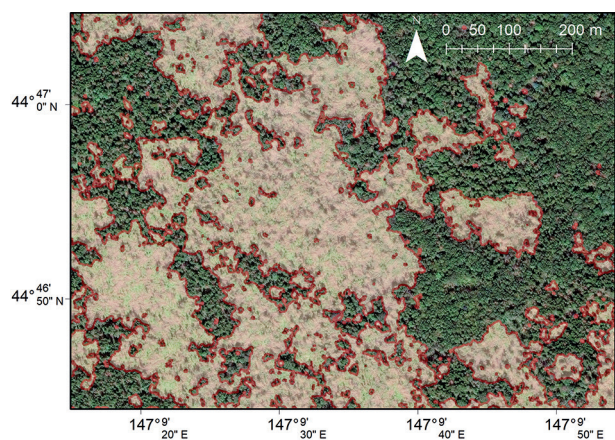


Figure 3 Example of windthrow delineation produced by the U-Net semantic-segmentation model applied to RGB+NIR very-high-resolution satellite imagery. Red outlines indicate identified windthrow patches

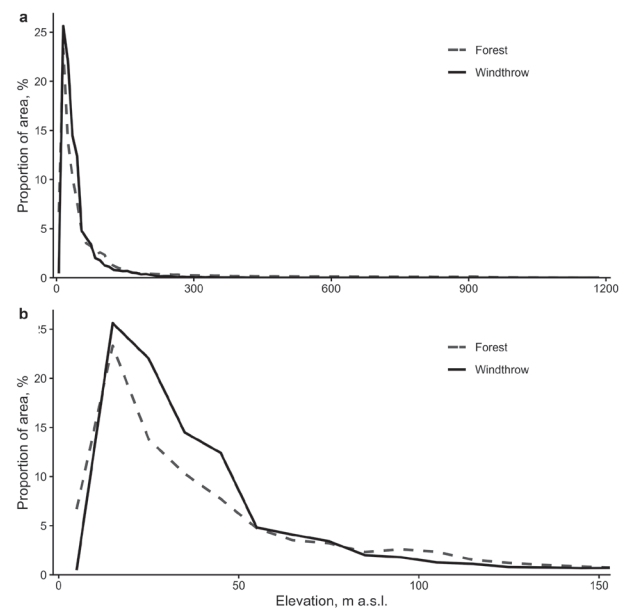


Figure 4 Elevational distribution of forest area and mapped windthrow area within the study area. (a) Distribution across the entire elevational gradient. (b) Detailed distribution within the lower elevation range from 0 to 150 m a.s.l. Values represent the proportion of the total area of each class occurring within each elevation interval

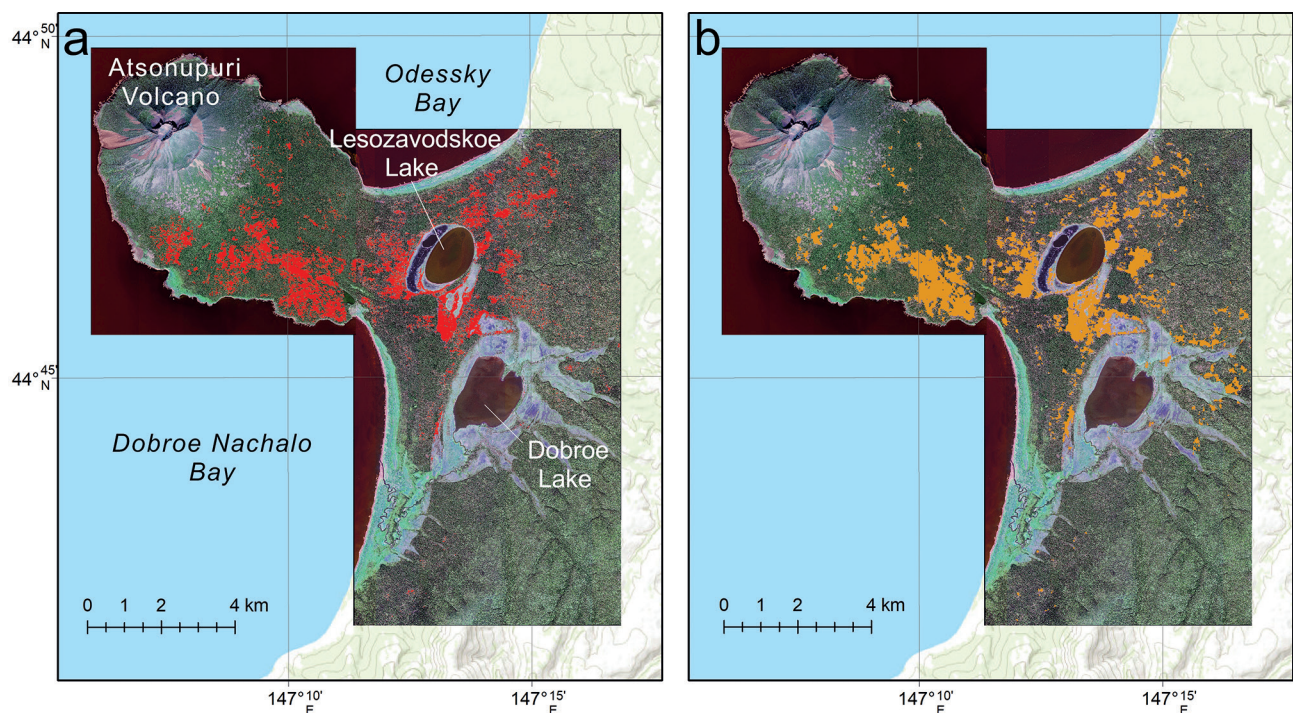


Figure 5 Comparison of windthrow detection from very-high-resolution imagery and the Global Forest Change product. (a) Windthrow patches mapped by the U-Net semantic-segmentation model using RGB+NIR very-high-resolution satellite imagery. (b) Forest-loss areas recorded by the Global Forest Change forest loss layer within the 2015–2017 temporal window

reference by only 6.4 %, approximately 42.4 % of VHR-mapped windthrow area was omitted.

The 2015–2017 temporal window increased overlap to 470.51 ha (Fig. 6a) and recall to 0.594, but commission error also increased. GFC mapped 416.12 ha of loss outside the VHR windthrow polygons, while 321.57 ha of VHR windthrow remained undetected (Fig. 6b). Precision was 0.531, Dice 0.561, and IoU 0.389 (Fig. 6c). Total GFC loss area was 11.9 % greater than the VHR estimate. Thus, a broadly similar landscape-scale total concealed substantial spatial disagreement between the two products.

GFC detection depended strongly on the amount of windthrow within a medium-resolution cell. For 2015–2017, the logistic-regression coefficient for within-cell windthrow fraction was positive (slope = 2.95, SE = 0.038, $P < 0.001$) (Fig. 6c). The fitted model reached a 50% probability of GFC detection when approximately 58.9 % of a cell was occupied by VHR windthrow. The equivalent threshold based on absolute disturbed area was approximately 324 m² per GFC cell (Fig. 7).

DISCUSSION

Additional spectral information improved spatial transfer

The paired experiment provides direct evidence that the NIR channel improved recognition of windthrow in spatially independent image subsets. Because RGB and RGB+NIR models used the same folds, architecture, patch-sampling scheme, augmentation, and random seeds, the mean increase in Dice from 0.607 to 0.722 is unlikely to reflect stochastic training differences alone. The additional NIR band most likely provided information complementary

to visible colour and texture, particularly where windthrow altered canopy continuity, exposed woody material or understory, and changed the amount and condition of photosynthetically active vegetation.

This result is consistent with the broader value of convolutional neural networks for vegetation and forest-disturbance mapping from high-resolution optical imagery (Kattenborn et al. 2021) and extends earlier windthrow work that relied predominantly on RGB VHR data (Kislov & Korznikov 2020, Kislov et al. 2021). The remaining fold-to-fold variation is equally informative. One fold performed better with RGB alone, and even the final RGB+NIR model varied markedly among spatial subsets. Spectral information therefore improved transfer but did not eliminate landscape heterogeneity.

The grouped validation design is central to this interpretation. Randomly separating neighbouring tiles would allow highly similar canopy structure, illumination, and disturbance morphology to occur simultaneously in training and test data, producing an accuracy estimate that is less relevant to geographical transfer. Spatially structured cross-validation has been recommended for exactly this reason in ecological modelling (Roberts et al. 2017). The nested out-of-fold estimate reported here should therefore be viewed as a more conservative measure of expected performance in unsampled parts of the same VHR image domain.

Spatial context improved segmentation without additional training

The second clear methodological result was the benefit of larger inference context. Increasing the inference window from 512 to 768 pixels raised nested Dice from 0.790 to

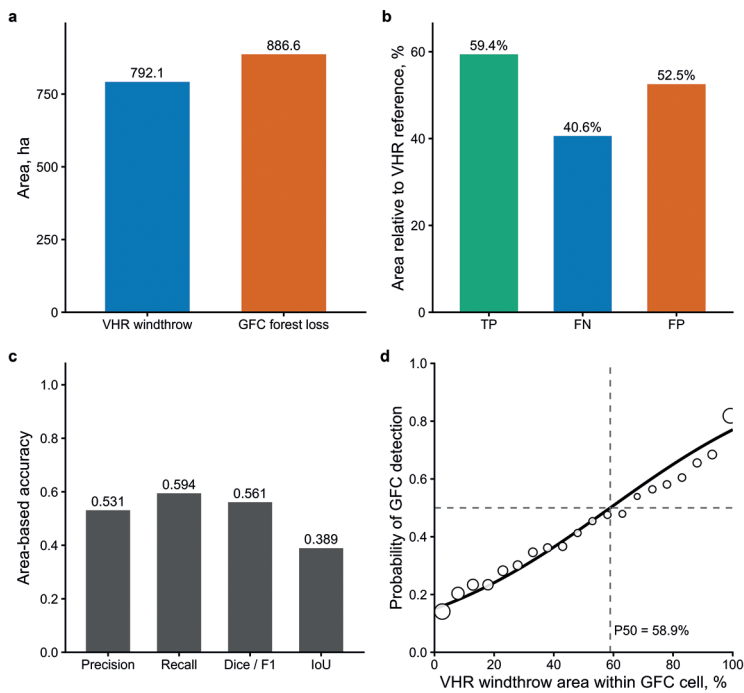


Figure 6 Area-based comparison between very-high-resolution (VHR) windthrow mapping and the Global Forest Change (GFC) forest-loss layer for the 2015–2017 temporal window. (a) Total mapped area of VHR windthrow and GFC forest loss within the common analysis domain. (b) Spatial agreement expressed as true-positive (TP), false-negative (FN), and false-positive (FP) area relative to the VHR reference. (c) Area-based accuracy metrics for GFC detection, including precision, recall, Dice/F1, and intersection-over-union (IoU). (d) Logistic relationship between the proportion of a GFC cell occupied by VHR-mapped windthrow and the probability of GFC detection. Open circles show observed detection frequencies by windthrow-coverage class, the solid line shows the fitted logistic model, and dashed lines indicate the 50% detection probability and the corresponding windthrow-coverage threshold (P50 = 58.9%)

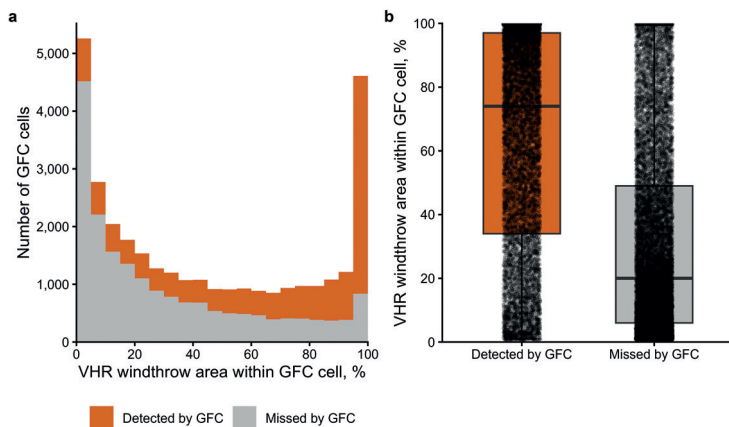


Figure 7 Relationship between very-high-resolution (VHR) windthrow extent and detection by the Global Forest Change (GFC) product for the 2015–2017 temporal window. (a) Distribution of GFC cells according to the proportion of each cell occupied by VHR-mapped windthrow, separated into cells detected by GFC and cells missed by GFC. (b) Distribution of the proportion of VHR windthrow area within GFC cells for detected and missed cells. Boxplots show the median and interquartile range; points represent individual GFC cells

0.800 and reduced area bias from +1.10 % to +0.92 %, whereas additional fine-tuning at the larger window size provided no improvement. At 0.30-m resolution, the two windows span approximately 154 and 230 m, respectively. This scale is ecologically meaningful because windthrow

is not only a pixel-level spectral feature: patch orientation, internal texture, boundaries, neighbouring intact canopy, and the spatial coherence of fallen trees are expressed over tens to hundreds of metres.

The result also illustrates a practical point for small VHR training datasets. More optimization is not necessarily the best route to better generalization. Here, the learned local features were already useful, but presenting them within a wider spatial context improved interpretation of patch structure. Conversely, test-time augmentation added computational complexity without improving Dice or area bias and was therefore excluded from production mapping.

Windthrow was concentrated in low-elevation forest

The Iturup disturbance was disproportionately concentrated in the lower part of the forested elevational distribution. The difference was expressed most clearly in the upper tail: 75 % of mapped windthrow occurred below 49 m a.s.l., compared with 87 m for 75 % of all forest in the VHR coverage. This pattern resembles the earlier Kunashir analysis, where cyclone windthrow was also concentrated in low-elevation forest and declined at higher elevation forests (Korzniakov et al. 2025a).

Elevation itself should not be interpreted as a single causal driver. Within the Iturup study area, elevation covaries with forest composition and terrain: spruce-fir forests dominate much of the lowland, whereas stone-birch forest becomes more important on the slopes of Atsonupuri. Differences among tree species in crown architecture, rooting, wood properties, and winter foliage can strongly influence storm mortality (Barrere et al. 2023), and the Kunashir study showed that evergreen conifer-dominated forests were especially susceptible to the same regional winter event (Korzniakov et al. 2025a). Our Iturup analysis is descriptive and does not separate vegetation effects from wind exposure, slope, or other terrain variables. Nevertheless, it identifies a clear landscape position in which most mapped forest disturbance occurred.

Windthrow dominated by a few large patches

The resulting windthrow distribution shows that patch number and disturbed area describe different dimensions of the event. Fine-scale canopy openings were extremely numerous, yet landscape-scale impact was controlled by a small number of large, contiguous blowdown areas. Similar dominance of large disturbance patches has been reported for severe cyclone damage elsewhere in the Kuril-Sakhalin region (Korzniakov et al. 2022, 2023). VHR imagery is particular-

ly useful in this context because it can retain both ends of the disturbance spectrum instead of forcing small gaps and large blowdowns into the same coarse grid representation.

Similar total GFC area masked substantial spatial error

The GFC comparison demonstrates why total disturbed area alone is an insufficient validation criterion. By 2015–2017, GFC mapped 886.6 ha within the analysis domain compared with approximately 792 ha in the VHR reference, a difference of only about 12 %. Yet less than 60 % of the VHR disturbance area overlapped GFC loss, precision was only 0.531, and IoU was 0.389. Omission of VHR windthrow was therefore partly offset by GFC loss outside the VHR disturbance, giving a superficially similar total for spatially different maps.

This does not imply that GFC performs poorly for its intended purpose. The Hansen et al. (2013) product was designed for globally consistent detection of stand-replacing forest-cover loss at Landsat scale, not for event-specific delineation of subpixel windthrow geometry. The Iturup comparison instead quantifies the scale mismatch that arises when such a product is interpreted as a detailed map of a particular disturbance. The fitted detection curve makes this mismatch explicit: approximately 59 % of a GFC cell, or about 324 m² in the analysed grid, had to be occupied by VHR windthrow before the probability of GFC detection reached 0.5.

The temporal offset provides the second warning. The December 2014 event produced no corresponding GFC loss in the 2015 layer, while apparent agreement increased only when 2016 and 2017 were added. Annual forest-loss products can therefore lag abrupt disturbance because detectable spectral change depends on image acquisition, cloud conditions, vegetation phenology, and the timing of canopy collapse or exposure. Medium-resolution time series remain highly valuable for regional monitoring (Hansen et al. 2013, Chen et al. 2021), but local estimates of heterogeneous windthrow require careful scale-aware validation (Korznikov et al. 2025b). This conclusion agrees with earlier work from Kunashir showing substantial underestimation or spatial misrepresentation of forest disturbance in medium-resolution imagery (Korznikov & Altman 2024).

Regional interpretation and limitations

The spatial pattern on Iturup is consistent with a northern expression of the same December 2014 winter storm disturbance documented on Kunashir (Korznikov et al. 2025a). The regional interpretation is supported by timing, cyclone trajectory, and similar windthrow patterns, but direct meteorological attribution on Iturup remains less secure because the detailed wind, precipitation, and temperature measurements currently available for this event are from Kunashir. The meteorological values reported here should therefore be interpreted as characterization of the regional event, not as local Iturup station observations.

Several additional limitations should be considered. First, the mapped area is restricted to the available VHR coverage and should not be interpreted as a complete island-wide estimate of Iturup windthrow. Our visual interpretation of later VHR imagery suggests that the same

cyclone also caused recognizable, although less extensive, forest disturbance farther south on Iturup, particularly on the northern slope of Berutarube Volcano and the southern slopes of the Lvinaya Past ("Lion's Maw") volcanic somma. In both areas, windthrow affected mixed conifer-broadleaved and evergreen conifer forests. Second, manual removal of unambiguous false positives introduces an expert-quality-control step that should be retained in any reproducible production workflow but separated conceptually from independent model validation. Third, the approximately 30-m DEM supports landscape-scale elevation comparisons but cannot resolve microtopographic exposure. Finally, the 2015–2017 GFC window may include some forest losses unrelated to the December 2014 event, so commission relative to the VHR map should not be interpreted exclusively as GFC classification error. Visual inspection of later VHR images did not reveal clear evidence of additional forest loss caused by subsequent cyclones, geomorphological processes, or anthropogenic disturbance within the study area. Some of the additional GFC loss may instead represent delayed mortality of trees that were damaged, but not immediately killed, by the 2014 cyclone. This cannot be resolved confidently from satellite imagery alone without field observations. Moreover, because more than a decade has now passed since the disturbance, retrospective field identification of the original causes of individual mortality patches would be increasingly difficult.

Despite these constraints, the combination of spatially independent segmentation validation, area-aware calibration, connectivity-corrected patch mapping, elevation analysis, and explicit cross-scale comparison provides a robust event-scale description of the mapped disturbance. It also establishes a reproducible framework for comparing VHR forest-damage maps with global forest-change products without assuming that agreement in total area implies agreement in spatial pattern (Korznikov et al. 2025b).

CONCLUSION

VHR multispectral imagery and semantic segmentation revealed extensive and strongly aggregated windthrow in southeastern Iturup associated with the December 2014 regional storm. Adding NIR information substantially improved spatial generalization relative to RGB alone, and increasing inference context to 768 pixels provided the best area-aware performance without additional fine-tuning. The final connectivity-corrected map contained 10,224 physical patches covering 798.56 ha, with more than half of the total area concentrated in the ten largest patches.

The disturbance was concentrated in the lower part of the forested elevational distribution, while comparison with GFC showed pronounced scale and timing effects. GFC produced a broadly similar total loss area over an extended 2015–2017 window but reproduced the VHR spatial pattern only weakly. Event-scale estimates based only on medium-resolution forest-loss totals can therefore obscure both omission of fine windthrow and commission outside the mapped event footprint. For heterogeneous cyclone disturbance in complex forest landscapes, spatially validated VHR mapping provides a substantially more informative basis for ecological interpretation.

ACKNOWLEDGEMENTS

We thank Dmitry Kislov for his contribution on the Python code development in the early stages of this study. The work of Kirill Korznikov was supported by the long-term research development project of the Czech Academy of Sciences RVO 67985939. The work of Tatyana Petrenko was supported by the scientific framework of the Botanical Garden-Institute FEB RAS (122040800089-2) and as part of the state assignment of Sakhalin State University (FEFF-2026-0010).

LITERATURE CITED

- Abrams, M., R. Crippen & H. Fujisada 2020. ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD). *Remote Sensing* 12:1156.
- Barrere, J., B. Reineking, T. Cordonnier, N. Kulha, J. Honkaniemi, M. Peltoniemi, K.T. Korhonen et al. 2023. Functional traits and climate drive interspecific differences in disturbance-induced tree mortality. *Global Change Biology* 29:2836–2851.
- Chen, N., N.-E. Tsendbazar, E. Hamunye, J. Verbesselt & M. Herold 2021. Sub-annual tropical forest disturbance monitoring using harmonized Landsat and Sentinel-2 data. *International Journal of Applied Earth Observation and Geoinformation* 102:102386.
- Hansen, M.C., P.V. Potapov, R. Moore, M. Hancher, S.A. Turubanova, A. Tyukavina, D. Thau, et al. 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342:850–853.
- Kattenborn, T., J. Leitloff, F. Schiefer & S. Hinz 2021. Review on convolutional neural networks (CNN) in vegetation remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 173:24–49.
- Kislov, D.E. & K.A. Korznikov 2020. Automatic windthrow detection using very-high-resolution satellite imagery and deep learning. *Remote Sensing* 12:1145.
- Kislov, D.E., K.A. Korznikov, J. Altman, A.S. Vozmishcheva & P.V. Krestov 2021. Extending deep learning approaches for forest disturbance segmentation on very high-resolution satellite images. *Remote Sensing in Ecology and Conservation* 7:355–368.
- Korznikov, K. & J. Altman 2024. Pitfalls of forest damage detection using medium-resolution imagery: A response to the paper of Karpov et al. (2024). *Forest Ecology and Management* 569:122073.
- Korznikov, K., D. Kislov, J. Dolezal & J. Altman 2023. Poleward migration of tropical cyclones induced severe disturbance of boreal forest above 50 degrees. *Science of the Total Environment* 890:164376.
- Korznikov, K., D. Kislov, J. Dolezal & J. Altman 2025a. Climate-linked cascading disturbances drive conifer forest decline in a boreal-temperate ecotone: evidence from Kunashir Island (Northwest Pacific). *Forestry: An International Journal of Forest Research* 99(3):cpaf078.
- Korznikov, K., D. Kislov, J. Dolezal & J. Altman 2025b. Hidden forest loss: challenges in detecting wind- and insect-driven forest disturbances with global forest change land-sat-based products in mixed southern boreal forests. *Remote Sensing Applications: Society and Environment* 40:101798.
- Korznikov, K., D. Kislov, J. Dolezal, T. Petrenko & J. Altman 2022. Tropical cyclones moving into boreal forests: Relationships between disturbance areas and environmental drivers. *Science of the Total Environment* 844:156931.
- Krestov, P.V. 2003. Forest vegetation of easternmost Russia (Russian Far East). In: *Forest Vegetation of Northeast Asia* (J. Kolbek, M. Srutek & E.O. Box, eds), pp. 93–180, Springer, Dordrecht.
- Lindroth, A., F. Lagergren, A. Grelle, et al. 2009. Storms can cause Europe-wide reduction in forest carbon sink. *Global Change Biology* 15:346–355.
- Little, A.S., M.D.K. Priestley & J.L. Catto 2023. Future increased risk from extratropical windstorms in northern Europe. *Nature Communications* 14:4434.
- Loshchilov, I. & F. Hutter 2019. Decoupled weight decay regularization. *International Conference on Learning Representations (ICLR)*. Available at: <https://arxiv.org/abs/1711.05101>.
- Mitchell, S.J. 2013. Wind as a natural disturbance agent in forests: a synthesis. *Forestry* 86:147–157.
- Priestley, M.D.K. & J.L. Catto 2022. Future changes in the extratropical storm tracks and cyclone intensity, wind speed, and structure. *Weather and Climate Dynamics* 3:337–360.
- Roberts, D.R., V. Bahn, S. Ciuti, M.S. Boyce, J. Elith, G. Guiller-Arroita, S. Hauenstein et al. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* 40:913–929.
- Ronneberger, O., P. Fischer & T. Brox 2015. U-Net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 (Lecture Notes in Computer Science, vol. 9351)* (N. Navab, J. Hornegger, W.M. Wells & A.F. Frangi, eds), pp. 234–241, Springer, Cham.
- Seidl, R., D. Thom, M. Kautz, D. Martin-Benito, M. Peltoniemi, G. Vacchiano, J. Wild, D. Ascoli, M. Petr, J. Honkaniemi, M.J. Lexer, V. Trotsiuk, P. Mairota, M. Svoboda, M. Fabrika, T.A. Nagel & C.P.O. Reyer 2017. Forest disturbances under climate change. *Nature Climate Change* 7:395–402.
- Wu, Y. & K. He 2018. Group Normalization. In: *Computer Vision – ECCV 2018 (Lecture Notes in Computer Science, vol. 11217)* (V. Ferrari, M. Hebert, C. Sminchisescu & Y. Weiss, eds), pp. 3–19, Springer, Cham.