21 Spatial data science: extracting knowledge from geo-environmental data +

22 Virtual Representation of Forests: Methods and Applications

Convenors session 21: Marj Tonini (UNIL), Tom Beucler (UNIL), David Ginsbourger (UNIBE), Fabian Guignard (UNIBE), Amanda Burton (Agroscope), Cédric Métraux (swisstopo).
Convenors session 22: Arnadi Murtiyoso, Amanda Mathys, Monika Niederhuber, Verena Griess

Swiss Geocomputing Centre

TALKS:

21.1 Carreaud A., Deillon Y., Gressin A.: Automatic determination of land objects – Use of oblique drone images

21.2 Fol C.R., Murtiyoso A., Griess V.C.: Inspection of Forest Point Cloud in Virtual Reality

21.3 Hari C., Davin E., Fischer M.: Combining future projections of land-use and climate change to assess their impact on biodiversity

21.4 Schmutz L., Thao S., Vrac M., Mariethoz G.: A multivariate approach to combine general circulation models using graph cuts

21.5 Tonini M., Trucchia A., Izadgoshasb H., Fiorucci P.: Comparison of the performance of machine learning models for wildfire susceptibility mapping

21.6 Weber M.S., Griess V.C.: Visualizing future forest changes in VR

21.7 Zermatten V., Santacroce N., Gros-Daillon J., Tuia D., Ackermann N.: Characterization of the Uncertainty in Land Cover Predictions near Train Tracks

POSTERS:

P 21.1 Amaya M., Linde N., Meles G.: Bayesian inversion using adaptive sequential Monte Carlo combined with surrogate modelling

P 21.2 Eiholzer L., Tonini M., Losapio G.: Modelling the Impact of Glacier Retreat on Vegetation Dynamic

P 21.3 Friedli L., Linde N., Ginsbourger D., Doucet A.: Sequential Monte Carlo for posterior risk assessment

P 21.4 Jouvet G.: Generative adversarial network for paleo glacial landscape reconstruction

P 21.5 Levy S., Laloy E., Linde N.: Combining inverse autoregressive flows with deep generative networks for improved efficiency and scalability of inverse problems

P 21.6 Poloni A., Tonini M., Lambiel C.: Advanced spatial learning technique for automatic mapping of geomorphological features in alpine periglacial environment

P 21.7 Jingjing Wang Gang Chen, Michel Jaboyedoff , Marc-Henri Derron: Mountain Landslide susceptibility assessment by Multi-sampling strategy combining optimization of BiLSTM algorithm

21.1
Automatic determination of land objects - Use of oblique drone images

Antoine Carreaud[1], Yves Deillon[1], Adrien Gressin[1]

[1] *University of Applied Sciences Western Switzerland (HES-SO / HEIG-VD), Insit Institute - (antoine.carreaud, yves.deillon, adrien.gressin)@heig-vd.ch*

The determination of land objects and their updating for the different cadastre plans or BIM projects is still carried out with significant human resources: by field surveys and/or manual digitalization from orthoimages or laser surveys. Such operations are therefore expensive and time-consuming. In this context, we propose to study the possibilities of automating the recognition of land objects based on drone images using artificial intelligence (AI). Today, the use of orthoimages in AI is widely developed, however, it has limitations such as the determination of objects located in hidden areas (under the trees, at the foot of buildings...).

Futhermore, aerial oblique images are already widely used in areas like cities for complex 3D reconstruction and provide a large amount of information, which is no longer available in the orthoimage. Therefore, we propose to use such oblique images in a prototype project to automate the periodic updating of land cover objects in the Swiss official cadastral survey. This study was conducted in collaboration with the canton of Vaud (Switzerland).

For this study, we used 1700 nadir and oblique images, acquired with the fixed-wing drone Ebee, equipped with the SODA 3D sensor. Images have been acquired with an average resolution of 4 cm, over a 210 hectare area of the city of Lausanne. To automatically label the training images, the classes of the objects concerned (buildings, vegetation, hard surfaces, railway, water basins), from an existing cadastral plan, have been projected onto the oriented images (nadir and oblique). Then, the Deeplab v3+ convolution network has been trained to perform semantic segmentation on the drone images (Figure 1 shows an example of the inference result on an image from the validation set, and Figure 2 the resulting labelled 3D point cloud). Finally, we use the predictions in the form of masks to perform a 3D reconstruction by class on the validation areas of our dataset, using a Structure-From-Motion algorithm.

In these areas, we obtain a semantic accuracy of 94% with a recall rate of 91% on the 3D point cloud. Such precision is not yet sufficient to update the land cover objects of the Swiss official cadastral survey in a fully automatic way, but it allows us to imagine interesting perspectives such as detection of changes, semi-automatic updating methods, or land use statistics.

Figure 1. Inference on one validation image (batiment = building, dur = hard surface, vert = vegetation).

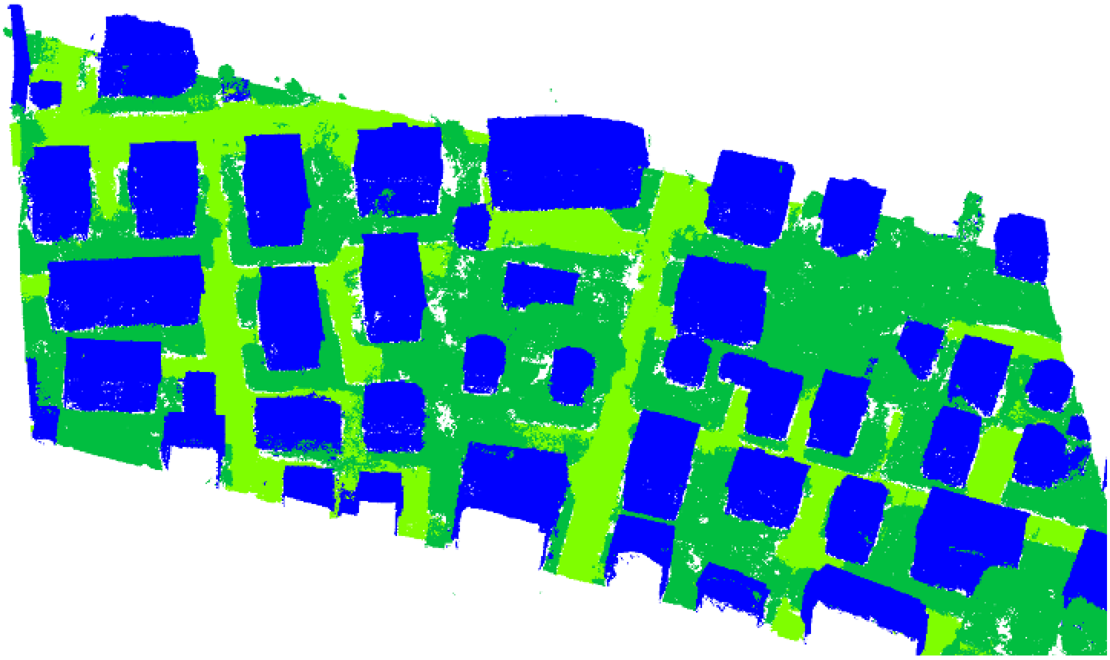

Figure 2. The resulting labelled 3D point cloud.

21.2
Inspection of Forest Point Cloud in Virtual Reality

Cyprien R. Fol*, Arnadi Murtiyoso, Verena C. Griess

Forest Resources Management, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zurich, Switzerland (cyprien.fol, arnadidhestaratri.murtiyoso, verena.griess)@usys.ethz.ch

1) Introduction
Forest digitalisation opened up a completely new interpretation of forest ecology based on direct measurements in the 3D digital twin rather than using the classical DBH (diameter at breast height) centred approach that does not include any information on the forest architecture (Lines et al. 2022). Although the processing of 3D forest data has already allowed researchers to increase their understanding of different complex process such as radiative transfer modelling (Calders et al. 2018) and canopy microclimate (Zellweger et al. 2019), very little research has been done to enhance the visualisation and inspection of forests 3D data. In fact, the classical approach still involves computer software that display the 3D model on a 2D monitor which makes difficult the manipulation of large dataset. For this reason, VR (Virtual Reality) is a great alternative to facilitate the forest inspection by immersing the user inside the 3D model and allowing him to explore rapidly the virtual forest using the several VR motions options, i.e., walking and teleporting. In this research, a VR application was developed to turn the task of 3D forest data inspection into a more user-friendly and pleasant experiment.

2) Methods
To begin with, the data type used in this study is point clouds because they represent most 3D data in the field of forestry. By reviewing the different state-of-the art solution for virtual forest visualisation, two projects stood out from the rest. First, Michael Hertz developed a GitHub repository called PointCloudXR, which is an open-source Unity3D application specifically designed for forest 3D data. In addition, the user can interact with the points, e.g., for distance measurements or removing noise. However, this application only handles point clouds with a maximum of 10 to 15 million points, which is adequate to work with single tree point clouds but far from sufficient for forest point clouds made of around 1 billion points. Nevertheless, a study from the cultural heritage field written by Kharroubi et al., 2019, achieved to visualise 2.3 billion points in real time and continuously in VR. These 2 projects share a common aspect: they both employ Unity and the GPU (Graphics Processing Unit) computing power to optimise the rendering process. Hence, we opted for a similar approach

3) Results

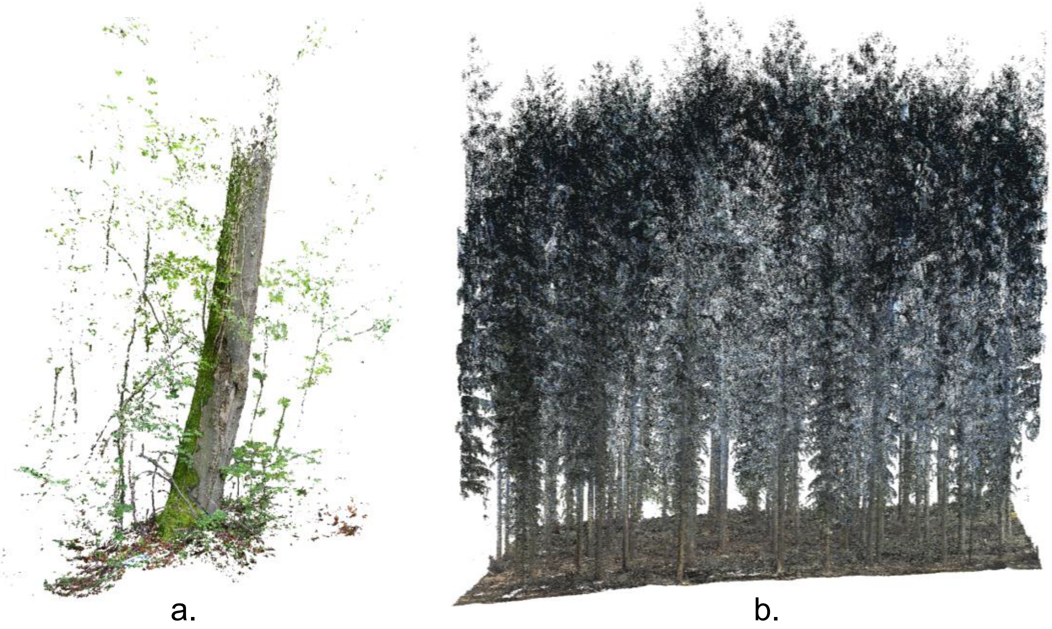

Figure 1: a) single tree point cloud of 27 314 314 points. b) forest point cloud of 28 000 000 points.

4) Conclusion
In conclusion, VR gives users the opportunity to be immersed in the 3D model, allowing a more intuitive and thorough inspection of virtual forest. Furthermore, VR has the potential to be used for measuring directly forest parameters such as DBH (diameter at breast height) or total height of a tree (Mikael Hertz, accessed: 08-07-2022). The addition of this features will be considered for the further development of our application.

REFERENCES

Calders K, Origo N, Burt A, Disney M, Nightingale J, Raumonen P, Åkerblom M, Malhi Y, Lewis P. 2018. Realistic Forest Stand Reconstruction from Terrestrial LiDAR for Radiative Transfer Modelling, Remote Sensing; 10(6):933.

Kharroubi, A., Hajji, R., Hajji, R., Billen, R., & Poux, F. 2019. Classification and integration of massive 3D points clouds in a virtual reality (VR) environment. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XLII-2/W17. 165-171

Lines, E. R., Fischer, F. J., Owen, H. J. F., & Jucker, T. 2022. The shape of trees: Reimagining forest ecology in three dimensions with remote sensing. Journal of Ecology, 110, 1730– 1745.

Mikael Hertz. GitHub Repository PointcloudXR. https://github.com/mhzse/PointcloudXR. Accessed: 08-07-2022.

Zellweger, F., De Frenne, P., Lenoir, J., Rocchini, D., & Coomes, D. 2019. Advances in microclimate ecology arising from remote sensing. Trends in Ecology & Evolution, 34(4), 327–341.

21.3
Combining future projections of land-use and climate change to assess their impact on biodiversity

Chantal Hari[123], Édouard Davin[123], Markus Fischer[34]

[1] *Wyss Academy for Nature at the University of Bern, University of Bern, Kochergasse 4, CH-3011 Bern (chantal.hari@wyssacademy.org)*
[2] *Climate and Environmental Physics, Physics Institute, University of Bern, Sidlerstrasse 5, CH-3012 Bern*
[3] *Oeschger Centre for Climate Change Research, Hochschulstrasse 4, CH-3012 Bern*
[4] *Institute of Plant Sciences, Plant Ecology, Altenbergrain 21, CH-3013 Bern*

Biodiversity loss, land degradation, and climate change are acknowledged environmental challenges faced by humanity (Newbold 2018; Rosenzweig et al. 2008). Climate change's impacts on biodiversity will likely intensify in the future (Pereira et al. 2020). Additionally, land use is a current key stressor of biodiversity (IPBES 2019). Generally, human activities including land-use changes are the cause of global change, thus, future projections of biodiversity impacts need to include both, climate change and land-use change. Here we aim to combine climate change and land-use change scenarios to assess their impacts on biodiversity.

Extensive research has been made on mapping and projecting the vulnerability of multiple species based on different climate mitigation scenarios or warming levels (e.g. Jenkins et al. 2021; Seddon et al. 2021; Strassburg et al. 2012; Thomas et al. 2004; Warren et al. 2013). What is lacking, is the inclusion of land-use trajectories in these projections (Cabral et al. preprint.). Hof et al., (2018) made a first step to overcome these deficiencies. They evaluated the potential future impacts of climate and land-use change on global species richness of terrestrial vertebrates. They used climate and land-use impact projections from the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) under a low-and high emission scenario. However, for both emission scenarios, they used the same land-use projections. This study aims to go further and combine future climate scenarios and a matrix of integrated assessment modeling (IAM) derived land-use projections to assess the impact on biodiversity through the application of species distribution modeling.

Species distribution data will be obtained for the world's amphibians, mammals, and reptiles from the IUCN Red List of Threatened Species as expert range maps (IUCN 2022). An ensemble model will be built to combine the information from individual models fitted with different modeling techniques at a 0.5° resolution. Namely, General Linear Models (GLM), General Additive Models (GAM), and the Random Forest (RF) algorithm will be used. For historical climate data, the W5E5 v.2.0 data will be used, which is a global dataset at 0.5° resolution covering the period 1979-2019, made available through ISIMIP (Lange et al. 2021). For future climate data, this study will use bias-corrected global scenarios produced by ISIMIP phase 3b (Lange & Büchner 2021). ISIMIP3b provides future scenarios for three Representative Concentration Pathways (RCPs) combined with Shared Socioeconomic Pathways (SSPs), SSP1-2.6, SSP3-7.0, SSP5-8.5, and five CMIP6 global climate models (GCMs), GFDL-ESM4, IPSL-CM6A-LR, MPI-ESM1-2-HR, MRI-ESM2-0, UKESM1-0-LL.

The Land Use Harmonization dataset v2 (LUH2; see Hurtt et al. 2020) will be used to include future projections of land-use change. The probabilities of occurrence from the species distribution models will then be filtered by the land-use data to determine the implications for biodiversity among different scenarios. LUH2 reconstructs and projects changes in land use among 12 categories. To match the species' habitat preferences, data from IUCN Habitat and Classification Scheme for each species will be downloaded and then mapped onto the 12 land-use types represented in the LUH2 dataset according to the conversion table from Carlson et al. (2022). The land-use data is then used to refine the climatic envelope and filter out regions, where species cannot survive based on the land-use projections.

This approach then allows to quantify the uncertainties between different climate and land-use scenarios and combinations of both and provides quantitative information on the effects of also considering land-use change projections when running climate-driven species distribution models to determine the impact of future climate change on biodiversity.

REFERENCES

Carlson, C. J., Albery, G. F., Merow, C., Trisos, C. H., Zipfel, C. M., Eskew, E. A., Olival, K. J., Ross, N., & Bansal, S. 2022: Climate change increases cross-species viral transmission risk. Nature 2022 607:7919, 607(7919), 555–562.

Hof, C., Voskamp, A., Biber, M. F., Böhning-Gaese, K., Engelhardt, E. K., Niamir, A., Willis, S. G., & Hickler, T. 2018: Bioenergy cropland expansion may offset positive effects of climate change mitigation for global vertebrate diversity. Proceedings of the National Academy of Sciences of the United States of America, 115(52), 13294–13299.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., … Zhang, X. 2020: Harmonization of global land use change and management for the period 850-2100 (LUH2) for CMIP6. Geosci. Model Dev, 13, 5425–5464.

Stefan Lange, Matthias Büchner (2021): ISIMIP3b bias-adjusted atmospheric climate input data (v1.1). ISIMIP Repository.

21.4
A multivariate approach to combine general circulation models using graph cuts

Lucas Schmutz[1], Soulivanh Thao[2], Mathieu Vrac[2], Grégoire Mariethoz[1]

[1] *Institute of Earth Surface Dynamics (IDYST), UNIL-Mouline, Geopolis, University of Lausanne, 1015 Lausanne, Switzerland (lucas.schmutz@unil.ch)*
[2] *Laboratoire des Sciences du Climat et l'Environnement (LSCE-IPSL) CNRS/CEA/UVSQ, UMR8212, Université Paris-Saclay, Gif-sur-Yvette, France*

General circulation models (GCMs) are of extreme importance to make future climate projections. Their predictions are used extensively by policymakers to manage their response to anthropogenic global warming and climate change.

To extract a robust global signal and evaluate uncertainties, individual models are often assembled in Multi-Model Ensembles (MMEs). Various approaches to combine individual models have been developed, such as the Multi-Model Mean (MMM) or its weighted variants.

Recently, Thao et al. (2022) developed a model comparison approach based on graph cuts. Graph cut optimization was developed in the field of computer vision to efficiently approximate a solution for low-level computer vision tasks such as image segmentation (Boykov et al., 2001). Applied to MMEs, it allows selecting for each pixel a set of best-performing models and produces a patchwork of models that maximizes performances while maintaining pixel-to-pixel continuity. It thus allows considering the local performance of ensemble members in contrast with approaches such as MMM or similar methods that use global weights.

Here we propose a new multivariate combination approach of MMEs based on graph cuts. Compared to the existing univariate graph cuts method, our approach ensures that the relationships between variables, that are present in ensemble members, are locally preserved while providing coherent spatial fields. Moreover, we exploit the design of the graph cut optimization to propose a stochastic version of model combinations that provides multiple similarly good solutions.

We demonstrate the efficiency of our approach by co-optimizing multi-decadal means of multiple variables. We compare the performance to univariate optimization and show that the loss of performance is small with a negligible increase in computational cost.

REFERENCES

Boykov, Y., Veksler, O., & Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*(11), 1222–1239. https://doi.org/10.1109/34.969114

Thao, S., Garvik, M., Mariethoz, G., & Vrac, M. (2022). Combining global climate models using graph cuts. *Climate Dynamics*, *February*. https://doi.org/10.1007/s00382-022-06213-4

21.5
Comparison of the performance of machine learning models for wildfire susceptibility mapping

Marj Tonini[1], Andrea Trucchia[2], Hamed Izadgoshasb[2], Paolo Fiorucci[2]

[1] *Institute of Earth Surface Dynamics, Faculty of Geosciences and Environment, University of Lausanne, CH-1015 Lausanne, Switzerland (marj.tonini@unil.ch)*
[2] *International Center for Environmental Monitoring CIMA Research Foundation, 17100 Savona, Italy*

Developing Wildfire Susceptibility Maps (WSM) and investigating the main driving factors for wildfire occurrence are fundamental to support forest protection and management plans. Namely, WSM indicate the likelihood for an area to be exposed to fire in the future based solely on the observed past events and on the local properties of a site (environmental and anthropogenic factors). Machine Learning (ML) based approaches are capable of learning from and make predictions on data by modelling the hidden/non-linear relationships between a set of input variables (i.e., the driving factors) and output observations; thus, they are particularly suited for WSM.

With the present study, Authors continue a research framework developed in a pioneristic work at local scale for Liguria Region, and lately adapted to national scale (Italy). In these earlier investigations, a Random Forest-based modelling workflow was developed to assess the susceptibility to wildfires under the influence of several environmental driving factors (land cover, type of vegetation, altitude and its derivatives, nearby infrastructures). The main novelties and contributions of the present study are: (i) we compared three ML models, namely Random Forest, Multi-layer Perceptron, and Support Vector Machine, to assess their prediction capabilities and to estimate which model performs better; (ii) we used a more accurate vegetation map as input, allowing to evaluate the impact of different types of local and neighbouring vegetation on wildfires occurrence; (iii) we improved the selection of the testing dataset, in order to take into account both the spatial and the temporal variability of the burning seasons.

As main results, WSMs were elaborated based on the output probabilistic predicted values from the three ML models, and the spatial distribution of the more susceptible areas was discussed. The three ML models were compered by means of the AUC (Area Under the Curve) ROC (Receiver Operating Characteristics), evaluated over the testing dataset.

In addition, the variable importance ranking was estimated as by-product of Random Forest, which can handle both numerical variables (as for the percentage of neighbouring vegetation) and native categorical variables (as for the types of vegetation at pixel level). Vegetation resulted by far to be the most important driving factor; the marginal effect of each single type of vegetation was also evaluated and discussed.

REFERENCES

Tonini, M.; D'Andrea, M.; Biondi, G.; Degli Esposti, S.; Trucchia, A.; Fiorucci, P. (2000): A Machine Learning-Based Approach for Wildfire Susceptibility Mapping. The Case Study of the Liguria Region in Italy. *Geosciences*, 10, no. 3: 105. https://doi.org/10.3390/geosciences10030105.

Trucchia, A.; Meschi, G.; Fiorucci, P.; Gollini, A.; Negro, D. (2022): Defining Wildfire Susceptibility Maps in Italy for Understanding Seasonal Wildfire Regimes at the National Level. *Fire*, 5,30. https://doi.org/10.3390/fire5010030.

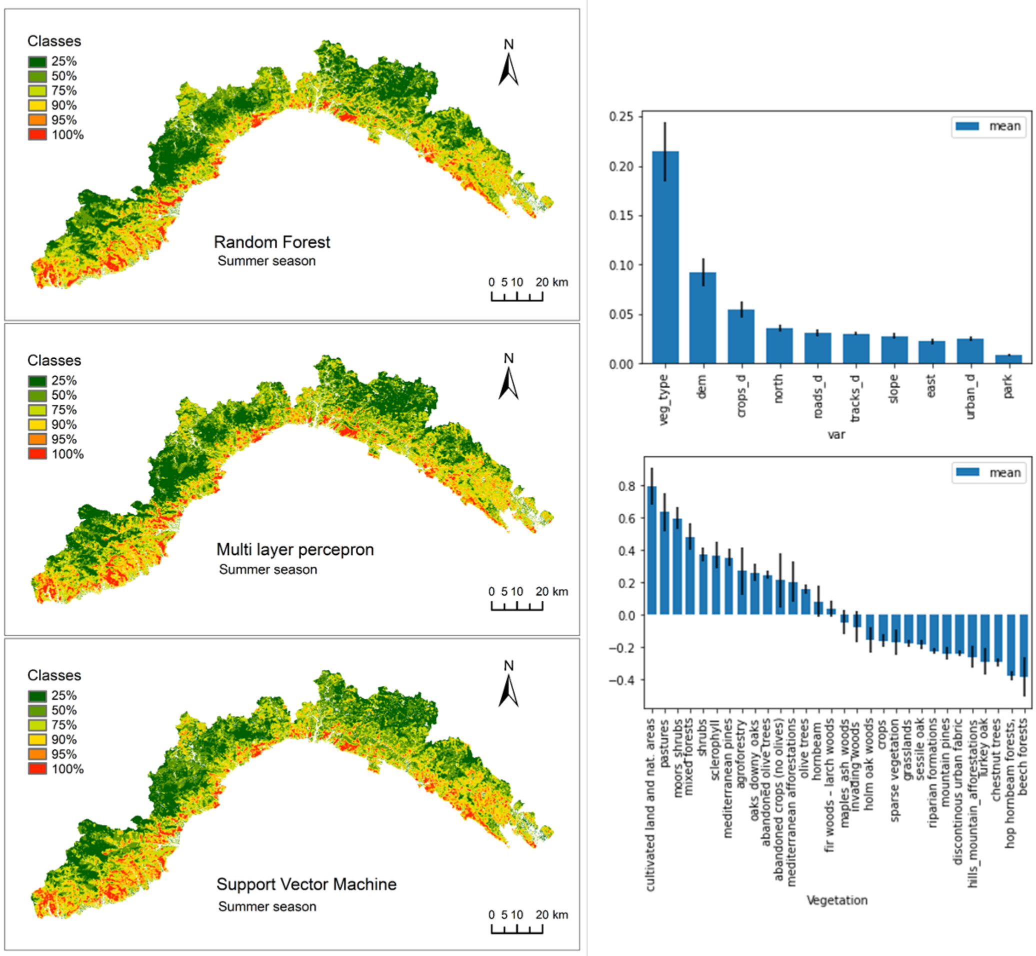

Figure 1. Wildfire Susceptibility Maps for Liguria region (Italy) (on the left). The variable importance ranking (top-right) and the partial plot of the types of vegetation (bottom-right).

21.6
Visualizing future forest changes in VR

Michelle S. Weber*, Verena C. Griess

Forest Resources Management, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zurich, Switzerland
(michelle.weber, verena.griess)@usys.ethz.ch

This presentation shows the potential of the iVR (immersive Virtual Reality) technology for the visualization of research results on future climate-related scenarios. The project was developed at ZHdK, as part of a graduate project (BA Design, Scientific Visualization), in close collaboration with the FORM group at ETH Zürich.

Figure 1. Rendering of a sparse forest built in Unity.

Current predictive models have a great potential to help practitioners and forest owners foresee changes inside the forest. Based on their experience and the outcome of the predicted scenario, they can decide which forest intervention strategy to apply. However, most of today's forest interventions will have impacts that they will not experience in their entire lifetime. Therefore, iVR is particularly relevant to help them visualize the impact of their management strategy. Indeed, by immersing the user inside the future forest, this could greatly improve their understanding of forest development.

The *VR-Experience* prototype "*Here Comes the Sun*" showcases a common forest intervention - the thinning of the forest – and its potential influence on the forest. Visitors embark on a journey through time. They first explore a dense forest from the present, and then visit the same forest in 100 years' time. Several stations highlight different topics related to the forestry interventions, e.g., the use of wood as a resource. Users can teleport to each station, where a narrator provides further insight into the station's topic. This allows users to interactively learn about the advantages and disadvantages of thinning the forest.
In conclusion, this session will focus on the presentation of the *VR-Experience* prototype with an emphasis on the concept, the implementation and the potential of the project.

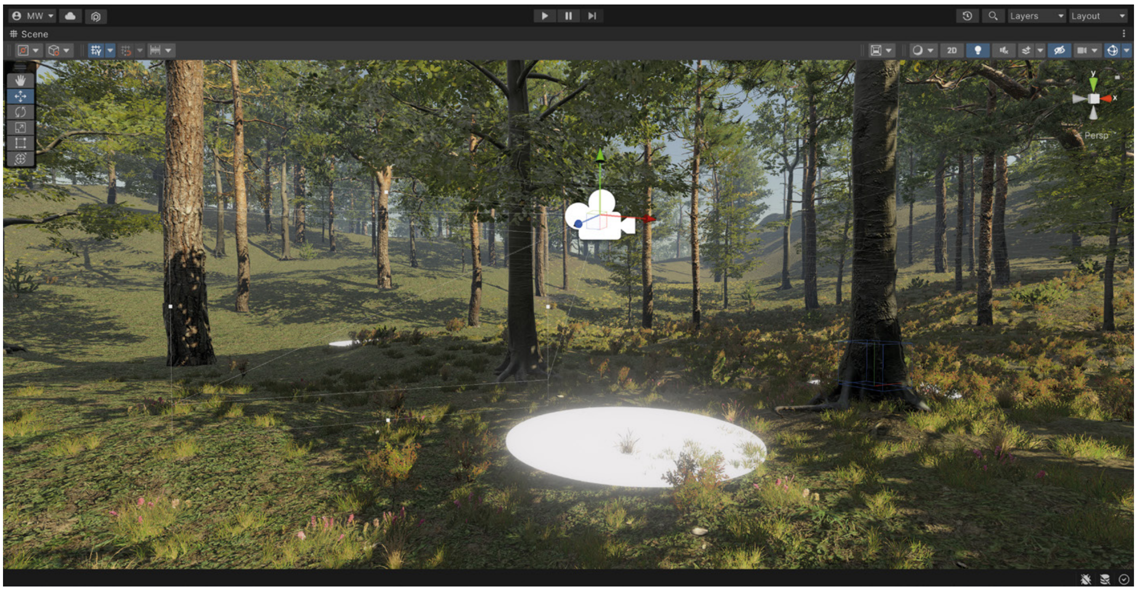

Figure 2. Setting up the forest environment with various information stations in Unity.

Figure 3. People testing the *VR-Experience* in the graduation exhibition.

21.7
Characterization of the Uncertainty in Land Cover Predictions near Train Tracks

Valérie Zermatten[1], Nicola Antonio Santacroce[1,2,3], Jules Gros-Daillon[1,2,3], Devis Tuia[1], Nicolas Ackermann[3]

[1] *Environmental Computational Science and Earth Observation Laboratory, EPFL, Ronquoz 35, 1950 Sion, (valerie.zermatten@epfl.ch)*
[2] *Environmental Sciences and Engineering, EPFL, Station 2, 1015 Lausanne*
[3] *Drone Development Department, SBB, Stöckackerstrasse 25, 3000 Berne*

The Swiss Federal Railways (SBB) manage a rail network approximately 3,000 kilometres long. The maintenance of the vegetation in the surroundings areas is an important part of this work. A classification of the land cover in proximity of the train tracks is used to monitor the current state of vegetation and determine the location where fieldwork must be undertaken. This land cover map focuses especially on vegetation types and height and requires a high level of accuracy. Until now, the land cover maps were derived through a time consuming and costly process involving extended manual corrections.

The objective of this work is to benefit from the recent advances in computer vision to provide a land cover map in automated manner and reduce the need for manual corrections. Our work involved the setup of a semantic segmentation model to predict the land cover classification and a second module providing the uncertainty of the predictions.

Input data include orthoimages with RGB and infrared bands, digital surface model, a vegetation high model derived from lidar data and a partially complete building layer.

We use UNet (Ronneberger et al., 2015), a convolutional neural network to classify the land cover at every pixel location. To account for both model and data uncertainty, we used two approaches: test time augmentation (Gawlikowski et al., 2021) was used to observe the data uncertainty. Model uncertainty is predicted through a Bayesian network with a method called Monte Carlo Dropout (Gal and Ghahramani, 2016) that slightly modifies the neural network by randomly turning off some of the model connections at each forward pass. The resulting variations in the predictions are interpreted as the model uncertainty.

Our results showed that our baseline segmentation model reaches a high level of accuracy (>95%) for categories such as building, bare soil or wooded areas, but performs less well on vegetation classes with rule-based separation such as trees (>2m height) and bushes (<2m height), or classes defined by external information such as ruderal areas (see Figure 1). Uncertainty maps highlight the uncertainty in the prediction close to shadowed areas, class boundaries or unusual appearance of a class. Combining the uncertainty maps (see Figure 2) with the predictions accuracy make it possible to select areas where human verification is needed: some potential wrong predictions (low confidence, low accuracy), or mislabelled areas (high confidence, low accuracy).

Figure 1: Left: an aerial image given as input; Center: the ground truth land cover labels; right: the classes predicted by our model.

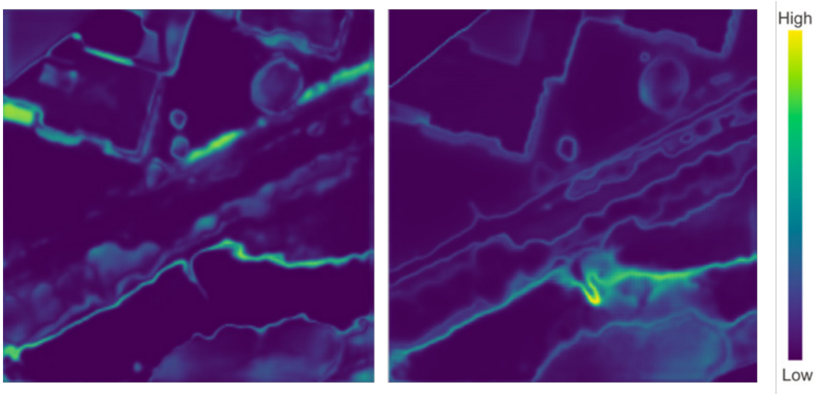

Figure 2: Left: Data uncertainty; Right: Model uncertainty for the example in Figure 1.

REFERENCES

Gal, Y., Ghahramani, Z., 2016. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning, in: Proceedings of The 33rd International Conference on Machine Learning. Presented at the International Conference on Machine Learning, PMLR, pp. 1050–1059.

Gawlikowski, J., Tassi, C.R.N., Ali, M., Lee, J., Humt, M., Feng, J., Kruspe, A., Triebel, R., Jung, P., Roscher, R., Shahzad, M., Yang, W., Bamler, R., Zhu, X.X., 2021. A Survey of Uncertainty in Deep Neural Networks. arXiv.org.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

P 21.1
Bayesian inversion using adaptive sequential Monte Carlo combined with surrogate modelling

Macarena Amaya[1], Niklas Linde[1], Giovanni Meles[1]

[1] *Institute of Earth Sciences, University of Lausanne, Switzerland (macarena.amaya@unil.ch)*

In the context of Bayesian inversion, traditional Markov chain Monte Carlo (MCMC) methods are commonly used to estimate posterior probability density functions (PDFs). Their efficiency can be limited when the inverse problems is highly non-linear and high-dimensional, thereby, failing to properly explore and sample the posterior PDF given realistic computation budgets. One way to circumvent such issues is to rely on tempering, which gives less weight to the likelihood function and, therefore, enhances the freedom of exploration. In particular, Sequential Monte Carlo methods (SMC) are particle approaches that build a sequence of importance sampling steps between neighbouring tempered distributions called power posteriors, thereby, bridging the prior and the posterior PDF. To ensure high quality estimation, SMC methods perform resampling steps when the variance of the particle weights become too high. To optimize the tempering schedule, we rely herein in adaptive sequential Monte Carlo (ASMC), a method which tunes the reduction in temperature between neighbouring power posteriors that work well for many problem-settings where advanced state-of-the-art MCMC methods struggle or fail. Still, its computational cost is high as many expensive forward simulations are required. Surrogate modelling are used to emulate the behaviour of the expensive forward solvers at much faster computational times. In this study, we incorporate a polynomial chaos expansion (PCE) surrogate model to ASMC inversion and evaluate it for a cross-hole tomography problem. We compare our preliminary results with previously obtained results obtained with standard MCMC (with or without surrogate modelling) and explore the value of updating the surrogate model as the ASMC tempered sequence progresses.

P 21.2
Modelling the Impact of Glacier Retreat on Vegetation Dynamic

Luca Eiholzer[1], Marj Tonini[1], Gianalberto Losapio[1]

[1] Institute of Earth Surface Dynamics, University of Lausanne, Quartier UNIL-Mouline, Bâtiment Géopolis, CH-1015 Lausanne (luca.eiholzer@unil.ch)

The retreat of glaciers is one of the emblematic effects of global warming, especially in the Alps, which leave the mountain landscape progressively uncovered by ice (Joerin et al., 2006) and colonized by plants. However, plants that inhabit the areas in front of retreat glaciers are exposed to climate warming and glacier retreat, with extinction risk increasing for early colonizers and pioneers (Losapio, 2021). This master project aims at developing a spatio-temporal model to predict the vegetation dynamic following glacier retreat. Species distribution maps for plants most sensitive to glacier retreat and glacier disappearance will be developed. We aim to estimate the risk of extinction and biodiversity loss under different scenarios and factors. This interdisciplinary work, involving biology, botany, and geo-informatics, is important in the field of geo-environmental sciences and represents a promising application in conservation biology and ecosystem management. Indeed, political authorities are increasingly using ecological data modelling tools to establish protected areas.

The study area is located at Glacier du Mont Miné (Val d'Herens, Valais, Switzerland), where its foreland extend in the subalpine zone over 2 km from 1'900 m to 2'000 m asl. A second area in Anniviers Valley was chosen to evaluate the spatial generalization capability of the model, which is, if the model trained in Ferpecle Valley is able to make good prediction in another area holding similar characteristics. A vegetation survey with a stratified random sampling approach, was carried out in July and August 2022. First, by means of a GPS, we elaborated a spatial inventory of key plant species and plant functional types (herb, graminoid, shrub, tree seedling, tree sapling, and tree adult) in 120 plots of 1 x 1 m^2. Tree diameter and height were measured too. We also developed environmental maps including different factors such as year of glacier retreat and geomorphology. The applied methodology is based on standard protocols for species distribution models (Araújo et al., 2019; Zurell et al., 2020) involing model fitting (selection and complexity), assessment and predictions. The final goal is to reconstruct the colonization and local extinction of plant species and vegetation during the last 170 years and to forecaste their distribution for the next 50 years. This work shall ultimately contribute to mitigate and anticiopate the impact of glacier retreat on biodiversity and ecological systems.

REFERENCES

Araújo, M. B., Anderson, R. P., Márcia Barbosa, A., Beale, C. M., Dormann, C. F., Early, R., Garcia, R. A., Guisan, A., Maiorano, L., Naimi, B., O'Hara, R. B., Zimmermann, N. E., & Rahbek, C. 2019: Standards for distribution models in biodiversity assessments. Science Advances 5(1), eaat4858. https://doi.org/10.1126/sciadv.aat4858

Joerin, U. E., Stocker, T. F., & Schlüchter, C. 2006: Multicentury glacier fluctuations in the Swiss Alps during the Holocene. The Holocene 16(5), 697–704. https://doi.org/10.1191/0959683606hl964rp

Losapio, G., Cerabolini, B. E. L., Maffioletti, C., Tampucci, D., Gobbi, M., & Caccianiga, M. 2021: The Consequences of Glacier Retreat Are Uneven Between Plant Species. Frontiers in Ecology and Evolution 8. https://www.frontiersin.org/article/10.3389/fevo.2020.616562

Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Guisan, A., Lahoz-Monfort, J. J., Leitão, P. J., Park, D. S., Peterson, A. T., Rapacciuolo, G., Schmatz, D. R., Schröder, B., Serra-Diaz, J. M., Thuiller, W., … Merow, C. (2020). A standard protocol for reporting species distribution models. Ecography 43(9), 1261–1277. https://doi.org/10.1111/ecog.04960

P 21.3
Sequential Monte Carlo for posterior risk assessment

Lea Friedli[1], Niklas Linde[1], David Ginsbourger[2] and Arnaud Doucet[3]

[1] Institute of Earth Sciences, University of Lausanne, Switzerland;
[2] Institute of Mathematical Statistics and Actuarial Science and Oeschger Center for Climate Change Research, University of Bern, Switzerland;
[3] Department of Statistics, Oxford University, United Kingdom.

We consider non-linear Bayesian inversion problems in which we infer an unknown hydrogeological or geophysical property field by using hydrogeological or geophysical data. In practice, one is often not interested in the field itself, but in a quantity depending on the field through non-linear relationships, for instance, in the probability of this quantity exceeding a critical threshold. We estimate a hydraulic conductivity field using pumping tests and aim at predicting the probability of a contamination level exceeding a critical threshold at given time since the release of a pollutant. The classical approach to infer Bayesian posteriors is to apply a sampling method such as the Metropolis-Hastings (MH) algorithm or the Sequential Monte Carlo (SMC) method. A straightforward extension to risk estimation is then to perform forward modeling on the posterior samples in order to estimate the probability of exceeding the threshold of interest. However, since the rare events of interest are likely associated with the tails of the posterior, it is challenging to estimate their occurrence with such a brute force approach. Instead, we propose a sequential application of SMC algorithms. The first stage relies on a particle estimation of the posterior using a sequence of probability density functions giving increasing weights to the likelihood. To subsequently estimate the risk of the rare event, a second SMC is applied. Using an increasing sequence of thresholds, in each iteration the particles exceeding the threshold are selected and propagated. We compare the results of this Doubly Sequential Monte Carlo (DSMC) method with the brute force approach. Preliminary results show that the DSMC method requires significantly less samples to guarantee an accurate estimate of the risk.

P 21.4
Generative adversarial network for paleo glacial landscape reconstruction

Guillaume Jouvet[1,2]

[1] *Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland (guillaume.jouvet@unil.ch)*
[2] *Department of Geography, University of Zurich, Zurich, Switzerland*

What would a satellite have seen if it had passed over the Alps during the last glacial maximum? To try to answer this question, we present a deep-learning generative model that predicts plausible glacial landscape texture from climate and glacier related variables. For that purpose, we use climate, topography and Sentinel 2 satellite imagery data of present-day glaciers worldwide to ``instruct'' a Generative Adversarial Network (GAN) how the texture of glaciers and their surroundings is explained from predictive variables. The GAN permits to re-interpret modeled glacier reconstructions into manufactured satellite-like images, which greatly improve the perception and visualization of the landscape. To illustrate the potential of the method, we reconstruct the texture of the Alpine landscape during the last glacial maximum, when most of it was cover by ice, from coupled climate-glacier model results presented in a companion SGM contribution (Jouvet and al., 2022).

REFERENCES

Jouvet G., D. Cohen, E. Russo, J. Buzan, Christoph C. Raible, U. H. Fischer, W. Haeberli, M. Imhof, J. K. Becker, A. Landgraf. Coupled climate-glacier modelling of the last glaciation in the Alps. 2022 SGM contribution in Symposia “10. Quaternary Environment”.

P 21.5
Combining inverse autoregressive flows with deep generative networks for improved efficiency and scalability of inverse problems

Shiran Levy[1], Eric Laloy[2], Niklas Linde[1]

[1] *Institute of Earth Sciences, University of Lausanne, Lausanne, Switzerland (shiran.levy@unil.ch)*
[2] *Institute for Environment, Health and Safety, Belgian Nuclear Research Centre, Mol, Belgium*

We study the combination of two deep learning methods: normalizing flows and deep generative models, within the context of geophysical inverse problems as an efficient alternative to Markov chain Monte Carlo (MCMC) sampling. Normalizing flows is a series of transport maps transforming an initial density of random variables into a target density. Specifically, we use inverse autoregressive flows (IAF), in which transformation of one instance is conditioned on previous instances. In this context, the flows are blocks of autoregressive neural networks for which the outputs are scale and shift functions applied on the input variables. The target density resulting from the transformation is parameterized by the neural parameters, allowing for more expressive and complex distributions. The IAF transformation is trained by optimizing the neural parameters using variational inference (VI). The objective in VI is to approximate some target distribution parametrically for a given family of distributions. It provides a computationally-efficient approach that scales well to high-dimensional problems. In each training iteration of the IAF, samples from a normal distribution are pushed forward through the invertible, differentiable transformation onto a variational density approximating the posterior of interest. The parameters of the IAF are learned by maximizing the evidence lower bound which is essentially equivalent to minimizing the Kullback-Leibler divergence between the variational density and the target posterior distribution. In our study, we use two different deep generative models: a generative adversarial network (GAN) and a variational autoencoder (VAE), to encode the high-dimensional prior model into a low-dimensional, latent space of normally distributed variables. We compare the presented approach against popular methods for solving geophysical inverse problems such as deterministic gradient-based methods and MCMC sampling. Due to the nonlinearity of GANs, previous attempts incorporating GANs and deterministic gradient-based inversion failed. Nonetheless, when tested on channelized subsurface models given nonlinear physics the posterior approximation resulting from the IAF is in agreement with the true model for both types of generative models and provides a reliable uncertainty quantification when the VAE is used. Moreover, the training of the IAF is seven times faster than an equivalent MCMC inversion.

P 21.6
Advanced spatial learning technique for automatic mapping of geomorphological features in alpine periglacial environment

Alessio Poloni[1], Dr. Marj Tonini[1], Dr. Christophe Lambiel[1]

[1] *Institut des Dynamiques de la Surface Terrestre (IDyST), University of Lausanne, Batiment Geopolis, CH-1015 Lausanne (alessio.poloni@unil.ch)*

Permafrost corresponds to unconsolidated sediments or bedrock that remains frozen for at least two consecutive years (Rogger et al. 2017). This phenomenon is difficult to observe because it occurs in the sub-surface and is therefore invisible to the naked eye. Permafrost is an important component of high mountain regions, and its investigation is of current interest, especially with regard to its evolution in relation to ongoing climate changes. The main controlling factors are the mean annual air temperature, the solar radiation, the characteristics of the ground surface and of the active layer, and the snow duration and depth.

In mountainous regions such as the Alps, characterised by highly variable micro-climatic, topographical and terrain roughness conditions, the spatial distribution of permafrost is very heterogeneous. This complex pattern makes it difficult to map at the local scale. Several permafrost distribution models, based on physical, empirical or statistical approaches, have been developed, but these models tend to overestimate the actual extent of permafrost. In recent years, interest about prediction of the presence or absence of permafrost determined through the use of machine learning algorithms has grown.

Machine learning methods are increasingly applied in geo-environmental applications to predict the occurrence of a phenomenon based on the observations of its presence in a given area and on a set of predictive factors. The Machine learning domain includes algorithms capable of learning from data, by modelling the hidden relationships linking a set of input and output variables. For permafrost distribution, models based on Random Forest (RF) have proven to be efficient and accurate (Deluigi et al. 2017), but the algorithm is generally applied in a “non-spatially” way, and it cannot consider the spatial relationships between variables.

In this research, an advanced classification algorithm is developed with the aim of determining the occurrence of permafrost at the local scale with high accuracy. This algorithm is based on two methods derived from RF, which additionally allow a spatial analysis of the phenomenon: Geographical RF (Georganos et al. 2019) and Spatial RF (Benito 2021). The novelty of this approach consists also in the addition of new parameters related to the terrain roughness (Figure 1). The input dataset includes twelve control variables (air temperature, solar radiation, plan curvature, profile curvature, grain size, altitude, northness, eastness, slope degree, ndvi, poly and lithology). The final output will be a high-precision permafrost susceptibility map of the selected study area, which will also consider the spatial contiguity among the pixels.

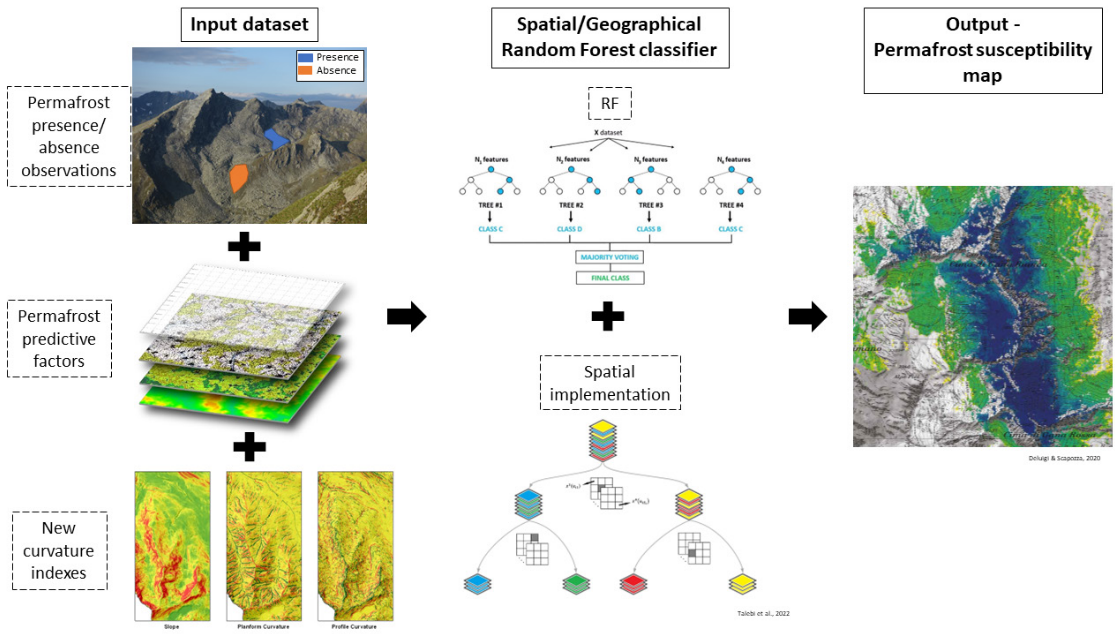

Figure 1. Methodology workflow diagram.

REFERENCES

Blas M. Benito. 2021. R package spatialRF: Easy Spatial Regression with Random Forest. doi: 10.5281/zenodo.4745208. https://blasbenito.github.io/spatialRF/

Deluigi, N., Lambiel, C., & Kanevski, M. 2017. Data-driven mapping of the potential mountain permafrost distribution. Science of The Total Environment, 590-591, 370-380. https://doi.org/10.1016/j.scitotenv.2017.02.041

Georganos, S., Grippa, T., Niang Gadiaga, A., Linard, C., Lennert, M., Vanhuysse, S., . . . Kalogirou, S. 2021. Geographical random forests: A spatial extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population modelling. Geocarto International, 36(2), 121-136. https://doi.org/10.1080/10106049.2019.1595177

Rogger, M., Chirico, G. B., Hausmann, H., Krainer, K., Brueckl, E., Stadler, P., & Bloeschl, G. 2017. Impact of mountain permafrost on flow path and runoff response in a high alpine catchment. Water Resources Research, 53(2), 1288-1308.

P 21.7
Mountain Landslide susceptibility assessment by Multi-sampling strategy combining optimization of BiLSTM algorithm

Jingjing Wang[1 3], Gang Chen[2], Michel Jaboyedoff [3], Marc-Henri Derron[3]

[1.] Key Laboratory of Geological Survey and Evaluation of Ministry of Education, China University of Geosciences, Wuhan 430074, China (wangjingjing_cug@163.com)
[2.] College of Marine Science and Technology, China University of Geosciences, Wuhan 430074, China
[3.] Institute of Earth Sciences, University of Lausanne, Lausanne 1015, Switzerland

Machine learning models have been widely used in landslide susceptibility assessment. However, landslide feature extraction, prediction quantification, and representation are still challenges. This paper proposes a multi-sampling strategy combining optimization of the BiLSTM（Bidirectional Long Short-Term Memory） algorithm to evaluate landslide susceptibility in mountainous areas. The proposed method mainly involves three steps:（1）The spatial correlations were analyzed between historical landslides and geographic environment factors, such as topography (DEM, slope, aspect, curvature, surface relief, faults, land subsidence, etc.), hydrogeology (terrain moisture index, rivers, etc.), vegetation cover, earthquake, and human activity (roads, land use) in Sichuan province. Then landslide Inducing factors were selected and expressed quantitatively. The landslide density within the statistical grid cell is used as a predictor to define the landslide probability. The landslide density statistical index refers to the ratio of the number of historical landslide points counted cell grid area. The index was used as the classification basis of susceptibility, which can better reflect the spatial clustering characteristics of landslides than traditional machine learning using a single landslide point as a classification. （2）Construction of multi-channel landslide feature engineering. With multi-band fusion technology, multi-band raster image files in datasets were read, processed, and visualized. The sample information was stored as an image with multiband information and then fed into the algorithm. The method can achieve pixel-by-pixel feature extraction of landslides from high-dimensional multi-source geographic data. Compared with the traditional pixel points represented by a set of matrices as input, it avoids the loss of landslide information because the landslide is a surface and not just a point. (3) The dataset is derived from the Qinba Mountains and the eastern Hengduan Mountains in Sichuan Province. The ratio of the number of samples in the training set, validation set, and test set is 6:2:2. The total number of examples in the data set varies, depending on different sampling strategies. Based on transfer learning, we use a realistic landslide distribution scenario simulation and Bayesian optimization variable strategy to analyze the influence of the proportion of positive and negative samples in the BiLSTM algorithm training process on the prediction results.
The consequence shows that the deep neural network model jointly optimized by multiple sampling strategies can effectively extract the spatial distribution of landslides in complex environments and accurately predict the spatial location of potential landslides.

REFERENCES
Pham B T, Pradhan B, Bui D T, et al. J.2016; A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India). Environmental Modelling & Software, 84: 240-250.
Orhan O, Bilgilioglu S S, Kaya Z, et al. J.2022; Assessing and mapping landslide susceptibility using different machine learning methods. Geocarto International, 37(10): 2795-2820.