

Reviews of Manuscript: egosphere-2025-670

Title: Harnessing Machine Learning methods for climate multi-hazard and multi-risk assessment

Special issue: Methodological innovations for the analysis and management of compound risk and multi-risk, including climate-related and geophysical hazards (NHES/ESD/ESSD/GC/HES inter-journal SI)

The manuscript was revised according to the Reviewers' suggestions and recommendations, implementing specific changes related with the research questions, the methodology, the structure of the paper, and selected topics as requested. Responses to each general and detailed comment are provided in this letter for your convenience. Furthermore, we enclose a new version of the Manuscript with changes performed to improve the manuscript according to the reviewers' suggestions in track change, noting that the reference lines in our responses refer to the line in the manuscript WITH track change (see the file Manuscript_egosphere_670_reviewed_with_track_changes.doc).

Reviewer 1

Answers to general comments

Comment 1: Lack of Scientific Novelty or Conceptual Contribution: The paper lacks scientific novelty and conceptual contribution. It does not introduce new concepts, frameworks, or theoretical insights. Instead, it compiles existing literature without offering a critical synthesis or identifying research gaps. The review does not significantly advance the understanding of multi-hazard or multi-risk modelling compared to previous reviews, and its largely descriptive discussion limits its value as a synthesis resource.

Response. We thank the reviewer for this comment that gives us the opportunity to better explain the relevance of our study. We have better positioned the paper in relation to similar studies, given that this paper focuses on the specific role of ML and statistical approaches in advancing multi-hazard and multi-risk assessment, rather than single-hazard or sector-specific applications. [lines 156-174; 202-225]

Revised text:

[154-195] Unlike other recent reviews that have focused on ML (particularly DL) for specific hazards or sectors – such as extreme events (Salcedo-Sanz et al., 2022), hydrology (Tripathy & Mishra, 2024), geophysics (S. Yu & Ma, 2021), wildfires (Jain et al., 2020), and climate risk (Zennaro et al., 2021) – this paper takes a cross-cutting perspective on multi-hazard and multi-risk assessment. By structuring the discussion around successive stages of risk analysis – data processing, hazard prediction, risk assessment, and future scenarios – we connect climate risk and data-driven methods while also identifying critical gaps, particularly in linking hazard interactions with vulnerability.

[215-300]

1. Data: How can Machine Learning improve data collection and processing?
2. Multi-Hazard: How can Machine Learning and statistical tools be used to analyse extreme events, and model hazard interactions?
3. Multi-Risk: How can Machine Learning applications integrate vulnerability and exposure in multi-risk analysis?
4. Future: How can Machine Learning and statistical tools be used to predict long-term future multi-hazard and multi-risk? [...]

The first research question examines how ML can help process diverse types of data, extracting and harmonising the information needed to analyse multi-hazard and multi-risk by addressing current gaps such as data sparsity, inconsistency across sources, and the lack of harmonised formats. This contributes to improving the quality and comparability of risk assessments by enabling integrated use of climate, EO, and textual datasets. In particular, the sub-themes are divided based on the type of data analysed:

- I. Climate data (time series of geospatial climate data), which describe the characteristics of climate-related hazards across space and time. Preparing this data for multi-hazard and multi-risk applications often requires ML methods (i.e. feature engineering) to increase spatial and temporal resolution, harmonise and extend the time coverage of the datasets or correct for biases (Schneider et al., 2023).
- II. EO, which can be used to characterise exposure and vulnerability layers and extract information on impacts (Ghaffarian & Emtehani, 2021; Novellino et al., 2024).
- III. Textual data, such as newspapers or social media, which in the last years have been leveraged for extracting information on diverse impacts (Sodoge et al., 2023).

The second research question investigates how ML and statistical tools improve the identification and modelling of hazard dynamics by capturing complex spatio-temporal patterns, compounding effects, and non-linear interactions that traditional approaches often overlook. This helps advance multi-hazard and multi-risk analysis by providing more accurate detection, classification, and modelling of extreme events. In particular, the key sub-themes are:

- I. Analyse which methods can be used to identify, classify and cluster extreme events, producing spatio-temporal footprints of multi-hazard events (H. Yu et al., 2022).
- II. The prediction of (multi-)hazard events, for example through early warning systems or seasonal predictions (Bhowmik et al., 2023).
- III. The analysis of hazard interactions, for example characterising joint distributions through copulas (Bevacqua et al., 2021) or multi-hazard susceptibility maps (Pourghasemi et al., 2019).

The third research question concerns the application of ML for the integration of vulnerability and exposure into multi-risk analysis addressing the current gap where vulnerability and exposure are often treated as static or secondary layers rather than dynamic drivers of risk. This integration strengthens the ability of multi-risk assessments to capture how socio-economic conditions and adaptation measures interact with hazards to shape overall risk. In particular, the key themes are:

- I. Multi-hazard exposure and vulnerability on assessments, integrating susceptibility mapping with information on specific exposure layers, such as buildings and population (Rusk et al., 2022).
- II. Modelling risk from past impacts data, often through supervised ML approaches that use hazard, vulnerability and exposure indicators as predictors (Dal Barco et al., 2024).

The fourth research question investigates the possible contribution of ML and statistical tools into the analysis of (long-term) future multi-hazard and multi-risk, where uncertainty associated with the representation of future extremes in climate projections further complicates risk modelling, highlighting a critical gap in existing approaches, which often fail to adequately capture compound and cascading extremes under changing climate conditions. This research question clarifies how ML can enhance scenario building, improve uncertainty quantification, and support more robust long-term multi-risk assessments. In particular, the key sub-themes are:

- I. Modelling future multi-hazard trends and spatial patterns using statistical methods, in particular for compound and consecutive events (Zscheischler et al., 2018).
- II. Assessing future impacts based on climate change projections, often using methods trained on historical data and applied to ensembles of RCP projections (S. J. Park & Lee, 2020).

Moreover, we have refined each of the results sections by adding a summary table outlining the main methods, gaps, and opportunities, together with a paragraph explicitly explaining how the section contributes to the field of multi-hazard and multi-risk analysis.

Below we provide the tables that were produce for each section.

Section 1	Methods	Gaps	Opportunities
3.1.1 Climate datasets	LSTM for soil moisture & hydrology; RF for runoff/precipitation; ConvNPs for sparse sensors; ANNs for interpolation; CNNs, GANs, diffusion models for downscaling; GPs for uncertainty quantification	Struggles with sparse/irregular data; poor scalability (GPs); extremes misrepresented; limited uncertainty treatment	Hybrid ML–physics models; scalable probabilistic methods; better uncertainty quantification; generative models for ensembles
3.1.2 Earth observations (EO)	SVM, RF, LSTM for soil moisture; CNNs/autoencoders for land cover, impacts, disaster recovery; transfer learning for data-scarce regions; ML for water quality (RF, ANN, XGBoost)	Bias toward data-rich regions; revisit gaps/clouds limit detection; false positives; weak multi-hazard integration	Robust models for missing/noisy data; near-real-time EO pipelines; integrate EO with socio-economic data; transfer learning for vulnerable regions
3.1.3 Textual data	NLP + ML (Naïve Bayes, RF, SVM, CNN, BERT, LSTM); multimodal (text + images); rule-based for small datasets	Few labelled datasets; language/cultural bias; imprecise spatial info; noisy social media inputs	Multilingual/transfer learning; improved geolocation extraction; integrate with EO/sensor data; robust methods for noisy/misinformation-prone data

Section 2	Methods	Gaps	Opportunities
3.2.1 Identify, classify & cluster	Thresholding (empirical & percentiles) to build multi-hazard catalogs; return periods & GEV; EOFs; CNNs (semi-/supervised) for extreme-weather object detection in reanalyses; DBSCAN for spatio-temporal footprints and compound clusters.	Under-detection of joint (non-univariate) extremes when hazards are merged post-hoc; label scarcity & class imbalance for supervised DL; skewed datasets; sensitivity to spatial/temporal non-stationarity.	Unified pipelines that detect compound signatures directly (multivariate thresholds + clustering); semi-/self-supervised DL to mitigate label scarcity; robust cluster tracking of compound hotspots under change.
3.2.2 Hazard forecasting & prediction	LSTM/CNN for hydrology, storm surge, drought/heat; Transformers for floods; GNN/GRU for river-network dynamics; classical ML (RF/SVM/XGB) for local extremes when data are limited.	High data demands; generalisation beyond observed regimes; limited interpretability; performance varies with spatial context and input windowing.	Physics-informed/graph-aware DL for better extrapolation; attention/attribution to expose drivers; global-to-local transfer learning; benchmarking vs. process models for trust.
3.2.2 Modelling	Copulas (pair/vine/Joe) for joint extremes; copula-BNs for river–coastal compounding; XAI on	Copula family selection & tail-dependence in high dimensions; ML black-box limits causal	Hybrid ML–copula stacks (ML to predict/characterise events, copulas to quantify joint probabilities);

hazard interactions	LSTMs/CNNs/Transformers (gradients, attention, sensitivity) to reveal drivers and shifts.	insight; difficulty linking physical drivers to dependence structures.	benchmarking ML-learned dependencies against copula baselines; conditional vines for cascades.
Susceptibility mapping (multi-hazard)	Supervised ML (LR/GLM, RF, SVM, BRT, CART, ANN, CNN) to build single-hazard susceptibility then combine into multi-hazard maps; feature importance to rank drivers.	Often “multi-layer single-hazard” (weak interaction modelling); training skew (few positives); sampling bias & autocorrelation; static inputs ignore evolving exposure/vulnerability.	Spatio-temporal CV (block) to curb leakage; dynamic susceptibility that updates with sequences/adaptation; explicit interaction terms or graph-based fusion; extend beyond the usual quartet (fire/landslide/flood/quake).

Section 3	Methods	Gaps	Opportunities
3.3.1 Risk via susceptibility + exposure + vulnerability	Overlay of single-hazard susceptibility (RF, SVM, ANN, BRT, CART, MaxEnt, CNN with patch context) with exposure (buildings, population, infrastructure) and simple vulnerability layers; AHP/MCDM weighting; feature importance/SHAP to rank drivers.	Often “multi-layer single-hazard” (weak/implicit interaction modelling); vulnerability treated as static; label imbalance and sampling bias; spatial/temporal autocorrelation leakage; limited hazard breadth beyond fire/landslide/flood/quake.	Dynamic vulnerability/exposure updates using EO and time-sequenced hazards; spatio-temporal block cross-validation; interaction-aware fusion (graphs, learned weights); extend to wind, hail, heat, surge; probabilistic risk surfaces with uncertainty bands.
3.3.2 Predicting impacts	Ensemble and hybrid ML approaches (RF, XGBoost, SVM, DL, copulas, causal ML) applied to health, food, environmental, and economic impacts; explainable AI (SHAP) and probabilistic modelling for driver attribution.	Impact labels are sparse, coarse, biased, and confounded; scale mismatches and aggregation blur signals; extremes and tails poorly represented; DL tends to overfit and transfer poorly across cities/regions/climates; uncertainty quantification and causal attribution often limited.	(1) Data & catalogs: build geocoded, event-level, cross-sector impact datasets and standardized labels (health, yields, biodiversity, losses). (2) Causal & lag-aware stacks: combine DLNM / explicit-lag models with ML and causal discovery to capture delayed and causal pathways. (3) Multi-source fusion & transfer: integrate EO, in-situ, socio-economic and market data; use domain-adaptation/transfer learning for cross-region generalization.

Section 4	Methods	Gaps	Opportunities
3.4.1 Predicting future hazards	Bias correction for projections (incl. quantile mapping); hotspot/trend detection via percentile thresholds (e.g., 95th–99th), return periods; uncertainty sources and propagation; vine copulas for joint tails; SMILE large ensembles; storyline event-based scenarios for HILP analysis.	Univariate bias correction can distort extremes/joint tails; regional skill varies; limited direct detection of compound signals; uncertainty treatment often partial.	Adopt multivariate bias correction; combine SMILEs to separate internal variability vs. structural model differences; scale up vine copulas for compound events; embed storylines for preparedness.
3.4.2 Modelling future impacts	Trained on historical impacts and applied to future ensembles; ensemble ML methods (RF, XGBoost, ...) for coastal risk, conflict risks, crop yield and adaptation scenarios; Bayesian Networks for multi-model chains (hydrodynamics–waves–shoreline); distributed-lag models for future health impacts; future susceptibility integrating land use changes	Impact data often coarse, biased, and sparse; studies often rely on few years → low representativeness; causal discovery hinges on strong assumptions; biases due to scale mismatch in climate–exposure–impact data.	Use multi-model, multi-year ensembles and probabilistic models to quantify and propagate uncertainty.; evaluate adaptation portfolios with standardized costs; integrate dynamic land use + socio-economic pathways (SSPs); uncertainty-aware learning, and open, geocoded impact datasets.

Comment 2: No PRISMA flow diagram is provided, and Search strings, filtering criteria, and quality assessment processes are not disclosed. Suggest including a PRISMA diagram and methodological appendix (even in supplementary materials) to ensure transparency.

Response. We appreciate the reviewer’s suggestion. In response, we have now updated the PRISMA flow diagram (Figure 1) to better illustrate the literature selection process.

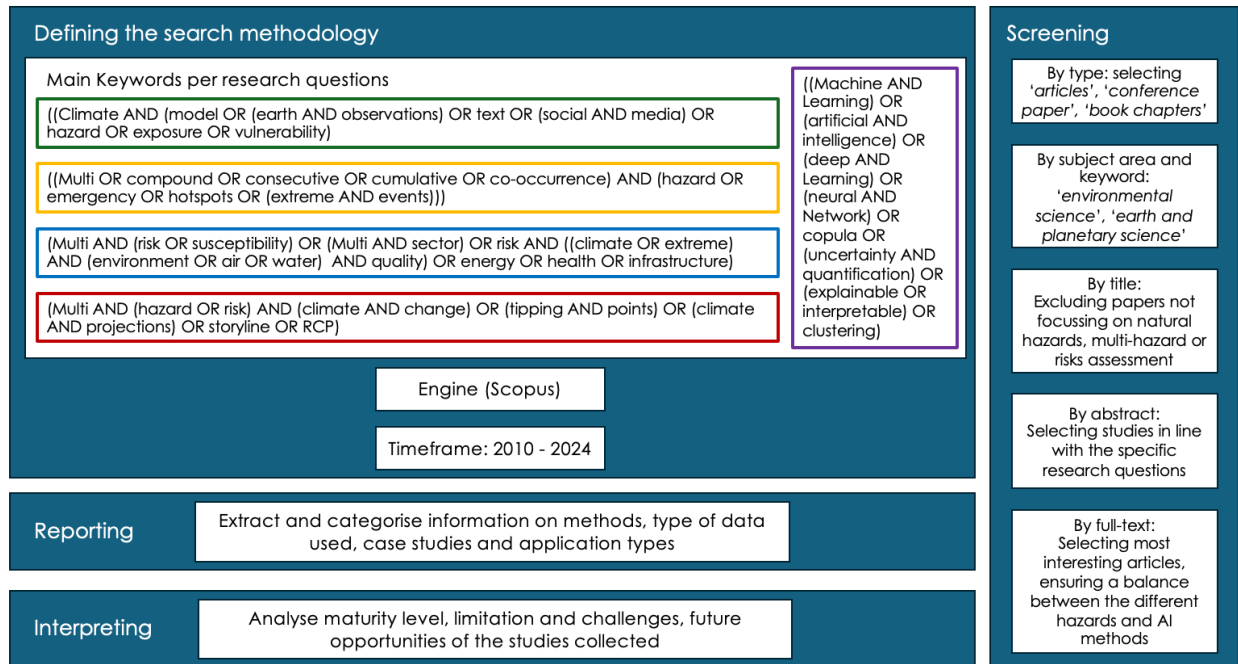


Figure 1: Literature review methodological steps

In addition, we have included a methodological appendix (see lines 1150 -1175) detailing the search strings, filtering criteria, and quality assessment procedures used in the review.

Revised text: The literature review followed the PRISMA guidelines to ensure transparency and reproducibility in the identification, screening, and selection of studies. The process is summarized in the PRISMA flow diagram and detailed as follows. First, records were retrieved from major scientific databases (Scopus) and filtered by type, retaining only *articles*, *conference papers*, and *book chapters* and language (English). Next, documents were filtered by subject area and keyword, selecting only those classified under *Environmental Science* and *Earth and Planetary Science* as subjects and considering machine learning, climate change, risk assessments (and their synonyms and variations) as keywords.

In the third step (title screening stage), studies not focusing on *natural hazards*, *multi-hazard*, or *risk assessment* were excluded. During the abstract screening stage, each paper was evaluated for its relevance to the review’s research questions, focusing particularly on the use of machine learning (ML) techniques and their application to multi-hazard or multi-risk contexts. Studies were retained if they explicitly applied ML, AI, or statistical learning methods to the modelling, characterization, or assessment of natural hazards, or if they addressed interactions between multiple hazard types (e.g., cascading or compound events) and their associated risks. Papers focusing solely on single hazards without methodological innovation or on unrelated environmental modelling were excluded. This step ensured that the final selection captured studies advancing methodological understanding of ML-driven hazard analysis, as well as those integrating multiple hazard processes or risk dimensions. Finally, the full-text review identified the most relevant and representative papers, ensuring balanced coverage across different hazard types and AI methodologies. The final selections were based on diversity in data sources, geographical coverage, hazard types and machine learning methods used. This process ensured that the resulting corpus reflects the breadth of current research at the intersection of AI, Earth observation, and multi-hazard risk assessment.

The number of studies retained at each step is summarized below (numbers correspond to the four main research questions):

Table B1: summary of the screening step results

Screening step	RQ1	RQ2	RQ3	RQ4
Initial retrieval	24,335	9,542	22,054	1,961
After type filtering	17,676	8,731	3,548	344
After title screening	6,999	801	2,215	104
After abstract screening	376	107	214	67
Full text screening	52	50	29	22

These additions improve the transparency and reproducibility of our literature selection and review methodology.

Comment 3: Poor Language Quality: The manuscript contains many grammatical errors, awkward phrasing, and redundant or overly long sentences. The language undermines clarity and makes the manuscript difficult to read. Given its role as a review article, this significantly reduces its accessibility and utility to the scientific community. A full professional language revision is essential.

Response. We thank the reviewer for this important comment. We have carefully revised the entire manuscript to improve its readability and clarity. In particular, we have:

- Checked and corrected grammar and syntax throughout the text.
- Simplified and shortened overly long or complex sentences.
- Removed redundancies and improved the overall flow and coherence.
- Ensured consistent and precise terminology.

These edits have substantially improved the language quality and accessibility of the manuscript. We believe the revised version now meets the standards expected for a review article.

Comment 4: Weak Structure and Inconsistent Framing: The manuscript shifts terminology from "machine learning methods" from the title to "data-driven methods" in the text, without clear justification. The manuscript includes a statistical method (copula), which is not a machine learning method. Poor coherence across sections: Overlap and redundancy between sections, and not clearly defined (e.g., 3.1.2 vs. 3.1.1 regarding satellite images observed soil moisture belongs to EO or climate data).

Response. We thank the reviewer for this valuable feedback. In response, we have revised the manuscript to ensure consistent framing and clearer structure. Specifically, we have:

- Adopted "data-driven methods" as the unifying term throughout the paper, including the title, to explicitly encompass both machine learning and statistical approaches such as copulas, which are relevant to the review's scope.
- Clarified the distinction between copulas and machine learning in the Introduction, explaining the rationale for including both classes of methods under the broader "data-driven" framework. [Lines 134-143]

Revised text: In addition to ML methods, this review briefly considers the role of copulas as multivariate statistical tools in multi-risk assessment. Copulas enable explicit modelling of the

dependence structure between variables, making them particularly valuable for analysing compound events in which multiple hazards occur simultaneously or sequentially (see, for example, Agrawal, 2022; Hochrainer-Stigler et al., 2019). They have, for instance, been used to characterise the joint occurrence of droughts and heatwaves, yielding insights into their combined impacts on agriculture and water resources (see e.g. Ribeiro et al., 2020). Although their application is more specialised than most ML approaches, copulas provide critical information about inter-hazard dependencies, supporting a deeper understanding of cascading and interacting risks. Their inclusion in this review therefore highlights their importance in contexts requiring precise statistical modelling of hazard interactions and underscores how they complement broader ML-based strategies in climate-risk analysis.

- We have reviewed the section organization to reduce redundancy and improve coherence, particularly between Sections 3.1.1 and 3.1.2. In the revised version, studies focusing on the production and processing of climate and hydrological variables (e.g., soil moisture, precipitation) are grouped in the first section, while those applying AI and data-driven methods to assess exposure, vulnerability, and impacts (e.g, land cover changes) are discussed in the second. This reorganization aims at clarifying the distinction between climate data and impact-oriented applications and improving consistency and readability.
- Reviewed the section organization to reduce redundancy and improve coherence, particularly between Sections 3.1.1 and 3.1.2, ensuring that data categories (e.g., Earth Observation vs. climate data) are consistently defined.
- Streamlined overlapping content and refined transitions to improve the logical flow of the manuscript.

We believe these revisions have substantially strengthened the manuscript’s clarity, consistency, and conceptual framing.

Other comments:

We thank the reviewer for the careful reading and detailed feedback. We have revised the manuscript thoroughly to address all noted issues:

1. We rephrased the section: “can advance multi-hazard and multi-risk” for greater precision and formality: “Supervised machine learning approaches trained on historical impact data to project future climate risks have also emerged as a significant research area...” (now lines 27-32)
2. Corrected the typo “his information” → “This information” (now line 320-321).
3. Replaced incorrect conjunction “even if” with “although” or “even though,” and applied the same correction consistently throughout the manuscript (e.g., line 869).
4. Corrected “images form” → “images from” (now line 558).
5. Revised subsection formatting for clarity (“3.1 Multi-hazard...” → “3.2 Multi-hazard...”).
6. Corrected “With regards to” → “With regard to,” consistently across the manuscript (e.g., line 637, 663, ...).
7. Reframed informal expression “showing popular results” as “Another popular DL architecture is GNN, applied in weather forecasting.” (now line 672)

Reviewer 2:

Major:

Comment 1: The manuscript outlines four key research questions, but their link to data-driven climate

multi-hazard analysis is not clear. Please clarify how these questions address existing gaps, improve risk assessments, and contribute quantitatively or qualitatively to the field.

Response: We thank the reviewer for the comment, allowing us to better clarify our research questions. In particular, we highlighted how each of the four research questions directly addresses existing gaps and contributes to advancing multi-hazard and multi-risk analysis in section 2.1. The first highlights how ML can harmonise diverse data sources to overcome issues of sparsity and bias; the second focuses on capturing spatio-temporal patterns and non-linear hazard dynamics beyond traditional methods; the third addresses the integration of vulnerability and exposure, often overlooked or treated as static; and the fourth explores future scenarios, where ML can enhance uncertainty quantification and scenario building. These refinements strengthen the link between the research questions, data-driven methods, and their contribution to improving the robustness and relevance of multi-risk assessments. (lines 212-300)

Revised text:

Section 2.1: Each of the four research questions (Figure 2) is focussed on a specific topic and presents several sub-topics, offering a structured framework to explore current applications, address challenges, and pinpoint future opportunities. These research questions are:

5. Data: How can Machine Learning improve data collection and processing?
6. Multi-Hazard: How can Machine Learning and statistical tools be used to analyse extreme events, and model hazard interactions?
7. Multi-Risk: How can Machine Learning applications integrate vulnerability and exposure in multi-risk analysis?
8. Future: How can Machine Learning and statistical tools be used to predict long-term future multi-hazard and multi-risk?

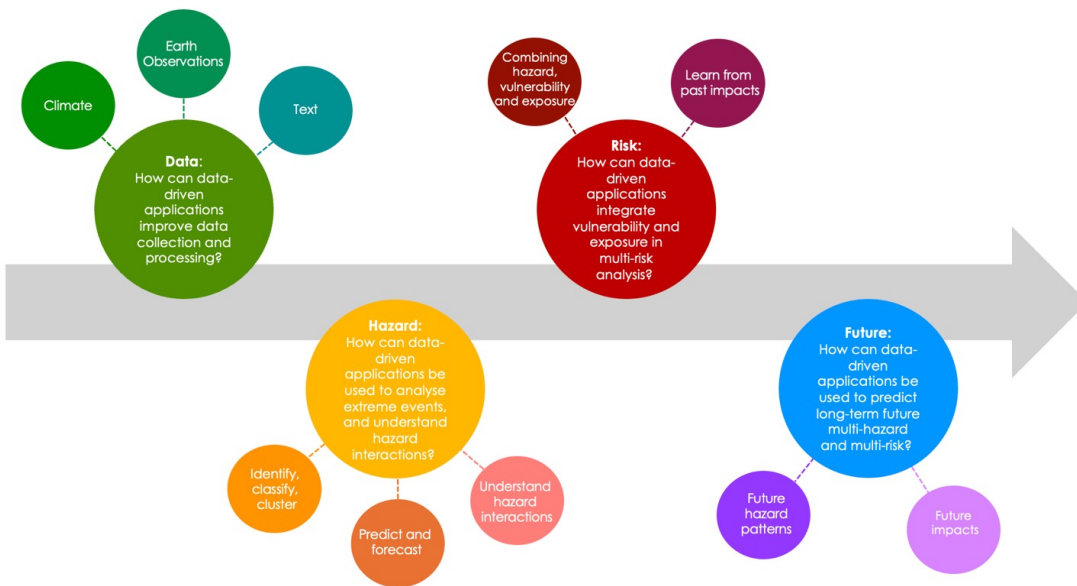


Figure 2: Research questions and sub-themes

The first research question examines how ML can help process diverse types of data, extracting and harmonising the information needed to analyse multi-hazard and multi-risk by addressing current gaps such as data sparsity, inconsistency across sources, and the lack of harmonised formats. This contributes to

improving the quality and comparability of risk assessments by enabling integrated use of climate, EO, and textual datasets. In particular, the sub-themes are divided based on the type of data analysed:

- IV. Climate data (time series of geospatial climate data), which describe the characteristics of climate-related hazards across space and time. Preparing this data for multi-hazard and multi-risk applications often requires ML methods (i.e. feature engineering) to increase spatial and temporal resolution, harmonise and extend the time coverage of the datasets or correct for biases (Schneider et al., 2023).
- V. EO, which can be used to characterise exposure and vulnerability layers and extract information on impacts (Ghaffarian & Emtehani, 2021; Novellino et al., 2024).
- VI. Textual data, such as newspapers or social media, which in the last years have been leveraged for extracting information on diverse impacts (Sodoge et al., 2023).

The second research question investigates how ML and statistical tools improve the identification and modelling of hazard dynamics by capturing complex spatio-temporal patterns, compounding effects, and non-linear interactions that traditional approaches often overlook. This helps advance multi-hazard and multi-risk analysis by providing more accurate detection, classification, and modelling of extreme events. In particular, the key sub-themes are:

- IV. Analyse which methods can be used to identify, classify and cluster extreme events, producing spatio-temporal footprints of multi-hazard events (H. Yu et al., 2022).
- V. The prediction of (multi-)hazard events, for example through early warning systems or seasonal predictions (Bhowmik et al., 2023).
- VI. The analysis of hazard interactions, for example characterising joint distributions through copulas (Bevacqua et al., 2021) or multi-hazard susceptibility maps (Pourghasemi et al., 2019).

The third research question concerns the application of ML for the integration of vulnerability and exposure into multi-risk analysis addressing the current gap where vulnerability and exposure are often treated as static or secondary layers rather than dynamic drivers of risk. This integration strengthens the ability of multi-risk assessments to capture how socio-economic conditions and adaptation measures interact with hazards to shape overall risk. In particular, the key themes are:

- III. Multi-hazard exposure and vulnerability on assessments, integrating susceptibility mapping with information on specific exposure layers, such as buildings and population (Rusk et al., 2022).
- IV. Modelling risk from past impacts data, often through supervised ML approaches that use hazard, vulnerability and exposure indicators as predictors (Dal Barco et al., 2024).

The fourth research question investigates the possible contribution of ML and statistical tools into the analysis of (long-term) future multi-hazard and multi-risk, where uncertainty associated with the representation of future extremes in climate projections further complicates risk modelling, highlighting a critical gap in existing approaches, which often fail to adequately capture compound and cascading extremes under changing climate conditions. This research question clarifies how ML can enhance scenario building, improve uncertainty quantification, and support more robust long-term multi-risk assessments. In particular, the key sub-themes are:

- III. Modelling future multi-hazard trends and spatial patterns using statistical methods, in particular for compound and consecutive events (Zscheischler et al., 2018).
- IV. Assessing future impacts based on climate change projections, often using methods trained on historical data and applied to ensembles of RCP projections (Park & Lee, 2020).

Comment 2: Data assimilation is a critical component of modern climate analysis yet is missing from the discussion. Include recent developments that demonstrate its role in enhancing multi-hazard/risk assessments.

Response: We thank the reviewer for highlighting this important aspect. We agree that data assimilation (DA) plays a key role in improving the quality and reliability of climate and hazard information. In response, we have expanded the discussion to explicitly acknowledge the relevance of ML for data assimilation, especially along with hybrid approaches (e.g., (He et al., 2023)). The revised text also discusses the challenges and opportunities of such hybrid frameworks, particularly their computational demands, calibration complexity, and interpretability issues, while emphasizing their relevance for multi-hazard prediction. However, given the distinct methodological scope of data assimilation, we have included a concise discussion rather than a full section, while acknowledging its importance and potential for future research and benchmarking efforts. (lines 407-409, - 419-421).

Comment 3: Uncertainty quantification in climate risk studies is necessary for the analysis of hazard/risk. However, they are not discussed.

Response: We thank the reviewer for this insightful comment. We fully agree that uncertainty quantification (UQ) is a critical component of climate and risk analysis. In response, we have added a dedicated discussion highlighting the sources of uncertainty in ML-based environmental datasets (aleatoric and epistemic) and the probabilistic and hybrid approaches developed to address them. Specifically, we now discuss applications using Bayesian Networks, Gaussian and Multi-fidelity Gaussian Processes, Convolutional Neural Processes, Monte Carlo Dropout, along with emerging deep generative models for ensemble-based UQ.

The revised text also emphasizes that while these methods significantly improve reliability, systematic quantification and propagation of uncertainty through the modelling chain with example for climate data (e.g., (Saha & Ravela, 2022; Tazi et al., (2024), line 422-479) and hazard estimates (e.g., Macdonald et al., (2025) and Nguyen et al., (2024), lines 657-658) – remain limited and warrant further research. This addition acknowledges the reviewer’s point and strengthens the manuscript’s coverage of methodological challenges in data-driven climate risk studies.

Revised text:

In section 3.1.1: Machine learning applications for climate and environmental datasets have greatly improved the reconstruction and downscaling of variables from sparse and irregular observations. However, a critical yet often under-addressed aspect in this field is uncertainty quantification (UQ), which is particularly relevant when these datasets are later used for hazard or risk assessments (Beven, 2018). Uncertainty in ML-based arises from multiple sources: Aleatoric uncertainty stems from the intrinsic variability and noise in the underlying measurements, such as sensor errors, missing satellite observations, or inconsistent temporal coverage; epistemic uncertainty originates from limited or biased training data and model structural choices (Xu et al., 2022). Several probabilistic approaches have been explicitly designed to represent spatial data uncertainty by learning distributions rather than deterministic predictions, mainly involving Bayesian Networks (BN) and Gaussian Processes (GP) (Siddique et al., 2022). For example, Multi-fidelity Gaussian Processes with a $5/2$ Matern kernel in particular, were used to downscale precipitation data from ERA-5 over high mountain terrain. Multi fidelity models combine low-fidelity observations (which are usually more numerous and less expensive to obtain) with high-fidelity ones. This makes the model more suited than other state-of-the-art machine learning methods for smaller datasets and able to quantify and narrow the uncertainty associated with the precipitation estimates, which is especially needed over ungauged areas and can be used to estimate the likelihood of extreme events that lead to floods or droughts (Tazi et al., 2024b). Andersson et al., (2023) applies Convolutional Neural Processes (ConvNPs), to suggest informative sensor placements by finding sites that maximally reduce prediction uncertainty, testing it for air temperature anomalies measurements in Antarctica. Convolutional Neural Processes (ConvNPs) extend the probabilistic framework of Gaussian Processes by learning flexible, data-driven covariance structures through neural networks. While traditional GPs provide robust uncertainty

estimates but suffer from scalability and stationarity constraints (M. Jiang et al., 2022), ConvNPs maintain a probabilistic foundation while scaling linearly with data size and accommodating irregular spatial inputs (Garnelo et al., 2018). DeepSensor1, a specific GitHub python package, was developed to facilitate the application of Neural Processes in environmental sciences, especially for downscaling, interpolation, sensor placement and data imputation. Monte Carlo Dropout (MCD) enhances epistemic uncertainty quantification in climate data and was tested on neural networks for probabilistic medium-range weather forecasting (Garg et al., 2022). Deep generative models such as diffusion or GAN frameworks can further approximate uncertainty by generating ensembles of plausible realisations that sample the predictive probability space (Ling et al., 2024; Saha & Ravela, 2022). Despite these advances, most studies still focus primarily on improving resolution and accuracy, while systematic approaches to quantifying and propagating uncertainty through the modelling chain, from data to hazard and risk estimates, remain limited (Beven, 2018). Addressing this challenge is crucial, as downstream risk assessments rely heavily on the reliability of the climate inputs that feed them.

In section 3.2.2: Monte Carlo dropout techniques have been employed to quantify epistemic uncertainty, for example in surge forecasts (Macdonald et al., 2025) and flood modelling (M. Nguyen et al., 2024).

Comment 4: The current focus is on pure data-driven models. Please discuss the emerging hybrid modelling that integrate physical laws, and compare their strengths and limitations.

Response: We thank the reviewer for highlighting the importance of hybrid modelling. In the revised manuscript, we have added a discussion on emerging hybrid approaches that integrate data-driven ML methods with physical or process-based models. We describe their ability to enforce physical constraints, reduce overfitting, improve generalization, and provide more interpretable insights, citing recent examples in climate modelling (He et al., 2023; Huynh et al., 2025; S. Yu et al., 2024). We also note the associated challenges, including the need for domain knowledge, computational demands, complex calibration, and potential interpretability issues (Read et al., 2019; Willard et al., 2022). These additions highlight hybrid modelling as an emerging frontier that complements purely data-driven approaches and are reflected in the revised manuscript (lines 403 – 421).

Revised text:

Section 3.1.1: In recent years, there has been growing interest in hybrid modelling: approaches that combine data-driven ML methods with physical or process-based models or constraints, as a way to benefit from both high flexibility and physical realism. Such hybrid / physics-informed ML methods help address several limitations of pure data-driven models: they can enforce conservation laws, reduce overfitting to noise, improve generalization especially under conditions outside the training domain, and provide more interpretable insights into underlying drivers. For instance, He et al., (2023) integrates ML corrections into a land-surface / atmospheric model using data assimilation, remote sensing LAI and soil moisture to improve climate simulations. Similarly, Huynh et al., (2025) combines process-based hydrological flux models with neural networks to correct for scale mismatches and to better capture spatial heterogeneity. Also, (S. Yu et al., 2024) provides benchmarks for ML emulators that mimic nested high-resolution physical simulations. Despite their promise, hybrid models also face important limitations. They often require substantial domain and physical knowledge to be formulated appropriately and to ensure physical consistency (Willard et al., 2022). Moreover, coupling ML architectures with numerical process models can remain computationally demanding, particularly for high-resolution simulations or large spatio-temporal domains (Reichstein et al., 2019). Calibration and validation can also be complex, as balancing the contributions of the physical and data-driven components often involves ad hoc or case-specific tuning (Read et al., 2019). Finally, interpretability may still be reduced when the ML component acts as a black box, obscuring how physical constraints shape predictions (Kashinath et al., 2021a). These challenges are

¹ <https://github.com/alan-turing-institute/deepsensor>

also relevant for hazard prediction, where process dynamics such as land–atmosphere feedback play central roles and require models that are both physically credible and statistically robust. Thus, hybrid models represent an emerging frontier at the interface of ML, process-based modeling, and data assimilation, particularly relevant for both climate data reconstruction and hazard modelling and deserve explicit consideration in future reviews and benchmarking efforts.

Comment 5: Clarify the differences and complementarities between ML and copula techniques. Using practical examples if possible.

Response: We thank the reviewer for this suggestion. In the revised manuscript, we clarify how ML and copula techniques provide complementary strengths for hazard and risk assessment. ML approaches excel at prediction and feature discovery, including identifying key drivers of hazards, but are limited by high data demands, training biases, and challenges in extrapolating beyond observed conditions. In contrast, copulas provide interpretable dependence structures and precise joint probability estimates for extremes, but scale poorly in high dimensions and often lack causal interpretability. We highlight examples where ML forecasts hazard variables, while copulas quantify the joint occurrence of extremes (e.g., Sadegh et al., 2017; Tilloy et al., 2019). We also note emerging hybrid approaches, such as T. Jiang et al. (2023), which combine ML classifiers with C-vine copulas to estimate the probability of consecutive drought events, illustrating how integrating ML-derived drivers into copula frameworks can enhance compound risk assessments. These revisions now emphasize the complementary roles of ML and copula methods and their potential synergy in practical applications (lines 692-746).

Revised text:

Section 3.2.2: Recent work has applied interpretable ML frameworks to hazard modelling, aiming not only at prediction but also at identifying key drivers. For instance, S. Jiang, Bevacqua, et al. (2022) and S. Jiang, Zheng, et al. (2022) used LSTMs to study river flooding in Europe, combining feature attribution methods such as Expected Gradients (Erion et al., 2021) and Additive Decomposition (Du et al., 2019) to disentangle the roles of snowmelt and precipitation. By running models across decades, they revealed shifts in dominant flood drivers, with precipitation becoming increasingly important. Other studies have applied gradient-based methods (Sun et al., 2021), CNN heatmaps (Patil et al., 2023), attention mechanisms (Castangia et al., 2023), and sensitivity analysis (Bentivoglio et al., 2023; Bonino et al., 2024; Kratzert et al., 2019). These advances improve interpretability, yet ML approaches remain limited by high data demands, sensitivity to training biases, and the difficulty of generalising beyond observed conditions (Bentivoglio et al., 2023). Their strength lies in prediction and uncovering nonlinear relationships, but the black-box nature of many models complicates causal modelling (Freeman et al., 2018).

While most ML studies focus on univariate hazards, compound events require methods that capture joint extremes. Copulas offer a flexible statistical framework to model dependence structures between variables, such as the co-occurrence of high river discharge, intense rainfall, and coastal surges (Hao & Singh, 2016; Nelsen, 2006). By decoupling marginal distributions from their dependence structure, copulas can assess joint probabilities of rare events with more precision than traditional multivariate models (Tilloy et al., 2019). Applications include pair copulas for compound flooding in Italy (Bevacqua et al., 2017), Joe copulas for concurrent river–coastal extremes (Sadegh et al., 2017), and copula-based Bayesian networks for flood–drought interactions (Couasnon et al., 2018). However, several challenges remain: selecting appropriate copula families is non-trivial (since different families imply different tail dependencies, yet many common families assume simplistic dependency or exchangeability) (Oh & Patton, 2015); capturing joint tail dependence becomes increasingly difficult in high dimensions (vines, mixtures, or hierarchical copulas may help but bring computational and inference burdens) (Simpson et al., 2020); physical drivers (e.g. precipitation skew, changing climate forcings, watershed characteristics) are often only indirectly

represented through marginal or covariate models (Hochrainer-Stigler et al., 2019). Therefore, while copulas are powerful for probabilistic risk quantification, they are less suited to dynamic forecasting or process-based understanding without additional model structure or ensembles (Tootoonchi et al., 2022).

Comparison and complementarities.

ML and copula methods approach hazard interactions from distinct perspectives. ML excels at prediction and feature discovery but struggles with transparency and extrapolation, while copulas provide interpretable dependence structures and joint probability estimates but scale poorly with dimensionality and lack causal interpretability. ML can identify critical hazard predictors and generate inputs, while copulas rigorously quantify their joint occurrence. Yet, few studies combine these strengths; most rely on either predictive ML or probabilistic copulas in isolation. For example, an LSTM may forecast river discharge under given precipitation and snowmelt conditions, while a copula model can then quantify the probability that extreme discharge co-occurs with extreme rainfall or sea-level rise. Together, ML and copulas can provide a more complete picture: ML enables forecasting and driver attribution, while copulas ensure rigorous treatment of dependence structures and joint extremes (Sadegh et al., 2017; Tilloy et al., 2019). Combining both approaches offers a promising pathway for advancing compound risk assessments. Some approaches, such as, T. Jiang et al., (2023) used a hybrid ML-copula method to estimate the probability of consecutive drought events (in particular from meteorological to ecological droughts), combining several ML classifiers (KNN, RF, SVM, ...) to estimate the propagation probability of meteorological drought given its characteristics, and C-vine copulas to model conditional probability model of the paired meteorological and ecological drought events. Closing this gap, for instance, by integrating ML-derived drivers into copula frameworks, or benchmarking ML-learned dependencies against copula-based models, represents a promising but underexplored direction for compound risk assessment.

Comment 6: Reframe the discussion to emphasize how Earth observation data, combined with ML and copula techniques, leads to improved multi-hazard and risk assessments that benefit decision-makers. Try to rephrase from the perspective of assessments instead of technical comparisons.

Response: We thank the reviewer for the suggestion to reframe this discussion. In the revised section (lines 481-545), we emphasize how Earth Observation (EO) data, when combined with ML and statistical techniques, supports actionable multi-hazard and multi-risk assessments. We highlight applications across the disaster management cycle, including early warning, improved hazard characterization, rapid impact and damage estimation, and long-term recovery and vulnerability monitoring. For example, EO-based indicators such as vegetation stress or surface temperature anomalies, combined with ML models, enable early detection of droughts, floods, or wildfires, even in remote or data-scarce regions. EO also enhances hazard and environmental datasets, mitigates biases in ML-based risk models, and supports assessment of environmental quality, urban recovery, and exposure tracking using proxies such as night-time lights. These additions illustrate the direct benefits of EO-ML integration for decision-makers and are now reflected in the revised manuscript.

Revised text:

Section 3.1.2:

EO data, when combined with ML is increasingly recognised for its critical role in supporting actionable multi-hazard and multi-risk assessment, as evidenced by new initiatives from ESA and NOAA's Centre for AI, where particular attention is devoted to the use of EO for discovering impacts in remote areas and developing early warning systems.

Remote sensing images are used to improve climate datasets, for example increasing the spatial coverage in areas with sparse measurements or providing real data to bias-correct/downscale modelled data. Multiple AI methods, such as Support Vector Machine (SVM) (Ahmad et al., 2010; Jing et al., 2016a), Ridge Regression (Kang et al., 2018), RF (Han et al., 2023; Jing et al., 2016b) and LSTM (Fang et al., 2017) are applied for developing soil moisture datasets.

EO provides consistent, near-real time observations of environmental conditions that are critical for early warning and hazard characterisation. For instance, indicators such as vegetation stress (Miyoshi et al., 2020; Schiefer et al., 2020; Veras et al., 2022), surface temperature anomalies can enable the early detection of droughts (Barrett et al., 2020), floods (Dasgupta et al., 2022) or wildfires (Jain et al., 2020) especially in remote and data scarce areas. DL and Physics Informed Neural Networks can leverage radar (e.g., Sentinel-1 SAR), to estimate water levels for flood forecasting (Dasgupta et al., 2022; Gierszewska & Berezowski, 2024) or fused into predictive models that refine hazard forecasts for severe weather and anticipate cascading impacts (Flora et al., 2021). Remote sensing plays a crucial role in hazard dataset development by helping mitigate bias that may be inherited by ML-based risk models. These models are often trained on datasets calibrated with data from resource-rich regions, where the majority of weather stations are located. As a result, they may struggle to generalize effectively to underdeveloped areas, which are frequently the most vulnerable to extreme events (McGovern et al., 2019, 2022).

EO combined with ML is also used in assessing environmental quality, such as water quality (Sagan et al., 2020; Sit et al., 2020). These applications mainly showcase simpler models, such as short neural networks and SVM (Nazeer et al., 2017), Decision Trees (DT), RF, Cubist Regression and Extreme Gradient Boosting (XGBoost), due to their ease of implementation and relative scarcity of ground measurement data (Liu et al., 2023). They focus on optically parameters, such as chlorophyll-a, turbidity and suspended solids, but also others such as of nutrients and other non-optical parameter can be predicted relying on models integrating meteorological and hydrological variables (S. Chen et al. 2022).

A central application of EO is in supporting impact and damage assessments: change detection techniques that compare pre- and post- event imagery are used to estimate physical impacts (T. Bai et al., 2023). This includes building damage (Y. Bai et al. 2018), infrastructure collapse (Sublime & Kalinicheva 2019) due to earthquakes or tsunamis (Ji et al. 2018), but also flood extent (Munawar et al., 2021), landslides (Lei et al., 2019) and wildfire scars (Bo et al., 2022; Tran et al., 2020). The main challenges encountered in these applications are due to the return periods of satellites, which may limit their ability to detect fast changing impacts; to the presence of clouds, which can hamper visibility especially during the occurrence of extreme events likely to cause damages; and to changes in luminosity or season (Faiza et al., 2012).

Moreover, EO enables long-term recovery tracking and vulnerability/exposure monitoring, with applications using proxies such as night-time lights to measure recovery trajectories (Kabiru et al., 2023; Qiang et al., 2020). For examples, studies have used EO and ML to track how rapidly services return to urban slums post disaster, highlighting which population remain exposed and underserved (Ghaffarian & Emtchani, 2021). Similarly, UNET-based CNNs are used to identify deprivation pockets from satellite images and track during their recovery process (Wang et al., 2019), or to derive proxy indicators for poverty from satellite night lights (Jean et al., 2016), in combination with transfer learning to overcome scarcity of labelled data (Pan & Yang, 2010). At longer timescales, techniques like K-Nearest Neighbour (KNN), SVM, ANN and RF are used to classify urban and rural land cover, detect land use changes or informal settlements (Adam et al., 2014; Yuh et al., 2023; Zerrouki et al., 2019).

In summary, the integration of EO with ML and statistical techniques offers a powerful toolkit for multi-hazard and multi-risk assessment, supporting early warning, targeted preparedness, rapid impact estimation, and recovery monitoring.

Other comments:

1. Citation consistency: We have carefully reviewed and corrected all citations to ensure consistent formatting throughout the manuscript, including references such as “Linkov et al. (2022)” and “S. Yu & Ma (2021).”
2. Terminology (“understand”): We agree with the reviewer’s observation. The term “understand” has been replaced with more appropriate alternatives such as “modeling,” “characterization,” or “representation” where applicable, to better reflect the capabilities of data-driven models.
3. Grammar and style: We have thoroughly revised the manuscript for grammatical accuracy and stylistic clarity, improving overall readability.
4. Section numbering: The numbering has been corrected so that the “Multi-hazard” subsection now appears as Section 3.2.

References:

- Adam, E., Mutanga, O., Odindi, J., & Abdel-Rahman, E. M. (2014). Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers. *International Journal of Remote Sensing*, *35*(10), 3440–3458. <https://doi.org/10.1080/01431161.2014.903435>
- Ahmad, S., Kalra, A., & Stephen, H. (2010). Estimating soil moisture using remote sensing data: A machine learning approach. *Advances in Water Resources*, *33*(1), 69–80. <https://doi.org/10.1016/j.advwatres.2009.10.008>
- Andersson, T. R., Bruinsma, W. P., Markou, S., Requeima, J., Coca-Castro, A., Vaughan, A., Ellis, A.-L., Lazzara, M. A., Jones, D., Hosking, S., & Turner, R. E. (2023). Environmental sensor placement with convolutional Gaussian neural processes. *Environmental Data Science*, *2*, e32. <https://doi.org/10.1017/eds.2023.22>
- Bai, T., Wang, L., Yin, D., Sun, K., Chen, Y., Li, W., & Li, D. (2023). Deep learning for change detection in remote sensing: a review. *Geo-Spatial Information Science*, *26*(3), 262–288. <https://doi.org/10.1080/10095020.2022.2085633>
- Bai, Y., Mas, E., & Koshimura, S. (2018). Towards Operational Satellite-Based Damage-Mapping Using U-Net Convolutional Network: A Case Study of 2011 Tohoku Earthquake-Tsunami. *Remote Sensing*, *10*(10), 1626. <https://doi.org/10.3390/rs10101626>
- Barrett, A. B., Duivenvoorden, S., Salakpi, E. E., Muthoka, J. M., Mwangi, J., Oliver, S., & Rowhani, P. (2020). *Forecasting vegetation condition for drought early warning systems in pastoral communities in Kenya*. <http://arxiv.org/abs/1911.10339>
- Bentivoglio, R., Isufi, E., Jonkman, S. N., & Taormina, R. (2023). Rapid spatio-temporal flood modelling via hydraulics-based graph neural networks. *Hydrology and Earth System Sciences*, *27*(23), 4227–4246. <https://doi.org/10.5194/hess-27-4227-2023>
- Bevacqua, E., De Michele, C., Manning, C., Couasnon, A., Ribeiro, A. F. S., Ramos, A. M., Vignotto, E., Bastos, A., Blesić, S., Durante, F., Hillier, J., Oliveira, S. C., Pinto, J. G., Ragno, E., Rivoire, P., Saunders, K., Wiel, K., Wu, W., Zhang, T., & Zscheischler, J. (2021). Guidelines for Studying Diverse Types of Compound Weather and Climate Events. *Earth’s Future*, *9*(11). <https://doi.org/10.1029/2021EF002340>
- Bevacqua, E., Maraun, D., Hobæk Haff, I., Widmann, M., & Vrac, M. (2017). Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy). *Hydrology and Earth System Sciences*, *21*(6), 2701–2723. <https://doi.org/10.5194/hess-21-2701-2017>

- Beven, K. (2018). *Environmental Modelling*. CRC Press. <https://doi.org/10.1201/9781482288575>
- Bhowmik, R. T., Jung, Y. S., Aguilera, J. A., Prunicki, M., & Nadeau, K. (2023). A multi-modal wildfire prediction and early-warning system based on a novel machine learning framework. *Journal of Environmental Management*, *341*, 117908. <https://doi.org/10.1016/j.jenvman.2023.117908>
- Bo, W., Liu, J., Fan, X., Tjahjadi, T., Ye, Q., & Fu, L. (2022). BASNet: Burned Area Segmentation Network for Real-Time Detection of Damage Maps in Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–13. <https://doi.org/10.1109/TGRS.2022.3197647>
- Bonino, G., Galimberti, G., Masina, S., McAdam, R., & Clementi, E. (2024). Machine learning methods to predict sea surface temperature and marine heatwave occurrence: a case study of the Mediterranean Sea. *Ocean Science*, *20*(2), 417–432. <https://doi.org/10.5194/os-20-417-2024>
- Castangia, M., Grajales, L. M. M., Aliberti, A., Rossi, C., Macii, A., Macii, E., & Patti, E. (2023). Transformer neural networks for interpretable flood forecasting. *Environmental Modelling & Software*, *160*, 105581. <https://doi.org/10.1016/j.envsoft.2022.105581>
- Chen, S., Zhang, Z., Lin, J., & Huang, J. (2022). Machine learning-based estimation of riverine nutrient concentrations and associated uncertainties caused by sampling frequencies. *PLOS ONE*, *17*(7), e0271458. <https://doi.org/10.1371/journal.pone.0271458>
- Couason, A., Sebastian, A., & Morales-Nápoles, O. (2018). A Copula-Based Bayesian Network for Modeling Compound Flood Hazard from Riverine and Coastal Interactions at the Catchment Scale: An Application to the Houston Ship Channel, Texas. *Water*, *10*(9), 1190. <https://doi.org/10.3390/w10091190>
- Dal Barco, M. K., Maraschini, M., Ferrario, D. M., Nguyen, N. D., Torresan, S., Vascon, S., & Critto, A. (2024). A machine learning approach to evaluate coastal risks related to extreme weather events in the Veneto region (Italy). *International Journal of Disaster Risk Reduction*, *108*, 104526. <https://doi.org/10.1016/j.ijdr.2024.104526>
- Dasgupta, A., Hybbeneth, L., & Waske, B. (2022). *Towards Daily High-resolution Inundation Observations using Deep Learning and EO*. <http://arxiv.org/abs/2208.09135>
- Du, M., Liu, N., Yang, F., Ji, S., & Hu, X. (2019). On Attribution of Recurrent Neural Network Predictions via Additive Decomposition. *The World Wide Web Conference*, 383–393. <https://doi.org/10.1145/3308558.3313545>
- Erion, G., Janizek, J. D., Sturmfels, P., Lundberg, S. M., & Lee, S.-I. (2021). Improving performance of deep learning models with axiomatic attribution priors and expected gradients. *Nature Machine Intelligence*, *3*(7), 620–631. <https://doi.org/10.1038/s42256-021-00343-w>
- Faiza, B., Yuhaniz, S. S., Hashim, S. Z. M., & AbdulRahman, K. K. (2012). A Review and Analysis of Image Misalignment Problem in Remote Sensing. *International Journal of Scientific and Engineering Research*, *3*(3), 82–86.
- Fang, K., Shen, C., Kifer, D., & Yang, X. (2017). Prolongation of SMAP to Spatiotemporally Seamless Coverage of Continental U.S. Using a Deep Learning Neural Network. *Geophysical Research Letters*, *44*(21). <https://doi.org/10.1002/2017GL075619>
- Flora, M. L., Potvin, C. K., Skinner, P. S., Handler, S., & McGovern, A. (2021). Using Machine Learning to Generate Storm-Scale Probabilistic Guidance of Severe Weather Hazards in the Warn-on-Forecast System. *Monthly Weather Review*, *149*(5), 1535–1557. <https://doi.org/10.1175/MWR-D-20-0194.1>

- Freeman, B. S., Taylor, G., Gharabaghi, B., & Thé, J. (2018). Forecasting air quality time series using deep learning. *Journal of the Air & Waste Management Association*, 68(8), 866–886. <https://doi.org/10.1080/10962247.2018.1459956>
- Garg, S., Rasp, S., & Thuerey, N. (2022). *WeatherBench Probability: A benchmark dataset for probabilistic medium-range weather forecasting along with deep learning baseline models*. <http://arxiv.org/abs/2205.00865>
- Garnelo, M., Schwarz, J., Rosenbaum, D., Viola, F., Rezende, D. J., Eslami, S. M. A., & Teh, Y. W. (2018). *Neural Processes*. <http://arxiv.org/abs/1807.01622>
- Ghaffarian, S., & Emtehani, S. (2021). Monitoring Urban Deprived Areas with Remote Sensing and Machine Learning in Case of Disaster Recovery. *Climate*, 9(4), 58. <https://doi.org/10.3390/cli9040058>
- Gierszewska, M., & Berezowski, T. (2024). *A physics-guided neural network for flooding area detection using SAR imagery and local river gauge observations*. <http://arxiv.org/abs/2410.08837>
- Han, Q., Zeng, Y., Zhang, L., Wang, C., Prikaziuk, E., Niu, Z., & Su, B. (2023). Global long term daily 1 km surface soil moisture dataset with physics informed machine learning. *Scientific Data*, 10(1), 101. <https://doi.org/10.1038/s41597-023-02011-7>
- Hao, Z., & Singh, V. P. (2016). Review of dependence modeling in hydrology and water resources. *Progress in Physical Geography: Earth and Environment*, 40(4), 549–578. <https://doi.org/10.1177/0309133316632460>
- He, X., Li, Y., Liu, S., Xu, T., Chen, F., Li, Z., Zhang, Z., Liu, R., Song, L., Xu, Z., Peng, Z., & Zheng, C. (2023). Improving regional climate simulations based on a hybrid data assimilation and machine learning method. *Hydrology and Earth System Sciences*, 27(7), 1583–1606. <https://doi.org/10.5194/hess-27-1583-2023>
- Hochrainer-Stigler, S., Balkovič, J., Silm, K., & Timonina-Farkas, A. (2019). Large scale extreme risk assessment using copulas: an application to drought events under climate change for Austria. *Computational Management Science*, 16(4), 651–669. <https://doi.org/10.1007/s10287-018-0339-4>
- Huynh, N. N. T., Garambois, P.-A., Renard, B., Colleoni, F., Monnier, J., & Roux, H. (2025). A distributed hybrid physics–AI framework for learning corrections of internal hydrological fluxes and enhancing high-resolution regionalized flood modeling. *Hydrology and Earth System Sciences*, 29(15), 3589–3613. <https://doi.org/10.5194/hess-29-3589-2025>
- Jain, P., Coogan, S. C. P., Subramanian, S. G., Crowley, M., Taylor, S., & Flannigan, M. D. (2020). *A review of machine learning applications in wildfire science and management*. <https://doi.org/10.1139/er-2020-0019>
- Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301), 790–794. <https://doi.org/10.1126/science.aaf7894>
- Ji, Y., Sri Sumantyo, J., Chua, M., & Waqar, M. (2018). Earthquake/Tsunami Damage Assessment for Urban Areas Using Post-Event PolSAR Data. *Remote Sensing*, 10(7), 1088. <https://doi.org/10.3390/rs10071088>
- Jiang, M., Pedrielli, G., & Ng, S. H. (2022). Gaussian Processes for High-Dimensional, Large Data Sets: A Review. *2022 Winter Simulation Conference (WSC)*, 49–60. <https://doi.org/10.1109/WSC57314.2022.10015416>

- Jiang, S., Bevacqua, E., & Zscheischler, J. (2022). River flooding mechanisms and their changes in Europe revealed by explainable machine learning. *Hydrology and Earth System Sciences*, 26(24), 6339–6359. <https://doi.org/10.5194/hess-26-6339-2022>
- Jiang, S., Zheng, Y., Wang, C., & Babovic, V. (2022). Uncovering Flooding Mechanisms Across the Contiguous United States Through Interpretive Deep Learning on Representative Catchments. *Water Resources Research*, 58(1). <https://doi.org/10.1029/2021WR030185>
- Jiang, T., Su, X., Zhang, G., Zhang, T., & Wu, H. (2023). Estimating propagation probability from meteorological to ecological droughts using a hybrid machine learning copula method. *Hydrology and Earth System Sciences*, 27(2), 559–576. <https://doi.org/10.5194/hess-27-559-2023>
- Jing, W., Yang, Y., Yue, X., & Zhao, X. (2016a). A Comparison of Different Regression Algorithms for Downscaling Monthly Satellite-Based Precipitation over North China. *Remote Sensing*, 8(10), 835. <https://doi.org/10.3390/rs8100835>
- Jing, W., Yang, Y., Yue, X., & Zhao, X. (2016b). A Spatial Downscaling Algorithm for Satellite-Based Precipitation over the Tibetan Plateau Based on NDVI, DEM, and Land Surface Temperature. *Remote Sensing*, 8(8), 655. <https://doi.org/10.3390/rs8080655>
- Kabiru, P., Kuffer, M., Sliuzas, R., & Vanhuyse, S. (2023). The relationship between multiple hazards and deprivation using open geospatial data and machine learning. *Natural Hazards*, 119(2), 907–941. <https://doi.org/10.1007/s11069-023-05897-z>
- Kang, J., Jin, R., Li, X., Zhang, Y., & Zhu, Z. (2018). Spatial Upscaling of Sparse Soil Moisture Observations Based on Ridge Regression. *Remote Sensing*, 10(2), 192. <https://doi.org/10.3390/rs10020192>
- Kashinath, K., Mustafa, M., Albert, A., Wu, J.-L., Jiang, C., Esmailzadeh, S., Azizzadenesheli, K., Wang, R., Chattopadhyay, A., Singh, A., Manepalli, A., Chirila, D., Yu, R., Walters, R., White, B., Xiao, H., Tchelepi, H. A., Marcus, P., Anandkumar, A., ... Prabhat. (2021). Physics-informed machine learning: case studies for weather and climate modelling. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194), 20200093. <https://doi.org/10.1098/rsta.2020.0093>
- Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., & Nearing, G. (2019). Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. *Hydrology and Earth System Sciences*, 23(12), 5089–5110. <https://doi.org/10.5194/hess-23-5089-2019>
- Lei, T., Zhang, Q., Xue, D., Chen, T., Meng, H., & Nandi, A. K. (2019). End-to-end Change Detection Using a Symmetric Fully Convolutional Network for Landslide Mapping. *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3027–3031. <https://doi.org/10.1109/ICASSP.2019.8682802>
- Ling, F., Lu, Z., Luo, J.-J., Bai, L., Behera, S. K., Jin, D., Pan, B., Jiang, H., & Yamagata, T. (2024). Diffusion model-based probabilistic downscaling for 180-year East Asian climate reconstruction. *Npj Climate and Atmospheric Science*, 7(1), 131. <https://doi.org/10.1038/s41612-024-00679-1>
- Liu, J., Qiu, Z., Feng, J., Wong, K. P., Tsou, J. Y., Wang, Y., & Zhang, Y. (2023). Monitoring Total Suspended Solids and Chlorophyll-a Concentrations in Turbid Waters: A Case Study of the Pearl River Estuary and Coast Using Machine Learning. *Remote Sensing*, 15(23), 5559. <https://doi.org/10.3390/rs15235559>
- Macdonald, E., Tubaldi, E., & Patelli, E. (2025). Robust storm surge forecasts for early warning system: a machine learning approach using Monte Carlo Bayesian model selection algorithm. *Stochastic*

Environmental Research and Risk Assessment, 39(7), 2789–2816. <https://doi.org/10.1007/s00477-025-02993-3>

- McGovern, A., Ebert-Uphoff, I., Gagne, D. J., & Bostrom, A. (2022). Why we need to focus on developing ethical, responsible, and trustworthy artificial intelligence approaches for environmental science. *Environmental Data Science*, 1, e6. <https://doi.org/10.1017/eds.2022.5>
- McGovern, A., Lagerquist, R., John Gagne, D., Jergensen, G. E., Elmore, K. L., Homeyer, C. R., & Smith, T. (2019). Making the Black Box More Transparent: Understanding the Physical Implications of Machine Learning. *Bulletin of the American Meteorological Society*, 100(11), 2175–2199. <https://doi.org/10.1175/BAMS-D-18-0195.1>
- Miyoshi, G. T., Arruda, M. dos S., Osco, L. P., Marcato Junior, J., Gonçalves, D. N., Imai, N. N., Tommaselli, A. M. G., Honkavaara, E., & Gonçalves, W. N. (2020). A Novel Deep Learning Method to Identify Single Tree Species in UAV-Based Hyperspectral Images. *Remote Sensing*, 12(8), 1294. <https://doi.org/10.3390/rs12081294>
- Munawar, H. S., Ullah, F., Qayyum, S., Khan, S. I., & Mojtahedi, M. (2021). UAVs in Disaster Management: Application of Integrated Aerial Imagery and Convolutional Neural Network for Flood Detection. *Sustainability*, 13(14), 7547. <https://doi.org/10.3390/su13147547>
- Nazeer, M., Bilal, M., Alsahli, M., Shahzad, M., & Waqas, A. (2017). Evaluation of Empirical and Machine Learning Algorithms for Estimation of Coastal Water Quality Parameters. *ISPRS International Journal of Geo-Information*, 6(11), 360. <https://doi.org/10.3390/ijgi6110360>
- Nelsen, R. (2006). *An Introduction to Copulas*. Springer New York. <https://doi.org/10.1007/0-387-28678-0>
- Nguyen, M., Wilson, M., Lane, E., Brasington, J., & Pearson, R. (2024). Estimating uncertainty in flood model outputs using machine learning informed by Monte Carlo analysis. *2024 International Conference on Machine Intelligence for GeoAnalytics and Remote Sensing (MIGARS)*, 1–3. <https://doi.org/10.1109/MIGARS61408.2024.10544837>
- Novellino, A., Pennington, C., Leeming, K., Taylor, S., Alvarez, I. G., McAllister, E., Arnhardt, C., & Winson, A. (2024). Mapping landslides from space: A review. *Landslides*, 21(5), 1041–1052. <https://doi.org/10.1007/s10346-024-02215-x>
- Oh, D. H., & Patton, A. J. (2015). Modelling Dependence in High Dimensions with Factor Copulas. *Finance and Economics Discussion Series*, 2015.0(51), 1–41. <https://doi.org/10.17016/feds.2015.051>
- Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>
- Park, S. J., & Lee, D. K. (2020). Prediction of coastal flooding risk under climate change impacts in South Korea using machine learning algorithms. *Environmental Research Letters*, 15(9). <https://doi.org/10.1088/1748-9326/ABA5B3>
- Patil, K. R., Doi, T., & Behera, S. K. (2023). Predicting extreme floods and droughts in East Africa using a deep learning approach. *Npj Climate and Atmospheric Science*, 6(1), 108. <https://doi.org/10.1038/s41612-023-00435-x>
- Pourghasemi, H. R., Gayen, A., Panahi, M., Rezaie, F., & Blaschke, T. (2019). Multi-hazard probability assessment and mapping in Iran. *Science of The Total Environment*, 692, 556–571. <https://doi.org/10.1016/j.scitotenv.2019.07.203>

- Qiang, Y., Huang, Q., & Xu, J. (2020). Observing community resilience from space: Using nighttime lights to model economic disturbance and recovery pattern in natural disaster. *Sustainable Cities and Society*, 57, 102115. <https://doi.org/10.1016/j.scs.2020.102115>
- Read, J. S., Jia, X., Willard, J., Appling, A. P., Zwart, J. A., Oliver, S. K., Karpatne, A., Hansen, G. J. A., Hanson, P. C., Watkins, W., Steinbach, M., & Kumar, V. (2019). Process-Guided Deep Learning Predictions of Lake Water Temperature. *Water Resources Research*, 55(11), 9173–9190. <https://doi.org/10.1029/2019WR024922>
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743), 195–204. <https://doi.org/10.1038/s41586-019-0912-1>
- Rusk, J., Maharjan, A., Tiwari, P., Chen, T.-H. K., Shneiderman, S., Turin, M., & Seto, K. C. (2022). Multi-hazard susceptibility and exposure assessment of the Hindu Kush Himalaya. *Science of The Total Environment*, 804, 150039. <https://doi.org/10.1016/j.scitotenv.2021.150039>
- Sadegh, M., Ragno, E., & AghaKouchak, A. (2017). Multivariate Copula Analysis Toolbox (MvCAT): Describing dependence and underlying uncertainty using a Bayesian framework. *Water Resources Research*, 53(6), 5166–5183. <https://doi.org/10.1002/2016WR020242>
- Sagan, V., Peterson, K. T., Maimaitijiang, M., Sidike, P., Sloan, J., Greeling, B. A., Maalouf, S., & Adams, C. (2020). Monitoring inland water quality using remote sensing: potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. *Earth-Science Reviews*, 205, 103187. <https://doi.org/10.1016/j.earscirev.2020.103187>
- Saha, A., & Ravela, S. (2022). *Downscaling Extreme Rainfall Using Physical-Statistical Generative Adversarial Learning*. <https://doi.org/https://doi.org/10.48550/arXiv.2212.01446>
- Schiefer, F., Kattenborn, T., Frick, A., Frey, J., Schall, P., Koch, B., & Schmidtlein, S. (2020). Mapping forest tree species in high resolution UAV-based RGB-imagery by means of convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 170, 205–215. <https://doi.org/10.1016/j.isprsjprs.2020.10.015>
- Schneider, T., Behera, S., Boccaletti, G., Deser, C., Emanuel, K., Ferrari, R., Leung, L. R., Lin, N., Müller, T., Navarra, A., Ndiaye, O., Stuart, A., Tribbia, J., & Yamagata, T. (2023). Harnessing AI and computing to advance climate modelling and prediction. *Nature Climate Change*, 13(9), 887–889. <https://doi.org/10.1038/s41558-023-01769-3>
- Simpson, E. S., Wadsworth, J. L., & Tawn, J. A. (2020). *A geometric investigation into the tail dependence of vine copulas*. <https://doi.org/10.1016/j.jmva.2021.104736>
- Sit, M., Demiray, B. Z., Xiang, Z., Ewing, G. J., Sermet, Y., & Demir, I. (2020). A comprehensive review of deep learning applications in hydrology and water resources. *Water Science and Technology*, 82(12), 2635–2670. <https://doi.org/10.2166/wst.2020.369>
- Sodoge, J., Kuhlicke, C., & de Brito, M. M. (2023). Automated spatio-temporal detection of drought impacts from newspaper articles using natural language processing and machine learning. *Weather and Climate Extremes*, 41, 100574. <https://doi.org/10.1016/j.wace.2023.100574>
- Sublime, J., & Kalinicheva, E. (2019). Automatic Post-Disaster Damage Mapping Using Deep-Learning Techniques for Change Detection: Case Study of the Tohoku Tsunami. *Remote Sensing*, 11(9), 1123. <https://doi.org/10.3390/rs11091123>
- Sun, A. Y., Jiang, P., Mudunuru, M. K., & Chen, X. (2021). Explore Spatio-Temporal Learning of Large Sample Hydrology Using Graph Neural Networks. *Water Resources Research*, 57(12). <https://doi.org/10.1029/2021WR030394>

- Tazi, K., Orr, A., Hernandez-González, J., Hosking, S., & Turner, R. E. (2024a). Downscaling precipitation over High-mountain Asia using multi-fidelity Gaussian processes: improved estimates from ERA5. *Hydrology and Earth System Sciences*, 28(22), 4903–4925. <https://doi.org/10.5194/hess-28-4903-2024>
- Tazi, K., Orr, A., Hernandez-González, J., Hosking, S., & Turner, R. E. (2024b). Downscaling precipitation over High-mountain Asia using multi-fidelity Gaussian processes: improved estimates from ERA5. *Hydrology and Earth System Sciences*, 28(22), 4903–4925. <https://doi.org/10.5194/hess-28-4903-2024>
- Tilloy, A., Malamud, B. D., Winter, H., & Joly-Laugel, A. (2019). A review of quantification methodologies for multi-hazard interrelationships. *Earth-Science Reviews*, 196, 102881. <https://doi.org/10.1016/j.earscirev.2019.102881>
- Tootoonchi, F., Sadegh, M., Haerter, J. O., Rätty, O., Grabs, T., & Teutschbein, C. (2022). Copulas for hydroclimatic analysis: A practice-oriented overview. *WIREs Water*, 9(2). <https://doi.org/10.1002/wat2.1579>
- Tran, D. Q., Park, M., Jung, D., & Park, S. (2020). Damage-Map Estimation Using UAV Images and Deep Learning Algorithms for Disaster Management System. *Remote Sensing*, 12(24), 4169. <https://doi.org/10.3390/rs12244169>
- Veras, H. F. P., Ferreira, M. P., da Cunha Neto, E. M., Figueiredo, E. O., Corte, A. P. D., & Sanquetta, C. R. (2022). Fusing multi-season UAS images with convolutional neural networks to map tree species in Amazonian forests. *Ecological Informatics*, 71, 101815. <https://doi.org/10.1016/j.ecoinf.2022.101815>
- Wang, J., Kuffer, M., Roy, D., & Pfeffer, K. (2019). Deprivation pockets through the lens of convolutional neural networks. *Remote Sensing of Environment*, 234, 111448. <https://doi.org/10.1016/j.rse.2019.111448>
- Willard, J., Jia, X., Xu, S., Steinbach, M., & Kumar, V. (2022). *Integrating Scientific Knowledge with Machine Learning for Engineering and Environmental Systems*. <http://arxiv.org/abs/2003.04919>
- Xu, L., Chen, N., Yang, C., Yu, H., & Chen, Z. (2022). Quantifying the uncertainty of precipitation forecasting using probabilistic deep learning. *Hydrology and Earth System Sciences*, 26(11), 2923–2938. <https://doi.org/10.5194/hess-26-2923-2022>
- Yu, H., Lu, N., Fu, B., Zhang, L., Wang, M., & Tian, H. (2022). Hotspots, co-occurrence, and shifts of compound and cascading extreme climate events in Eurasian drylands. *Environment International*, 169, 107509. <https://doi.org/10.1016/j.envint.2022.107509>
- Yu, S., Hu, Z., Subramaniam, A., Hannah, W., Peng, L., Lin, J., Bhouri, M. A., Gupta, R., Lütjens, B., Will, J. C., Behrens, G., Busecke, J. J. M., Loose, N., Stern, C. I., Beucler, T., Harrop, B., Heuer, H., Hillman, B. R., Jenney, A., ... Pritchard, M. (2024). *ClimSim-Online: A Large Multi-scale Dataset and Framework for Hybrid ML-physics Climate Emulation*. <http://arxiv.org/abs/2306.08754>
- Yuh, Y. G., Tracz, W., Matthews, H. D., & Turner, S. E. (2023). Application of machine learning approaches for land cover monitoring in northern Cameroon. *Ecological Informatics*, 74, 101955. <https://doi.org/10.1016/j.ecoinf.2022.101955>
- Zerrouki, N., Harrou, F., Sun, Y., & Hocini, L. (2019). A Machine Learning-Based Approach for Land Cover Change Detection Using Remote Sensing and Radiometric Measurements. *IEEE Sensors Journal*, 19(14), 5843–5850. <https://doi.org/10.1109/JSEN.2019.2904137>
- Zscheischler, J., Westra, S., van den Hurk, B. J. J. M., Seneviratne, S. I., Ward, P. J., Pitman, A., AghaKouchak, A., Bresch, D. N., Leonard, M., Wahl, T., & Zhang, X. (2018). Future climate risk

from compound events. *Nature Climate Change*, 8(6), 469–477. <https://doi.org/10.1038/s41558-018-0156-3>