

GRAPHVSSM: Graph Variational State-Space Model for Probabilistic Spatiotemporal Inference of Dynamic Exposure and Vulnerability for Regional Disaster Resilience Assessment

Joshua Dimasaka^{1,2}, Christian Geiß^{3,4}, Emily So^{1,2}

¹Department of Architecture, University of Cambridge, United Kingdom

²Cambridge University Centre for Risk in the Built Environment, United Kingdom

³Earth Observation Center, German Aerospace Center, Weßling, Germany

⁴Institute of Geography, University of Bonn, Bonn, Germany

jtd33@cam.ac.uk

Abstract

Regional disaster resilience quantifies the changing nature of physical risks to inform policy instruments ranging from local immediate recovery to international sustainable development. While many existing state-of-practice methods have greatly advanced the dynamic mapping of exposure and hazard, our understanding of large-scale physical vulnerability has remained static, costly, limited, region-specific, coarse-grained, overly aggregated, and inadequately calibrated. With the significant growth in the availability of time-series satellite imagery and derived products for exposure and hazard, we focus our work on the equally important yet challenging element of the risk equation: physical vulnerability. Given this unique problem, we leverage machine learning methods that flexibly capture spatial contextual relationships, limited temporal observations, and uncertainty in a unified probabilistic spatiotemporal inference framework. We therefore introduce Graph Variational State-Space Model (**GRAPHVSSM**), a novel modular spatiotemporal approach that uniquely integrates graph deep learning, state-space modeling, and variational inference using time-series data and prior expert belief systems in a weakly supervised or coarse-to-fine-grained manner. We present three major results: a city-wide demonstration in Quezon City, Philippines; an investigation of sudden changes in the cyclone-impacted coastal Khurushkul community (Bangladesh) and the mudslide-affected Freetown (Sierra Leone); and an open geospatial dataset, **METEOR 2.5D**, that spatiotemporally enhances the existing global static dataset for 46 UN-recognized Least Developed Countries (as of 2020). Beyond advancing the practice of regional disaster resilience assessment and improving our understanding of global progress in disaster risk reduction, our method also offers a probabilistic deep learning approach, contributing to broader urban studies that require compositional data analysis in weakly supervised settings.

Code — <https://github.com/riskaudit/GraphVSSM>

Quezon City (Philippines) Dataset — <https://doi.org/pzj2>

METEOR 2.5D Dataset, Part 1/2 — <https://doi.org/pzq4>

METEOR 2.5D Dataset, Part 2/2 — <https://doi.org/pzrd>

Khurushkul-Freetown Dataset — <https://doi.org/pzkw>

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

1 Introduction

When a large earthquake, widespread flooding, raging wildfire, or any natural hazard strikes, we frequently hear the call for “resilience”, which has unfortunately become increasingly ambiguous and overused in many recent global climate conversations (Bogardi and Fekete 2019; Parker 2020). In contrast, when designing buildings and critical infrastructure, quantifying physical (or asset) resilience follows a clearer framework as a dynamic risk, which is a temporal convolution of three major changing elements – exposure, vulnerability, and hazard – through time (Bruneau et al. 2003; FEMA 2012; Almufti and Willford 2013). Varying temporal scales of dynamic risk analysis serve different yet interconnected purposes, ranging from immediate recovery efforts (Comerio 2006) at daily-to-annual scales to development-oriented climate finance (ADB 2017) at longer intervals, such as five-year periods. For example, low- and middle-income countries in the Caribbean, Pacific, and Africa have begun to benefit from modern international climate finance instruments, such as regional catastrophic risk pools (Ciullo et al. 2023). This underscores the importance of large-scale monitoring of changing patterns of exposure and vulnerability, amid intensifying climatic hazards, to better assess regional disaster resilience.

Despite recent advances in dynamic exposure and hazard modelling – such as Google Open Buildings 2.5D Temporal (Sirko et al. 2023), DLR World Settlement Footprint Evolution (Marconcini et al. 2021), Global Human Settlement Layer multitemporal products (Pesaresi et al. 2024), Microsoft Aurora (Bodnar et al. 2025), Google GraphCast (Lam et al. 2023), and ECMWF AIFS (Lang et al. 2024) – progress in mapping the dynamic nature of physical vulnerability remains limited (UNDRR 2025). Current approaches still rely heavily on the static information from the GEM Global Exposure Map and Vulnerability Model (Yepes-Estrada et al. 2023; Martins and Silva 2023) and the METEOR project (Huyck et al. 2019). Although analytical Bayesian probabilistic approaches (Pittore, Haas, and Silva 2020; Gómez Zapata et al. 2022) and various dasymmetric disaggregation techniques (Geiß et al. 2023; Dimasaka, Geiß, and So 2025) have enhanced spatial accuracy, the intricate subjectivity and coarse-grained representation of physi-

cal vulnerability as an ungeneralizable label across different construction practices indicate that solutions must be highly region-specific and consider the prevailing weakly supervised setup in both spatial and temporal domains.

With the significant growth in the availability of time-series satellite imagery, its rich information allows us to recognize relevant patterns and tackle these challenges using machine learning methods that flexibly capture spatial relationships, temporal dynamics, and uncertainty in a unified framework (Rolf et al. 2024). In particular, graph deep learning methods have enhanced our ability to model spatial relationships of irregularly structured data (e.g., connections between buildings) compared to conventional convolutional neural networks (Bronstein et al. 2021; Sheikh and Saha 2025). Additionally, deep state-space methods combine probabilistic sequential learning with prior knowledge specification and deep variational inference for uncertainty quantification (e.g., the changes in the likelihood of a highly vulnerable label given an existing coarse-grained belief system) (Girin et al. 2020; Lin and Michailidis 2024). Together, these advanced capabilities open new opportunities to address the gap in modeling the interpretable regional dynamics of exposure and physical vulnerability towards a more fine-grained and accurate large-scale mapping while incorporating existing and valuable prior domain knowledge.

In this paper, our key contributions are:

- **GRAPHVSSM** or Graph Variational State-Space Model, a novel spatiotemporal method that integrates graph deep learning, state-space modeling, and variational inference using time-series satellite imagery and prior expert belief systems in a weakly supervised manner.
- A probabilistic graph-based deep learning framework that advances the purely analytical Bayesian updating approach of previous work (Pittore, Haas, and Silva 2020), enabling the learning of complex relationships between building height, local spatial connectivity, and region-specific vulnerability labels.
- A city-wide validation of the assumed probability distributions for various indicators of regional exposure and physical vulnerability, using Quezon City in the Philippines as a case study.
- A high-resolution demonstration of yearly changes in physical vulnerability from 2016 to 2023 in the cyclone-impacted coastal Khurushkul community in Bangladesh (also known as “*the world’s largest climate refugee rehabilitation project*” (Khan et al. 2024)) and the mudslide-affected Freetown in Sierra Leone in 2017.
- A novel open geospatial dataset, **METEOR 2.5D**, that spatially refines the existing METEOR dataset (i.e., from 450-meter to 90-meter scale) and significantly extends it with country-wide dynamic evolution of regional exposure and physical vulnerability for 46 UN-recognized Least Developed Countries as of 2020, provided at five-year intervals between 1975 and 2030.

2 Related Work

This section reviews related literature across three key topics that motivate our development of **GRAPHVSSM** for mod-

eling the dynamic geospatial nature of physical vulnerability. We first examine current approaches and assumptions in analytical and statistical dynamic models used in disaster risk science, highlighting the distinct challenge of weak supervision. We then discuss how their limitations motivate the use of graph deep learning for efficient representation learning of building attributes. Next, we highlight the suitability of deep state-space models for handling short time series, which is another core aspect of our problem setting. Finally, we present the emerging concept of a graph state-space model and describe how extending it with variational inference supports a unified probabilistic framework for capturing uncertainty within disaster risk practice.

2.1 Dynamic Mapping of Physical Vulnerability

Along with the rapid growth of volunteered geographic information and satellite-derived products, many recent studies have developed a variety of analytical and probabilistic mapping approaches for dynamically updating the distribution of physical vulnerability across space and time (Cremen, Galasso, and McCloskey 2022). Relevant to our scope of large-scale mapping, the pioneering work of Porter et al. (2014) introduced the use of Beta probability distribution to perform Bayesian updating of the likelihood of a single building type using outcomes from field surveys. Building on categorical Bayesian inference (Agesti and Hitchcock 2005), Pittore, Haas, and Silva (2020) extended this approach using Dirichlet and Multinomial probability distributions to account for the correlation among different building types.

In contrast to these past studies that explicitly incorporate sparse in situ data, Lallemand (2015) applied a cellular automata approach and Markov chain modeling to dynamically update the physical vulnerability (e.g., incrementally expanding buildings), despite the lack of data calibration (Lallemand et al. 2017). Differently, Calderon and Silva (2022) proposed representing each location as a multi-agent system whose behavior depends on a geographically weighted regression model. As a more interpretable and manual yet costly approach, Schorlemmer et al. (2020) synthesized large volumes of crowdsourced information, such as OpenStreetMap, using a rule-based approach.

All of these past studies point toward the same prevailing challenge posed by the weak supervision setting, where we aim to refine coarse-grained vulnerability labels into finer spatiotemporal scales. Our study seeks to advance these state-of-practice methods by using deep learning to flexibly capture rich information from diverse datasets, including time-series satellite imagery, while also incorporating prior expert belief systems. This enables us to derive probabilistic representations that are better suited to disaster risk practice.

2.2 Graph Deep Learning of Building Attributes

Problems in weather, climate, and urban modeling that utilize satellite imagery have seen significant advances through the adoption of graph deep learning. The technique leverages contextual information and handles unstructured data, making it particularly well-suited for spatiotemporal graph-structured datasets (Zhao et al. 2024).

For instance, Fill, Eichelbeck, and Ebner (2024) constructed distance-based subgraphs to predict residential building type across multiple countries, outperforming convolutional neural networks. Likewise, Xu et al. (2022b) and Lei et al. (2024) demonstrated that graph deep learning achieved higher performance than support vector machines and random forests. Kong et al. (2024) further emphasized its ability to capture correlations among different building types. For other unstructured data, Dimasaka, Selvakumaran, and Marinoni (2024) effectively linked sparse settlement areas with complex road networks at large scales.

Therefore, considering the sparsity of the building footprints in our problem setting, our study combines deep learning with graph-structured representations to provide greater flexibility in modeling diverse representations and attributes of the built environment.

2.3 Deep State-Space Model for Short Sequences

In time-series forecasting, real-world scenarios with limited historical observations – particularly in spatiotemporal studies that rely on coarse temporal resolutions, such as yearly, five-year, or ten-year intervals – motivate the use of simpler models to avoid the overfitting tendency of more complex ones (Lim and Zohren 2021). In our problem setting where publicly available moderate-resolution satellite imagery has recently been introduced to inform key global frameworks through 2030, the characteristics of temporal data in terms of information quantity, irregularity, and sequence length are a crucial consideration in deciding between data-driven or deep forecasting models (Benidis et al. 2022).

To tackle these challenges, Benidis et al. (2022) further noted that deep state-space models outperform purely autoregressive and discriminative models, particularly in low-data regimes (de Bézenac et al. 2020) and when dealing with noisy or incomplete data (e.g., gaps in satellite imagery or sparsity of satellite-derived products). Specifically, Rangapuram et al. (2018) showed that hybridizing state-space models with deep learning facilitates an interpretable and generalizable framework by efficiently reducing the hypothesis space using input prior knowledge to the latent states (Grover, Kapoor, and Horvitz 2015) and combining the flexibility of deep neural networks with the useful structural assumptions of traditional state-space models via latent state representations (Durbin and Koopman 2012). Hence, in our study, we adopt a deep state-space model because its explicit structural assumptions align well with estimating the likelihood of physical vulnerability classes based on building height patterns and prior knowledge from expert-informed census-derived products.

Building on the preceding section on graph deep learning, our study therefore advances deep state-space modeling by integrating graph neural networks, which capture more inductive relational biases in terms of neighborhood effects among buildings that share the same local contextual characteristics. While early formulations of graph state-space models have primarily used encoder-decoder architectures (Zambon et al. 2023; Cini et al. 2025), we expand this approach by incorporating variational learning to parameterize the latent space representations probabilistically (Blei,

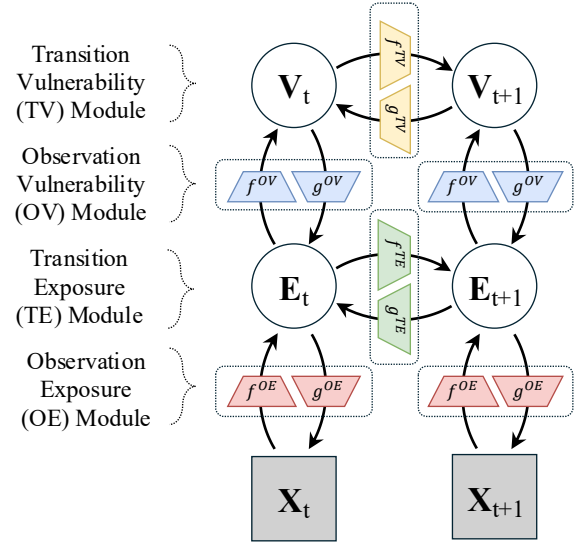


Figure 1: A schematic diagram of **GRAPHVSSM** consisting of four key modular components, where \mathbf{X} , \mathbf{E} , and \mathbf{V} represent any independent explanatory variables (e.g., signals of satellite imagery and other auxiliary data such as proximity to road networks), exposure (e.g., probability of building presence and height), and vulnerability (e.g., probability of a particular building typology), respectively. Each module has a variational autoencoder with graph convolutional networks (Kipf and Welling 2016) for encoder (i.e., $f \in \{f^{OE}, f^{TE}, f^{OV}, f^{TV}\}$) and decoder (i.e., $g \in \{g^{OE}, g^{TE}, g^{OV}, g^{TV}\}$), signifying relationships between and dynamics within \mathbf{X} , \mathbf{E} , and \mathbf{V} .

Kucukelbir, and McAuliffe 2017), leveraging the noisy signals of satellite imagery. This framing aligns well with the probabilistic treatment within disaster risk science, enabling our proposed approach to coherently quantify uncertainties across regional exposure, vulnerability, and hazard (Ward et al. 2020).

3 Graph Variational State-Space Model

This section presents the formulation of **GRAPHVSSM**, as illustrated in Figure 1. We first discuss the probabilistic representations of our exposure and physical vulnerability, followed by how these representations are integrated into each modular component of our graph state-space model. Next, we present the supervised variational learning that incorporates input prior knowledge. Finally, we show how this approach benefits from the expressive representation of spatiotemporal graphs, including its features and assumptions behind our design of the adjacency matrices.

3.1 Probabilistic Models

We define the exposure \mathbf{E} with two distinct probabilistic assumptions for the continuous and discrete variables of

building height and presence, respectively. Upon initial empirical investigation of their signal distributions and considering the non-negative constraint, we assume building height (\mathbf{E}_{BH}) and presence (\mathbf{E}_{BP}) to follow lognormal and Bernoulli probability distributions, respectively, as:

$$\ln \mathbf{E}_{\text{BH}} \in \{\ln \mathbf{E}_{\text{BH}}^{\text{OE}}, \ln \mathbf{E}_{\text{BH}}^{\text{TE}}\} \sim \mathcal{N}(\boldsymbol{\mu}_\theta(\mathbf{X}), \boldsymbol{\Sigma}_\theta(\mathbf{X})) \quad (1)$$

$$\mathbf{E}_{\text{BP}} \in \{\mathbf{E}_{\text{BP}}^{\text{OE}}, \mathbf{E}_{\text{BP}}^{\text{TE}}\} \sim \text{Bern}(\mathbf{p}_\theta(\mathbf{X})) \quad (2)$$

where $\boldsymbol{\mu}_\theta$, $\boldsymbol{\Sigma}_\theta$, and \mathbf{p}_θ (i.e., via logit ℓ_θ) are outcomes of our encoder networks, f^{OE} and f^{TE} , with learnable parameters, θ .

Moreover, following the analytical formulation of Pittore, Haas, and Silva (2020), we also represent the vulnerability \mathbf{V} as a categorical multinomial random variable:

$$\mathbf{V} \in \{\mathbf{V}^{\text{OV}}, \mathbf{V}^{\text{TV}}\} \sim \text{Mult}(\mathbf{p}_\theta^1(\mathbf{X}), \dots, \mathbf{p}_\theta^K(\mathbf{X})) \quad (3)$$

where \mathbf{p}_θ are outcomes of our encoder networks, f^{OV} and f^{TV} , with learnable parameters θ .

In a given single pixel that can have varying spatial resolutions in our experiments, we want to determine the probability of observing k^{th} building typology from all possible K types. Unlike their approach that uses discrete counts of buildings (Pittore, Haas, and Silva 2020), we improve this analytical approach to use the rasterized proportions of buildings, enabling the efficient use of geospatial data across large areal extents.

3.2 State-Space Model Components

For ease of demonstration, we sequentially train the modules in the following order: **OE**, **TE**, **OV**, and **TV**. This modularity provides flexibility, for cases where only the relationship between \mathbf{E} and \mathbf{V} is found to be more relevant than the well-studied task of predicting or forecasting \mathbf{E} . Nevertheless, our full training has still enabled us to establish and validate our probabilistic assumptions for \mathbf{E} and \mathbf{V} .

OE: Observation Exposure Module The encoder network f^{OE} inputs the \mathbf{X} covariates and jointly outputs the probabilistic parameters $\boldsymbol{\mu}_\theta^{\text{OE}}$, $\boldsymbol{\Sigma}_\theta^{\text{OE}}$, and $\mathbf{p}_\theta^{\text{OE}}$ for $\mathbf{E}_{\text{BH}}^{\text{OE}}$ and $\mathbf{E}_{\text{BP}}^{\text{OE}}$. The decoder network g^{OE} inputs the samples of $\mathbf{E}_{\text{BH}}^{\text{OE}}$ and $\mathbf{E}_{\text{BP}}^{\text{OE}}$ and outputs $\hat{\mathbf{X}}$, a reconstruction of our covariates.

TE: Transition Exposure Module The encoder network f^{TE} inputs the samples of $\mathbf{E}_{\text{BH}}^{\text{OE}}$ and $\mathbf{E}_{\text{BP}}^{\text{OE}}$ at time t , appended with the previous g^{TE} -reconstructed covariates (i.e., not the original \mathbf{X}) for the same time step, t .

Then, it jointly outputs a different set of probabilistic parameters $\boldsymbol{\mu}_\theta^{\text{TE}}$, $\boldsymbol{\Sigma}_\theta^{\text{TE}}$, and $\mathbf{p}_\theta^{\text{TE}}$ for $\mathbf{E}_{\text{BH}}^{\text{TE}}$ and $\mathbf{E}_{\text{BP}}^{\text{TE}}$ for the future time step, $t + 1$.

Similarly, the decoder network g^{TE} inputs the samples of $\mathbf{E}_{\text{BH}}^{\text{TE}}$ and $\mathbf{E}_{\text{BP}}^{\text{TE}}$ from a future time step, $t + 1$, and backward reconstructs our inputs to f^{TE} , for the preceding step, t .

After training the f^{TE} and g^{TE} , we sample $\mathbf{E}_{\text{BP}}^{\text{TE}}$ for the entire time horizon to identify the reduced set of pixels with a high likelihood of building presence (i.e., $\boldsymbol{\mu}_\theta^{\text{TE}} \geq 0.5$), instead of using the full map. This new set of pixels concludes the probabilistic modeling of \mathbf{E} and introduces a different set of adjacency matrices for the succeeding probabilistic modeling of \mathbf{V} .

OV: Observation Vulnerability Module The encoder network f^{OV} inputs the samples of $\mathbf{E}_{\text{BH}}^{\text{TE}}$ and the previously g^{TE} -reconstructed covariates. Then, it outputs the probabilistic K parameters $\mathbf{p}_\theta^{\text{OV}}$, corresponding to the K types of physical vulnerability. The decoder network g^{OV} inputs the samples from this multinomial probability distribution and outputs a further reconstruction of our covariates.

TV: Transition Vulnerability Module The encoder network f^{TV} inputs the samples of \mathbf{V}^{OV} at time t , appended with the previous g^{OV} -reconstructed covariates for the same time step, t . Then, it outputs a different set of probabilistic K parameters $\mathbf{p}_\theta^{\text{TV}}$ for the next time step, $t + 1$. Similarly, the decoder network g^{TV} inputs the samples of \mathbf{V}^{TV} from a future step, $t + 1$, and then backward reconstructs our inputs to f^{TV} , for the preceding step, t .

3.3 Variational Learning

Encoder Reparameterization Tricks To enable the variational nature, we implement the following reparameterization trick for the stochastic samplings of our building height variable using:

$$\mathbf{E}_{\text{BH}}^* = e^{\boldsymbol{\mu}_\theta + \boldsymbol{\Sigma}_\theta^{1/2} \boldsymbol{\epsilon}^*}, \quad \boldsymbol{\epsilon}^* \sim \mathcal{N}(0, 1) \quad (4)$$

Similarly, the variables for building presence and proportion of vulnerability types follow the same categorical parameterization trick via Gumbel-Softmax distribution for a continuous, differentiable approximation, proposed by Jang, Gu, and Poole (2016). In symbols,

$$\{\mathbf{E}_{\text{BP}}^*, \mathbf{V}^*\} = \frac{e^{(\boldsymbol{\ell}_\theta, k + \mathbf{g}_k^*)/\tau}}{\sum_{j=1}^K e^{(\boldsymbol{\ell}_\theta, j + \mathbf{g}_j^*)/\tau}} \quad (5)$$

where $K = 2$ for binary \mathbf{E}_{BP} and $K \geq 2$ for multiclass \mathbf{V} . τ is our scalar temperature input upon checking the shape of prior distribution across K classes, and \mathbf{g}^* is sampled as:

$$\mathbf{g}^* = -\log(-\log(\mathbf{u}^*)), \quad \mathbf{u}^* \sim \text{Uniform}(0, 1) \quad (6)$$

In addition, as introduced in Section 3.1, $\boldsymbol{\ell}_\theta$ are logits, which are direct outputs of our encoder networks. To determine the corresponding probabilistic parameter \mathbf{p}_θ for the k^{th} class or building typology for our Bernoulli or multinomial probabilistic distribution, respectively, we can also use the softmax operator but without the stochastic sampling part and the scalar temperature input, as:

$$\mathbf{p}_\theta^k = \frac{e^{\boldsymbol{\ell}_\theta, k}}{\sum_{j=1}^K e^{\boldsymbol{\ell}_\theta, j}} \quad (7)$$

Local Pruning for K Vulnerability Types Following a similar strategy for efficient training from multi-hazard impact studies (Xu et al. 2022a), the first of our two ways of incorporating prior knowledge information is through local pruning of our multinomial probabilistic distribution. For example, even though a region can have K vulnerability types, some areas can have a varying number of types. To illustrate, a more rural or more urbanized area might have only fewer types, while the rural-urban interface area can have more diverse types up to K types. In our experiment, the logit is simply set to a small value (i.e., -1000) if the k^{th} building typology is known a priori to be non-existent.

Loss Functions We train every module using the sum of reconstruction loss, \mathcal{L}^{rec} , and Kullback-Leibler divergence loss, \mathcal{L}^{KL} . The training of all decoder networks, g , is more straightforward than that of our encoder networks, f .

On one hand, our decoder networks minimize \mathcal{L}^{rec} in terms of a weighted mean-squared error.

$$\mathcal{L}^{\text{rec}} = \frac{1}{N} \sum_{i=1}^N w_i^x (x_i - \hat{x}_i)^2 \quad (8)$$

where, for N data points, $x \in \mathbf{X}$ and $\hat{x} \in \hat{\mathbf{X}}$. w^x is a scalar shape-dependent weight with values depending on their Euclidean distances from the defining characteristic points of the empirical distribution of \mathbf{X} . For instance, since \mathbf{X} seems to follow a normal distribution upon investigation, we assign higher weights to points located at the modal peak and at both extreme ends; otherwise, the weights for other points are linearly interpolated.

On the other hand, the encoder networks minimize \mathcal{L}^{KL} and use a similar weight assignment approach. For \mathbf{E}_{BH} , we take the average in such a way that, for i^{th} point,

$$\mathcal{L}_{\text{BH},i}^{\text{KL}} = \frac{1}{2} \left(\frac{(\mu_{\theta,i} - \mu_{0,i})^2}{\sigma_{0,i}^2} w_i^{\mu_0} + \left(\frac{\sigma_{\theta,i}^2}{\sigma_{0,i}^2} - \ln \left(\frac{\sigma_{\theta,i}^2}{\sigma_{0,i}^2} \right) - 1 \right) w_i^{\sigma_0} \right) \quad (9)$$

where $\mu_{\theta,i} \in \boldsymbol{\mu}_{\theta}$ and $\sigma_{\theta,i} \in \boldsymbol{\sigma}_{\theta}$. Similarly, μ_0 , σ_0 , w^{μ_0} , and w^{σ_0} are from the prior distribution of building height.

For \mathbf{E}_{BP} ,

$$\mathcal{L}_{\text{BP},i}^{\text{KL}} = \left(p_{\theta,i} \log \frac{p_{\theta,i}}{p_{0,i}} + (1-p_{\theta,i}) \log \frac{1-p_{\theta,i}}{1-p_{0,i}} \right) w_i^{p_0} \quad (10)$$

where $p_{\theta,i} \in \boldsymbol{p}_{\theta}$. Similarly, p_0 and w^{p_0} are from the prior distribution of building presence.

For \mathbf{V} where i^{th} point has K possible types,

$$\mathcal{L}_{\mathbf{V},i}^{\text{KL}} = \left(\sum_{k=1}^K p_{\theta,i,k} \log \frac{p_{\theta,i,k}}{p_{0,i,k}} \right) w_i^{p_0} \quad (11)$$

where $p_{\theta,i} \in \boldsymbol{p}_{\theta}$. p_0 and w^{p_0} also come from the prior distribution of vulnerability types. Specifically, w^{p_0} assigns higher importance to the most likely class via masking.

In \mathbf{V} , our initial findings suggest that the effects of weak supervision combined with many possible classes posed difficulties in learning diverse classifications. Hence, we added a supervised cross-entropy loss from a semi-supervised variational learning solution (Kingma et al. 2014).

$$\mathcal{L}_{\mathbf{V},i}^{\text{CE}} = \left(\sum_{k=1}^K p_{0,i,k} \log(p_{\theta,i,k}) \right) w_i^{p_0} \quad (12)$$

3.4 Spatiotemporal Graph

Whether at city or country scale, we divide the region into multiple square tiles and split them into training, testing, and validation sets with a balanced number of \mathbf{V} classes. Using

the following spatiotemporal graph representations, we train **OE**, **TE**, **OV**, and **TV** modules with graph convolutional networks (Kipf and Welling 2016).

For each tile, we create an undirected exposure graph $G_t^E = (N_t^E, A_t^E, X_t^E)$ at time step t . N_t^E is the set of nodes that represent all pixels in the square tile, A_t^E is the grid-based adjacency matrix or connectivity information between these nodes from all eight directions, and X_t^E is the feature covariates such as the signals and derived indices from satellite imagery. Through time, while the A_t^E remains unchanged, the values of X_t^E may vary to indicate changes in building presence and height.

To train the **OV** and **TV** modules, we then reduce the A_t^E if the sampled values of $\mathbf{E}_{\text{BP}}^{\text{TE}}$ exceeded 0.5. In this case, the A_t^E becomes A_t^V , which identifies N_t^V and also changes through time. Our X_t^V contains the sampled values of $\mathbf{E}_{\text{BH}}^{\text{TE}}$ and, as mentioned previously, the previously g^{TE} -reconstructed covariates, thereby creating the undirected vulnerability graph G_t^V .

3.5 Evaluation

To facilitate a comparable evaluation with previous studies on dynamic vulnerability modelling (Pittore, Haas, and Silva 2020), we calculate the Aitchison distance, d_A , which quantifies the difference between the compositions of two different models (i.e., prior \boldsymbol{p}_0 and ours \boldsymbol{p}_{θ}) (Aitchison 1982).

$$d_A = \sqrt{\frac{1}{2K} \sum_{i=1}^K \sum_{j=1}^K \left[\ln \left(\frac{\boldsymbol{p}_{0,i}}{\boldsymbol{p}_{0,j}} \right) - \ln \left(\frac{\boldsymbol{p}_{\theta,i}}{\boldsymbol{p}_{\theta,j}} \right) \right]^2} \quad (13)$$

4 Experimental Setup

We demonstrate the application of **GRAPHVSSM** through three distinct case studies, each with varying spatiotemporal resolutions. First, we implement full sequential training of all modules for Quezon City, Philippines, at an annual 10-meter resolution, conducted in close collaboration with the disaster risk managers. Next, to account for significant changes caused by recent disasters, we showcase the benefits of modularity by training only the **OV** module for the cyclone-impacted coastal Khurushkul community (Bangladesh) and the mudslide-affected Freetown (Sierra Leone), both at an annual 50-centimeter resolution. Finally, to illustrate the scalability at a coarser continental level, we train the **OV** module for 46 UN-recognized Least Developed Countries as of 2020, using a five-year interval at a 90-meter resolution. Further details are in the Supplementary Material section of our extended preprint version. (Dimasaka, Geiß, and So 2025).

5 Results and Discussion

5.1 Changing Physical Vulnerability Composition

In Figure 2, we illustrate, through two different time scales, the changes in the composition of physical vulnerability using our trained **OV** module with time-series satellite-derived products of building exposure.

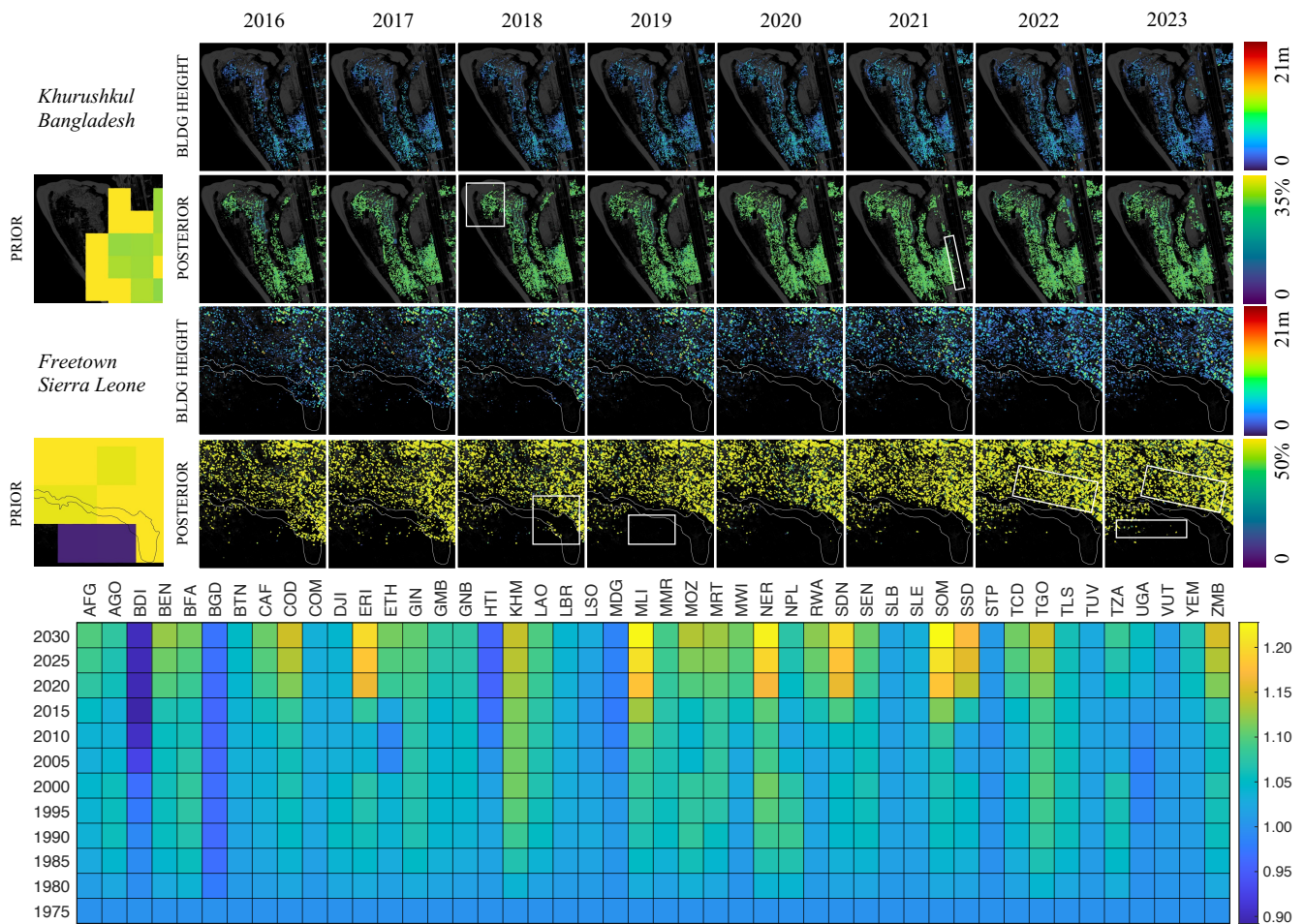


Figure 2: *Top*: Annual changes in building presence and height (Google Open Buildings 2.5D Temporal) and the corresponding METEOR prior and OV posterior probabilities for the informal settlement (INF) and unreinforced concrete block masonry (UCB) classes in cyclone-impacted Khurushkul (Bangladesh) and mudslide-affected Freetown (Sierra Leone), respectively. *Bottom*: Change in mean Aitchison distance between METEOR prior and OV posterior probabilities (from GHSL multitemporal products) of vulnerability classes across 46 UN Least Developed Countries (as of 2020), calculated as the ratio relative to the 1975 baseline (T/1975), at five-year intervals, 1975–2030.

In Khurushkul, Bangladesh, our findings show that prior to the disaster, our posterior maps for 2016 and 2017 indicated an increasing density of informal constructions around the coastal community. However, this pattern lessened in 2018 following the disaster and began to show signs of scattered rebuilding from 2019 onward, with a notable expansion near the northwest part of the map (i.e., close to the water body). The 2021 posterior map also captured the significant displacement caused by the construction of the airport. Since the building height and geometry of the new airport facilities have different characteristics from that of the surrounding dominant informal settlement patterns, the trained OV module recognized these differently, assigning lower probabilities of being classified as informal constructions. This result underscores the influence of local contextual information (i.e., how the houses are sized and irregularly arranged) as captured by the graph neural network, on

the probability of a particular building typology.

In Freetown, Sierra Leone, although the mudslide caused the affected area to be uninhabitable, our posterior maps revealed a post-disaster community-level behavior in which a densification of the same weak vulnerability class (unreinforced concrete block masonry) occurred around the perimeter and, in some instances, even within the previously devastated area. Since both cases show the persistence of highly vulnerable building types in areas with elevated climate and disaster risks, our evidence collectively suggest the critical importance of regularly monitoring the spatial and temporal distribution of physical vulnerability to better inform a holistic community-level disaster resilience planning.

On the bottom half of Figure 2, the change in mean Aitchison distance supports our prior expert belief systems regarding the compositions of vulnerability classes. Particularly, our resulting heatmap inflection points around 2000-

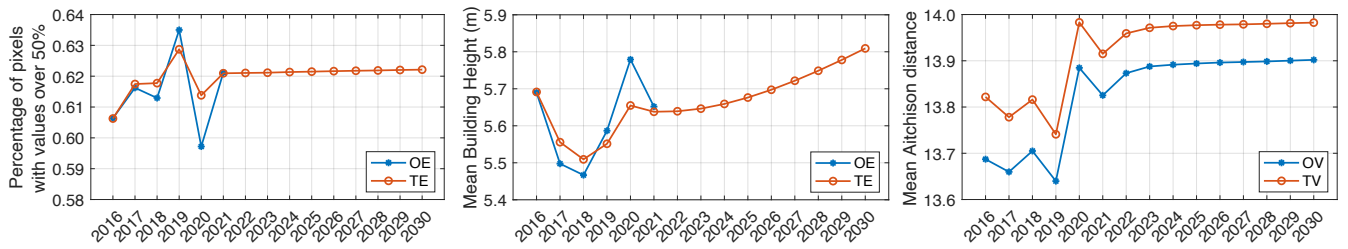


Figure 3: Annual changes in OE/TE posterior parameters for E_{BP} (left) and E_{BH} (middle), and the mean Aitchison distance between prior and OV/TV posterior probabilities of V classes (right) in Quezon City, Philippines.

2010 (i.e., shifting from decreasing to increasing) suggest that the METEOR data likely reflects vulnerability patterns during that decade. Between 1975 and 2000, the positive rate of change corresponds to the influence of a low built-up area characterized by a more homogeneous composition of vulnerability classes. After 2010, the rate of change accelerates beyond that of the 1975–2000 period, indicative of the anticipated urban expansion and substantial changes in the heterogeneous composition of physical vulnerability classes.

5.2 Uncovering Spatiotemporal Dynamics

Examining the derived posterior parameters across our four modules of the graph state-space model reveals insights into the regional spatiotemporal dynamics of building exposure and physical vulnerability. As shown in Figure 3, during the observable 2016–2021 period, the trends of posterior parameters for OE, TE, OV, and TV remained similar but began to diverge after 2021, implying the learned slow temporal behavior in building exposure and physical vulnerability.

5.3 Learning Interpretable Probabilistic Latents

Representing latent variables as parameters of assumed probability distributions has become feasible through the use of multiple explanatory covariates and derived indices from satellite imagery, as shown in Figure 4. Our results further demonstrate that incorporating prior knowledge into the Kullback-Leibler divergence and cross-entropy losses to address weak multi-class supervision allows the framework to balance supervised learning with deep Bayesian coarse-to-fine-grained updating. This yields interpretable posterior parameters and enables the reconstruction of covariates using reparameterization techniques for both continuous and categorical distributions.

6 Conclusion and Future Work

While machine learning and time-series satellite imagery are increasingly used to understand climate and disaster risks – especially for dynamic exposure and hazard mapping – our work tackles the equally important yet uniquely challenging aspect of modeling physical vulnerability at large spatiotemporal scales. Our contribution demonstrates a probabilistic deep learning approach, contributing to broader urban studies that require compositional data analysis in weakly supervised or coarse-to-fine-grained settings. For future study, we recommend the following next steps:

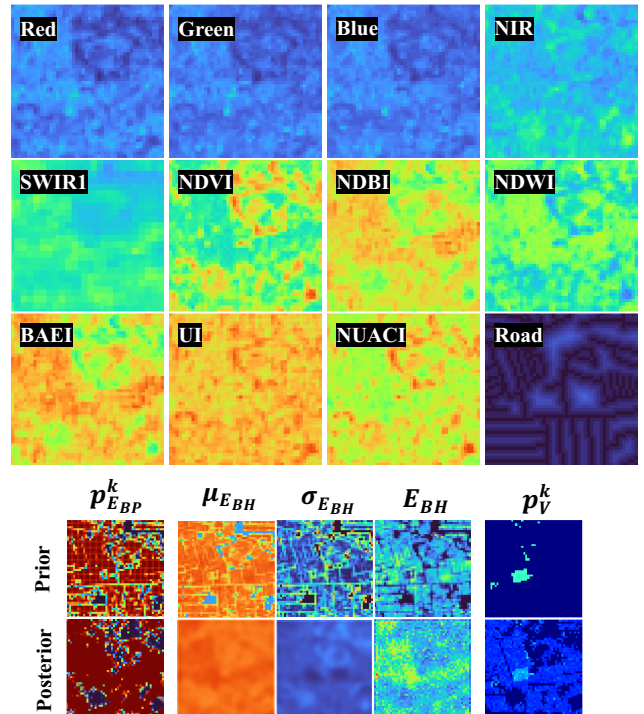


Figure 4: *Top twelve*: Explanatory covariates and derived indices from satellite imagery. *Bottom ten*: Comparisons between prior and OE/OV posterior parameters.

- In training the graph state-space model, while we understand that this may become computationally expensive, simultaneous, instead of sequential, learning can provide insight into a unified modelling framework.
- In vulnerability modelling, the use of vector-based representation of buildings can result in an interesting modification of the multinomial probability distribution that uses discrete counts, instead of rasterized proportions.
- Calibrating the TV module with monitored information on the changes of physical vulnerability (e.g., retrofit, demolition, or disaster damage) can improve the overall performance and flexibility with new observations.

Acknowledgments

This work is funded by the UKRI Centre for Doctoral Training in Application of Artificial Intelligence to the study of Environmental Risks (AI4ER) (EP/S022961/1). We are also thankful for the building exposure and physical vulnerability data from Quezon City Disaster Risk Reduction and Management Office (Ms. Maria Bianca Perez); the valuable practical guidance from Earthquakes and Megacities Initiative (Dr. Fouad Bendimerad and Dr. Renan Tanhueco); Charles Huyck for sharing technical limitations of the METEOR products; and Dr. Robert Muir-Wood for proposing the need for spatiotemporal disaster risk quantification auditing and technical suggestions on the relevance of case studies.

References

- ADB. 2017. *Climate Change Operational Framework, 2017-2030: Enhanced Actions for Low Greenhouse Gas Emissions and Climate-resilient Development*. Asian Development Bank.
- Agresti, A.; and Hitchcock, D. B. 2005. Bayesian inference for categorical data analysis. *Statistical Methods and Applications*, 14: 297–330.
- Aitchison, J. 1982. The statistical analysis of compositional data. *J. R. Stat. Soc. Ser. B (Methodol.)*, 44(2): 139–160.
- Almufti, I.; and Willford, M. 2013. The resilience-based earthquake design initiative (REDiTM) rating system. *Arup*.
- Benidis, K.; Rangapuram, S. S.; Flunkert, V.; Wang, Y.; Maddix, D.; Turkmen, C.; Gasthaus, J.; Bohlke-Schneider, M.; Salinas, D.; Stella, L.; et al. 2022. Deep learning for time series forecasting: Tutorial and literature survey. *ACM Computing Surveys*, 55(6): 1–36.
- Blei, D. M.; Kucukelbir, A.; and McAuliffe, J. D. 2017. Variational inference: A review for statisticians. *J. Am. Stat. Assoc.*, 112(518): 859–877.
- Bodnar, C.; Bruinsma, W. P.; Lucic, A.; Stanley, M.; Allen, A.; Brandstetter, J.; Garvan, P.; Riechert, M.; Weyn, J. A.; Dong, H.; et al. 2025. A foundation model for the Earth system. *Nature*, 641(8065): 1180–1187.
- Bogardi, J. J.; and Fekete, A. 2019. From intriguing concept (s) towards an overused buzzword: is it time for a requiem for resilience? *Resilience and Vulnerability: Conceptual*, 1996.
- Bronstein, M. M.; Bruna, J.; Cohen, T.; and Veličković, P. 2021. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv:2104.13478*.
- Bruneau, M.; Chang, S. E.; Eguchi, R. T.; Lee, G. C.; O'Rourke, T. D.; Reinhorn, A. M.; Shinozuka, M.; Tierney, K.; Wallace, W. A.; and Von Winterfeldt, D. 2003. A framework to quantitatively assess and enhance the seismic resilience of communities. *Earthq. Spectra*, 19(4): 733–752.
- Calderon, A.; and Silva, V. 2022. Forecasting seismic risk within the context of the Sendai framework: An application to the Dominican Republic. *Int. J. Disaster Risk Reduct.*, 82: 103364.
- Cini, A.; Marisca, I.; Zambon, D.; and Alippi, C. 2025. Graph Deep Learning for Time Series Forecasting. *ACM Comput. Surv.*
- Ciullo, A.; Strobl, E.; Meiler, S.; Martius, O.; and Bresch, D. N. 2023. Increasing countries' financial resilience through global catastrophe risk pooling. *Nat. Commun.*, 14(1): 922.
- Comerio, M. C. 2006. Estimating downtime in loss modeling. *Earthq. Spectra*, 22(2): 349–365.
- Cremen, G.; Galasso, C.; and McCloskey, J. 2022. Modelling and quantifying tomorrow's risks from natural hazards. *Science of The Total Environment*, 817: 152552.
- de Bézenac, E.; Rangapuram, S. S.; Benidis, K.; Bohlke-Schneider, M.; Kurlle, R.; Stella, L.; Hasson, H.; Gallinari, P.; and Januschowski, T. 2020. Normalizing kalman filters for multivariate time series analysis. *Adv. Neural Inf. Process. Syst.*, 33: 2995–3007.
- Dimasaka, J.; Geiß, C.; and So, E. 2025. GraphVSSM: Graph Variational State-Space Model for Probabilistic Spatiotemporal Inference of Dynamic Exposure and Vulnerability for Regional Disaster Resilience Assessment. *arXiv:2508.01310*.
- Dimasaka, J.; Geiß, C.; and So, E. 2025. DeepC4: Deep Conditional Census-Constrained Clustering for Large-scale Multitask Spatial Disaggregation of Urban Morphology. *arXiv:2507.22554*.
- Dimasaka, J.; Selvakumaran, S.; and Marinoni, A. 2024. Enhancing assessment of direct and indirect exposure of settlement-transportation systems to mass movements by intergraph representation learning. *Environ. Res. Lett.*, 19(11): 114055.
- Durbin, J.; and Koopman, S. J. 2012. *Time series analysis by state space methods*. Oxford University Press (UK).
- FEMA. 2012. *Seismic performance assessment of buildings*. Federal Emergency Management Agency.
- Fill, J.; Eichelbeck, M.; and Ebner, M. 2024. Predicting building types and functions at transnational scale. *arXiv:2409.09692*.
- Geiß, C.; Priesmeier, P.; Aravena Pelizari, P.; Soto Calderon, A. R.; Schoepfer, E.; Riedlinger, T.; Villar Vega, M.; Santa María, H.; Gómez Zapata, J. C.; Pittore, M.; et al. 2023. Benefits of global earth observation missions for disaggregation of exposure data and earthquake loss modeling: evidence from Santiago de Chile. *Natural Hazards*, 119(2): 779–804.
- Girin, L.; Leglaive, S.; Bie, X.; Diard, J.; Hueber, T.; and Alameda-Pineda, X. 2020. Dynamical variational autoencoders: A comprehensive review. *arXiv:2008.12595*.
- Gómez Zapata, J. C.; Zafirir, R.; Pittore, M.; and Merino, Y. 2022. Towards a sensitivity analysis in seismic risk with probabilistic building exposure models: an application in Valparaiso, Chile using ancillary open-source data and parametric ground motions. *ISPRS Int. J. Geo-Inf.*, 11(2): 113.
- Grover, A.; Kapoor, A.; and Horvitz, E. 2015. A deep hybrid model for weather forecasting. In *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 379–386.
- Huyck, C.; Hu, Z.; Amyx, P.; Esquivias, G.; Huyck, M.; and Eguchi, M. 2019. METEOR: exposure data classification, metadata population and confidence assessment. Report M3. 2/P. Technical Report M3. 2/P, British Geological Survey.

- Jang, E.; Gu, S.; and Poole, B. 2016. Categorical reparameterization with gumbel-softmax. *arXiv:1611.01144*.
- Khan, M.; Farid, Z. I.; Sojib, M. T. H.; Islam, M. A.; Gogon, M. I. R.; and Farid, S. I. 2024. How a Tropical Super-Cyclone Triggered an Intergenerational Vulnerability: A Case Study on a Cyclone Displaced Community in Cox's Bazar. Available at SSRN 4996026.
- Kingma, D. P.; Mohamed, S.; Jimenez Rezende, D.; and Welling, M. 2014. Semi-supervised learning with deep generative models. *Adv. Neural Inf. Process. Syst.*, 27.
- Kipf, T. N.; and Welling, M. 2016. Semi-supervised classification with graph convolutional networks. *arXiv:1609.02907*.
- Kong, B.; Ai, T.; Zou, X.; Yan, X.; and Yang, M. 2024. A graph-based neural network approach to integrate multi-source data for urban building function classification. *Comput. Environ. Urban Syst.*, 110: 102094.
- Lallemant, D. 2015. *Modeling the future disaster risk of cities to envision paths towards their future resilience*. Stanford University.
- Lallemant, D.; Burton, H.; Ceferino, L.; Bullock, Z.; and Kiremidjian, A. 2017. A framework and case study for earthquake vulnerability assessment of incrementally expanding buildings. *Earthq. Spectra*, 33(4): 1369–1384.
- Lam, R.; Sanchez-Gonzalez, A.; Willson, M.; Wirnsberger, P.; Fortunato, M.; Alet, F.; Ravuri, S.; Ewalds, T.; Eaton-Rosen, Z.; Hu, W.; et al. 2023. Learning skillful medium-range global weather forecasting. *Science*, 382(6677): 1416–1421.
- Lang, S.; Alexe, M.; Chantry, M.; Dramsch, J.; Pinault, F.; Raoult, B.; Clare, M. C.; Lessig, C.; Maier-Gerber, M.; Magnusson, L.; et al. 2024. AIFS–ECMWF's data-driven forecasting system. *arXiv:2406.01465*.
- Lei, B.; Liu, P.; Milojevic-Dupont, N.; and Biljecki, F. 2024. Predicting building characteristics at urban scale using graph neural networks and street-level context. *Comput. Environ. Urban Syst.*, 111: 102129.
- Lim, B.; and Zohren, S. 2021. Time-series forecasting with deep learning. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 379(2194): 1–14.
- Lin, J.; and Michailidis, G. 2024. Deep Learning-based Approaches for State Space Models: A Selective Review. *arXiv:2412.11211*.
- Marconcini, M.; Metz-Marconcini, A.; Esch, T.; and Gorelick, N. 2021. Understanding current trends in global urbanisation-the world settlement footprint suite. *GI Forum*, 9(1): 33–38.
- Martins, L.; and Silva, V. 2023. Global Vulnerability Model of the GEM Foundation. <https://doi.org/10.5281/zenodo.8391743>.
- Parker, D. J. 2020. Disaster resilience—a challenged science.
- Pesaresi, M.; Schiavina, M.; Politis, P.; Freire, S.; Krasnodebska, K.; Uhl, J. H.; Carioli, A.; Corbane, C.; Dijkstra, L.; Florio, P.; et al. 2024. Advances on the Global Human Settlement Layer by joint assessment of Earth Observation and population survey data. *Int. J. Digit. Earth*, 17(1): 2390454.
- Pittore, M.; Haas, M.; and Silva, V. 2020. Variable resolution probabilistic modeling of residential exposure and vulnerability for risk applications. *Earthq. Spectra*, 36(1_suppl): 321–344.
- Porter, K.; Hu, Z.; Huyck, C.; and Bevington, J. 2014. User guide: Field sampling strategies for estimating building inventories. *GEM Foundation*.
- Rangapuram, S. S.; Seeger, M. W.; Gasthaus, J.; Stella, L.; Wang, Y.; and Januschowski, T. 2018. Deep state space models for time series forecasting. *Adv. Neural Inf. Process. Syst.*, 31.
- Rolf, E.; Klemmer, K.; Robinson, C.; and Kerner, H. 2024. Mission Critical–Satellite Data is a Distinct Modality in Machine Learning. *arXiv:2402.01444*.
- Schorlemmer, D.; Beutin, T.; Cotton, F.; Garcia Ospina, N.; Hirata, N.; Ma, K.-F.; Nievas, C.; Prehn, K.; and Wyss, M. 2020. Global dynamic exposure and the OpenBuildingMap—a big-data and crowd-sourcing approach to exposure modeling. In *EGU Gen. Assem. Conf. Abstr.*, 18920.
- Sheikh, N.; and Saha, S. 2025. Graph neural networks for multi-sensor Earth observation. *Deep Learning for Multi-Sensor Earth Observation*, 211–230.
- Sirko, W.; Brempong, E. A.; Marcos, J. T.; Annkah, A.; Korme, A.; Hassen, M. A.; Sapkota, K.; Shekel, T.; Diack, A.; Nevo, S.; et al. 2023. High-resolution building and road detection from sentinel-2. *arXiv:2310.11622*.
- UNDRR. 2025. Global Assessment Report on Disaster Risk Reduction 2025: Resilience Pays: Financing and Investing for our Future. Technical report, United Nations Office for Disaster Risk Reduction, Geneva.
- Ward, P. J.; Blauhut, V.; Bloemendaal, N.; Daniell, J. E.; de Ruiter, M. C.; Duncan, M. J.; Emberson, R.; Jenkins, S. F.; Kirschbaum, D.; Kunz, M.; et al. 2020. Natural hazard risk assessments at the global scale. *Nat. Hazards Earth Syst. Sci.*, 20(4): 1069–1096.
- Xu, S.; Dimasaka, J.; Wald, D. J.; and Noh, H. Y. 2022a. Seismic multi-hazard and impact estimation via causal inference from satellite imagery. *Nat. Commun.*, 13(1): 7793.
- Xu, Y.; He, Z.; Xie, X.; Xie, Z.; Luo, J.; and Xie, H. 2022b. Building function classification in Nanjing, China, using deep learning. *Transactions in GIS*, 26(5): 2145–2165.
- Yepes-Estrada, C.; Calderon, A.; Costa, C.; Crowley, H.; Dabbeek, J.; Hoyos, M. C.; Martins, L.; Paul, N.; Rao, A.; and Silva, V. 2023. Global building exposure model for earthquake risk assessment. *Earthq. Spectra*, 39(4): 2212–2235.
- Zambon, D.; Cini, A.; Livi, L.; and Alippi, C. 2023. Graph state-space models. *arXiv:2301.01741*.
- Zhao, S.; Chen, Z.; Xiong, Z.; Shi, Y.; Saha, S.; and Zhu, X. X. 2024. Beyond Grid Data: Exploring graph neural networks for Earth observation. *IEEE Geosci. Remote Sens. Mag.*