Reliability Engineering and System Safety 272 (2026) 112575

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Analysing power system cascading failures and service disruptions in a data-scarce environment: A case study of Vietnam

Mengqi Ye [a,*], Surender Raj Vanniya Perumal [a], Miloš Pantoš [b], Sebastijan Ribnikar Cimerman [b], Philip J Ward [a,c], Elco E Koks [a]

[a] Institute for Environmental Studies, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, the Netherlands
[b] University of Ljubljana, Faculty of Electrical Engineering, Tržaška 25, 1000 Ljubljana, Slovenia
[c] Deltares, 2629 HV Delft, the Netherlands

ARTICLE INFO

Keywords:
Cascading failures
Service disruptions
Network analysis
Power flow model
Tropical cyclones

ABSTRACT

Cascading failures in power grids can escalate localised disruptions into extensive service outages. This study presents a transparent, reproducible framework for constructing large-scale power grid models from publicly available data. To assess service impacts on industrial users, we combine industrial points of interest with OpenStreetMap industrial zones and link them to substations using a capacity- and distance-informed allocation method. We examine two disruption scenarios: 1) random failures using percolation analysis and 2) hazard-driven failures using Monte Carlo sampling of wind-induced fragility curves derived from the 2024 Typhoon Yagi wind field. Results show that the Vietnam high-voltage grid is structurally vulnerable, with the giant component fragmenting rapidly under low levels of random bus removal, indicating limited redundancy. Although more than 95 % of transmission lines show no overload under random disruptions, critical corridors around major urban centres exhibit high overload probabilities exceeding 0.5. Under Typhoon Yagi, both direct and indirect failures contribute substantially to disruptions, with indirect impacts extending beyond wind exposed areas. Despite an average served load ratio of 0.91, 26 % of loads lose service in at least one simulation, revealing pronounced spatial disparities in resilience. These findings provide actionable insights for risk-informed resilience planning and targeted reinforcement investment.

1. Introduction

Power systems are critical infrastructure systems (CIS) that underpin modern society by enabling the delivery of essential services [1,2]. However, natural hazards can cause extensive damage to power grids, resulting in power outages for households and significant interruptions to essential services [3,4]. For example, in September 2024, the strongest typhoon over the past 70 years in Vietnam, Yagi, struck with wind speeds exceeding 213 km/h, affecting 26 provinces. The typhoon and its subsequent floods resulted in 320 deaths and overall damage around 1491 million USD, including 33 million USD damage in the electricity sector [5]. Similar serve consequences have been observed elsewhere. Hurricane Harvey in 2017 resulted in great power outages in Texas lasting over two weeks [6]. Hurricane Maria in 2017 caused severe damage to more than half of Puerto Rico's transmission towers and 50000 distribution poles, resulting in a near-complete blackout across the island [7]. Considering the fundamental role that power grids have in delivering services that are essential for household welfare, the functioning of society, and economic prosperity, it is crucial to take proactive measures to develop resilient infrastructure, which is often cost-effective in the long term [1,8,9].

In a changing environment, power systems are facing greater risks from hazardous events that are rare but can have severe impacts [10]. Failures in power systems rarely remain isolated technical incidents, instead, they can often propagate across interdependent networks, substantially amplifying the number of end users who experience service disruptions [11,12]. This cascading effect complicates post-hazard response and recovery and can result in large-scale economic losses. Understanding how initial failures evolve into cascading failures and service disruptions is therefore essential for risk assessment and resilience planning.

Interdependencies among components of the power grid, and between the power grid and other CISs, create complex and dynamic relationships. These interdependencies indicate that disruptions in one

* Corresponding author at: Room 8A-40, NU-building, Institute for Environmental Studies (IVM), De Boelelaan 1111, 1081 HV Amsterdam, the Netherlands.
E-mail address: mengqi.ye@vu.nl (M. Ye).

https://doi.org/10.1016/j.ress.2026.112575
Received 9 September 2025; Received in revised form 16 January 2026; Accepted 8 March 2026
Available online 9 March 2026

system can affect others, making it important to understand these links to identify vulnerabilities and potential intervention points. The interdependent nature can also intensify the impacts of disturbances through cascading failures. For example, a natural hazard event can disrupt other CISs that rely on electricity [4] and cause broad business interruptions [13]. To strengthen the resilience of power grids, it is important for stakeholders to create effective mitigation measures and manage both the interdependencies among systems and the potential impacts of failures. Understanding the interdependencies of these systems and assessing how they may influence the performance of power grids are essential steps in this process, often addressed using network modelling approaches.

Network modelling approaches generally fall into two categories: topological and serviceability methods [14]. Topological methods, which analyse network structure and connectivity using graph-theoretic measures, demonstrate that interdependencies introduce additional vulnerabilities to the network (e.g., [15–17]). Percolation analysis is widely used to examine the structure, functionality, and resilience of network systems [16–19]. In this method, system failures are simulated by gradually removing nodes or links from the network to evaluate how the system responds to disruptions. By employing different removal strategies, such as random, local, or targeted disruptions, this approach enables the exploration of various topological characteristics and system vulnerabilities, but it abstracts away operational constraints and therefore struggles to represent the real-world.

However, in real CISs, phenomena such as congestion in transportation networks and overloads in electricity transmission lines, may lead to gradual and continuous degradation in flow efficiency along the affected link instead of abrupt and complete failures [20]. Prior studies have demonstrated that more realistic vulnerability and resilience assessments can be achieved using serviceability methods. These approaches provide greater accuracy by evaluating how component failures affect network operations and service provision [14,21,22]. They incorporate the flow of passengers, goods, power, or information, while accounting for the capacity and other operational characteristics of different network components. Going beyond simple connectivity, they capture additional dynamics such as capacity loss, re-routing, and congestion. These methods offer a more realistic representation of infrastructure behaviour under stress, but typically require detailed data on network topology and operational information.

Appropriate network architecture is a critical and necessary foundation for the simulation and analysis of power grids [23], including the network topology and its service provision. The topological representation of existing energy infrastructure, such as the electricity grid, can have a deciding impact on future investments derived from energy system models [24]. A major barrier to applying the serviceability-based approach is the limited availability of open-source, high-quality power grid data, particularly in data-scarce contexts. While Transmission System Operators (TSOs) have detailed information on the high-voltage grid, these data are often not publicly available to the level of detail needed for academic research purposes [23,24]. Open-access data sources, such as OpenStreetMap (OSM), have therefore been used as an important alternative for constructing power grid representations in many existing studies [23–29], and have been validated with official data sources [24]. Although these data enable transparent and reproducible modelling, they often contain incomplete attributes, and lack of topology information.

Meanwhile, relatively few studies have examined the consequences of cascading failures on service provision, with the exception of [25], especially in the Global South context. In these regions, limited data availability and a lack of empirical evidence on climate-related power outages pose significant challenges for model development and validation. This gap is particularly important in regions with lower system reliability and high exposure to climate-related hazards, where service disruptions may exacerbate community inequities [30].

Regarding the service provision of the CISs, the conventional and widespread method used to allocate users to infrastructure assets is the Voronoi decomposition technique [13,23,31–33]. In this approach, service areas are defined as non-overlapping Voronoi cells, assuming that each user is equally served by the geographically closest infrastructure asset [13]. While the exact connections between infrastructure assets and end users are often unknown, it is reasonable to assume that demand centres are more likely to be connected to nearby assets. However, this method oversimplifies real-world dependencies, where actual service provision is shaped by proximity as well as by capacity constraints and operational decisions. For example, newly constructed facilities may be connected to a more distant substation if the nearest one lacks available capacity to accommodate additional demand. As a result, Voronoi-based assignments may misrepresent the true impact of infrastructure failures.

To address these gaps, this study develops a flow-constrained network modelling framework to analyse cascading failures and resulting service disruptions in power systems, particularly in data-scarce environments. We first develop a transparent framework for constructing a physically plausible network representation using open-source power grid data with empirical consistency constraints. Importantly, these constraints are introduced to regularise incomplete attributes in open-access data and to improve internal consistency. This data correction step addresses a key limitation in cascading failure studies, where incomplete network representations can substantially bias assessments of vulnerability and propagation dynamics.

Cascading failures are modelled through a two-stage process that includes failure initiation and failure propagation. By using the pandapower tool, we conduct power flow simulations to examine the network's response to two distinct types of initial disruptions: random disruption to simulate stochastic failures, and targeted disruption based on the impact of the 2024 Typhoon Yagi, representing a natural hazard-driven disruption to the network. This allows us to assess the grid's connectivity and robustness under both random disruptions and realistic extreme weather events, providing insights into its structural vulnerability and the potential for cascading failures that can lead to service interruptions.

To assess the consequences of cascading failures beyond the network level, the framework further links infrastructure assets to electricity demand (i.e., industrial areas) using a probabilistic, geospatial assignment that accounts for both spatial proximity and asset capacity. This approach improves on conventional proximity-based heuristics by enabling a more realistic representation of spatial dependency between infrastructure assets and end-users. By integrating this assignment with a power flow model, we are able to simulate cascading failures within the power system and trace their impact on downstream electricity provision. The proposed framework is demonstrated through a case study of the Vietnamese transmission network, proving its transparency and scalability for analysing power system risks in data-scarce environments. By integrating flow-constrained cascading failure modelling with a capacity-aware, geospatial representation of service dependencies, the framework offers a transferable and practically applicable approach for assessing service disruptions under both random and hazard-driven failures.

The remainder of this paper is structured as follows. Section 2 presents an integrated modelling framework that combines: (1) construction of power system model using open-source data; (2) geospatial modelling of dependencies between economic activities and infrastructure services, in order to quantify disruption to economic activities during infrastructure failures; and 3) cascading failures and service disruption analysis under random and targeted disruptions. In Section 3, we apply the framework to Vietnam as a case study to test the model and assess the vulnerability of power grids in Vietnam. Section 4 and 5 discuss the implications and limitations of the approach and conclude the paper.

2. Methods

We evaluate power system vulnerability using the framework illustrated in Fig. 1. First, a detailed network model is constructed from OSM data, supplemented with electricity generation and demand information from official statistics and Google Places. A power flow model is then built based on this network to simulate the operational behaviour under two disruption scenarios: random and targeted. In each scenario, component in-service statuses are updated accordingly, and the power flow analysis is run to evaluate system performance. Key performance metrics include component failure probability, line overloading probability, and the proportion of demand that remains served after initial disruption. By repeating the simulations 1000 times for each scenario, we assess how system performance degrades under different disruption patterns and loading conditions. The following sections describe each step of this framework in detail.

2.1. Network construction

2.1.1. Retrieving OSM power infrastructure data

OpenStreetMap is one of the largest open-access geospatial datasets of infrastructure assets. While it provides a valuable and consistent global database, it should be noted that the dataset is not necessarily fully complete, and there are gaps or inaccuracies in some regions [34–36]. Nonetheless, its open-access nature and global coverage make it a useful resource for preparing scalable and transferable models [26]. OSM has been widely used in energy system modelling [23,24,26,28,29,37] and risk analysis of various natural hazards to CIS [13,38,39]. Power grid data in OSM are represented by the data types of 'nodes', 'ways', and 'relations'. Nodes are points that are defined by their geographical coordinates. Ways are an ordered list of nodes, which define non-closed features (open ways, that is, transmission lines) or closed features (closed ways, that is, power plants and substations, which are rendered as polygons on map). Relations can contain nodes, ways, and a combination of either-or other relations. All mapped OSM data are associated with tags, which are dictionary-like entries having key and value text attributes [40].

In OSM data, the same asset can sometimes be represented either as a point (e.g., a node for a point of interest) or as its explicit shape (e.g., a closed-way or an area for a building), leading to duplicate entries. For example, a substation may be mapped as both a node and a way, each with a different OSM ID, despite referring to the same entity. To avoid duplicate representations while preserving meaningful sub-assets, we remove point features contained within polygons only when they represent the same asset, remove polygons nested within larger polygons if they refer to the same facility, and remove exact geometric duplicates using the OSM-flex tool [41,42].

2.1.2. Processing OSM data

First, we extract transmission lines at predefined voltage levels (i.e., 110 kV, 220 kV, and 500 kV), calculating line lengths adjusted by a slack factor (1.2 $\times$ geometric length) to approximate real-world routing [29]. Lines tagged with multiple voltages are cloned into distinct entries per voltage level (e.g., a line tagged with '220,000;500,000' is split into two lines with voltage tag of 220 kV and 500 kV, separately). To ensure unique line IDs, we append a suffix after the split (e.g., OSM ID 306271499 is split to 306271499a and 306271499b).

To eliminate redundant edges that do not contribute to power flow analysis, we remove lines tagged as busbars and lines that are entirely contained within substation boundaries. This process condenses substation internals (e.g., busbars) into single nodes, simplifying complex geometries while retaining topological information (Figure S1). Given the distinct electrical properties of direct current (DC) lines compared to alternating current (AC) lines, we identify potential DC lines based on specific criteria: lines with a cable count of 1, a frequency of 0 Hz, or keywords containing 'dc' are marked as DC cables.

In a three-phase AC high-voltage system, we assume that three cables constitute one AC circuit, six cables form two AC circuits, and so on [24,29]. For lines where circuit information is unavailable, we estimate the number of circuits by taking the floor of the total number of cables (if available) divided by three, and then further divided by the number of voltage levels. This approach estimates the number of circuits based on the assumption that each circuit typically consists of three cables. If the calculation results in zero circuits due to rounding down, or if cable data are missing and the number of circuits cannot be estimated, we assign a default minimum of one circuit per line. Otherwise, the circuit count is

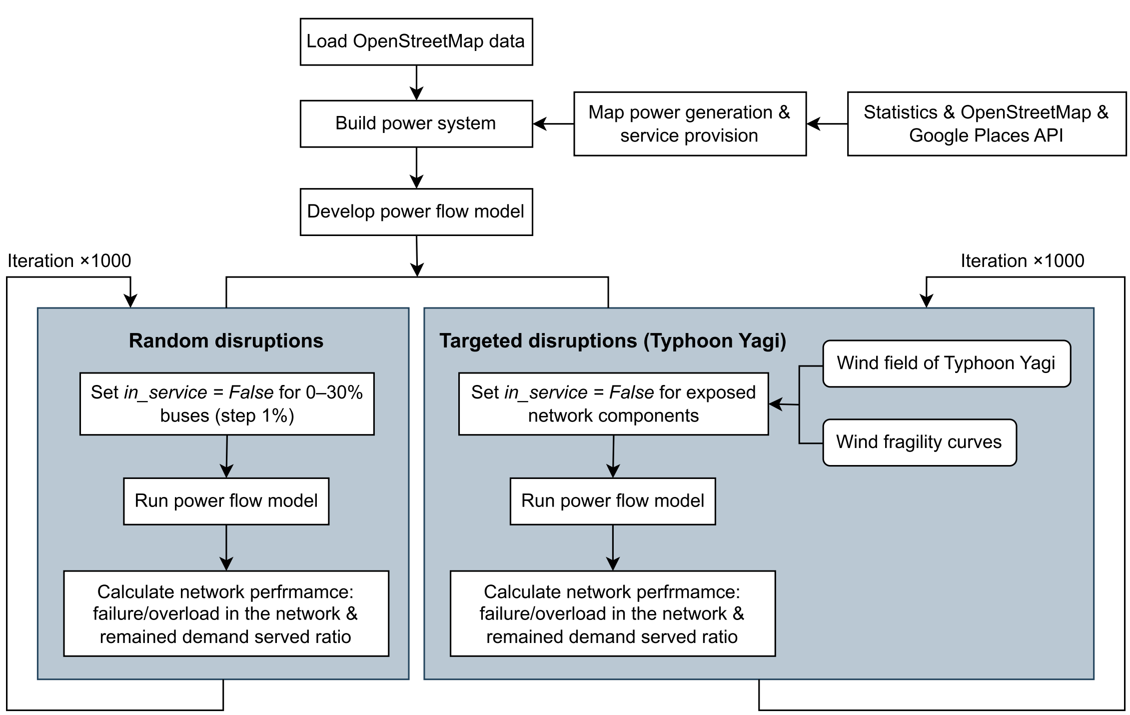

Fig. 1. Flowchart of the model procedure (adapted from [25]).

assigned the value of the calculated result. While this approach ensures that each line is associated with at least one circuit, it may sometimes underestimate the actual number of circuits, particularly in cases where the number of cables or voltage levels does not align perfectly with the assumptions of the model. Once the number of circuits is determined, each transmission line is assigned a unique identifier. Lines with a single circuit retain one record, while those with multiple circuits are duplicated, with alphabetical suffixes added to the base identifier. This cloning process allows each circuit to be modelled and analysed as a separate entity in subsequent simulations.

2.1.3. Constructing power network interdependencies

In this section, we explain how we establish topological connectivity among the grid components. To do so, we build a network graph, where buses/substations are represented as nodes and transmission lines as the connecting edges.

We start by extracting the end nodes of each line, referred to as ‘virtual buses’. Then, a 200 m buffer is created around each substation retrieved from OSM. To link virtual buses to physical substations, we implement a spatial attribution process. Each virtual bus that falls within the 200 m buffer zone of a substation inherits the substation’s information. The algorithm checks whether the virtual bus's coordinates are within any substation buffer using a spatial join. If multiple buffers overlap, the nearest substation is selected based on the minimum Euclidean distance to the virtual bus. Consequently, the line endpoints corresponding to the virtual bus are updated with the substation’s OSM ID, and the virtual bus itself is removed from the dataset.

Next, we apply a clustering algorithm based on physical distance between bus locations to aggregate spatially proximate buses, improving the topological representation of the power grid. This step is introduced as a spatial simplification procedure to reduce network complexity and computational costs arising from dense or overlapping node representations in original open-access data, rather than to represent electrical equivalence between aggregated buses. To achieve this, we first calculate the pairwise Euclidean distances between all buses and store the results in a distance matrix. This matrix allows to identify buses that can be grouped based on their proximity. Buses that are stacked on top of each other (i.e., with a distance of exactly 0) or fall within a predefined neighbouring threshold of 200 m are grouped accordingly [29]. The end nodes of lines are subsequently replaced by the aggregated buses, with their updated IDs and coordinates reflecting the new groupings. After grouping, lines that have identical start and end nodes are removed from the dataset.

To address missing voltage information, we assign the voltage levels of the connected buses to lines that lack this attribute. Any lines that still do not have voltage levels after this process are removed, as voltage level is a mandatory input parameter for a power flow model.

2.2. Mapping generators dependencies on power grids

To develop a power flow model, it is essential to know the power generation for each generator. However, the raw generator data extracted from OSM contain many missing or inconsistent entries, requiring additional data cleaning and completion. This process involves several steps to ensure consistency and accuracy in energy source classification and installed power capacity estimation. We first identify two types of generation facilities based on the key-value pair used to extract them from OSM, as summarized in Table S1. To create a unified generator dataset, we combine information on energy source and installed capacity from both power plants and generators. The identified energy sources in OSM data include hydro, coal, wind, biofuel, diesel, biogas, gas, gas;oil, oil, solar, and waste. These are harmonized and mapped to the fuel types used in the national statistical dataset[1], as detailed in Table S2.

The allocation of installed capacity is illustrated in the progression from Table S3 to Table S4. Updated or newly assigned values are shown in bold in Table S4, while entries removed due to insufficient information are also indicated. First, we categorize generators based on whether their installed capacity data are available. For those with recorded installed capacity, total and average values are calculated to evaluate the plausibility of the capacity distribution results. Next, we retained all generators for which an energy source is present, no matter whether installed capacity data are missing. Conversely, generators lacking an energy source are removed from the dataset, resulting in the elimination of 13 out of 1903 generators. Subsequently, the dataset is divided into two groups: one containing generators with recorded installed capacity (Group 1) and another containing generators without installed capacity (Group 2). The total installed capacity from Group 1 is subtracted from the statistical total capacity per energy source, and the remaining capacity is equally distributed among the generators in Group 2 based on their respective energy sources. The two groups are then recombined to form a complete dataset with estimated installed capacities.

Figure S2 shows the distribution of generators in Vietnam by energy source and their installed capacity as extracted from OSM data. In panel (A), the stacked bars represent the number of generators with and without installed capacity for each energy source. It indicates that a significant portion of wind turbines (901 out of 1233 units), solar panels (283 out of 328 units), and hydropower plants (126 out of 283 units) lack installed capacity data, highlighting gaps in publicly available OSM information. Panel (B) shows the total and average installed capacity by energy source.

To integrate the generators into the developed power grid developed in Section 2.1, we connected each generator to its nearest bus and assigned the corresponding bus ID. A distance threshold of 600 km was applied, allowing generators to be linked to any node within this range. This threshold was chosen to ensure that all generators could be connected to the network, as smaller thresholds would cause some generators to remain unlinked.

2.3. Mapping business dependencies on power grids

2.3.1. Extracting economic activities

Voltage requirements vary significantly across different end-users categories, reflecting diverse operational needs and infrastructure specifications. This study exclusively focuses on business consumers, particularly industrial entities. To execute our methodology, we require location data and electricity consumption information of economic activities.

To systematically map sites of economic activity, we integrated multiple data sources, including OSM land-use data and search results from Google Places API. Fig. 2 illustrates the multi-step process used to extract and spatially link industrial activities. In the first step (panel a), we query the Google Places API using a set of predefined search strings corresponding to specific industrial sub-sectors. This step produces a comprehensive set of locations potentially associated with industrial activity. The list of sector-specific search strings is provided in Table S5. To retrieve data from the Places API, we specified four key inputs: (1) a shapefile defining the study area boundary, (2) search strings tailored to the identified economic sectors in the study area, (3) clustering parameters, including land-use classifications and search radius constraints, and (4) an API key for accessing the Places API. The process fetches information on places that match the search strings within the specified search radius, including their locations. The clustering process is essential because the Places API returns a maximum of 60 results per

[1] Source of the statistical electricity generation data: https://ember-energy.org/data/electricity-data-explorer/

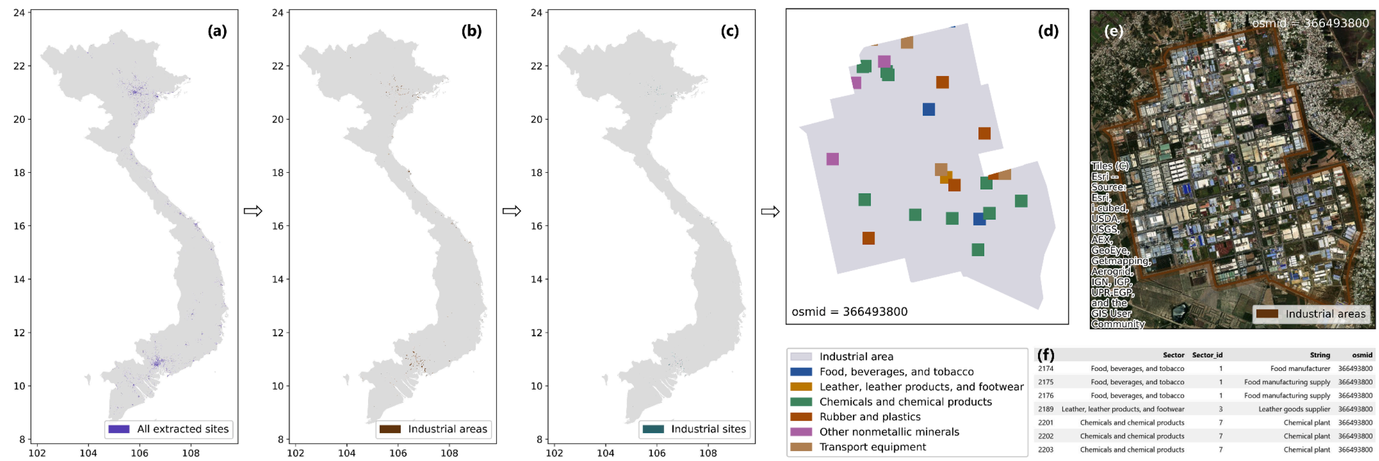

Fig. 2. Process of extracting economic activities.

query. To ensure comprehensive coverage of relevant sites across the entire study area, we select a mean search radius of 15 km to balance spatial resolution and query efficiency.

In the next step (panel b), we extract industrial areas from OSM, as industry is considered a primary electricity consumer in this study. These areas serve to spatially filter economic activities from Google Places API. Focusing on these areas helps to avoid capturing small-scale consultancies or suppliers operating in residential or commercial areas, who may self-identify as manufacturers in online listings but are not significant electricity consumers.

In the third step (panel c), we filtered extracted locations from Places API to retain only those that are operational and fall within the previously identified industrial areas. Panel (d), (e), and (f) in Fig. 2 present a representative example of this process for one industrial area (OSM ID: 366493800). Panel (d) shows the locations of sector-specific locations within the industrial area, and panel (e) presents the corresponding boundary overlaid on satellite imagery to confirm spatial accuracy. Panel (f) provides sample metadata for a subset of these locations, including economic sector classification, search string category, and OSM identifiers. It should be noted that these panels are intended for illustrative purposes and do not represent the full set of identified locations within the area.

Once economic activities are identified, we allocate sector-level electricity demand across all activities. Vietnam Energy Statistics [43] provides annual electricity sales by economic sector in gigawatt hour (GWh), representing the total energy consumed over the entire year. To meet the requirements of power flow modelling, it is necessary to convert this electricity consumption in GWh into power demand in megawatt (MW). Assuming the energy consumption of economic sectors is evenly distributed throughout the year, the conversion process involves: 1) converting the annual consumption from GWh to MWh by multiplying by 1000; and 2) dividing the result by the total number of hours in a year (8760 h).

The allocation process is structured as follows: 1) The total power demand is assigned to all identified economic activities based on sector classification. Each economic activity is assigned an equal share of the sectoral power demand; and 2) Economic activities are then grouped according to their corresponding industrial areas. Instead of treating each economic activity as a separate load, the total power demand within each industrial area is aggregated. By linking industrial sites to their corresponding industrial areas and then allocating power demand by economic sector among all sites, we obtain the necessary input data for loads in the power grid.

2.3.2. Developing dependency heuristics

In this section, we present the approach to establishing the dependencies between power grids and economic activities. We introduce a probabilistic-based approach for coupling substations and industrial areas, which incorporates both spatial proximity and substation capacity. In this method, factors such as the capacity of substations, which are challenging to include in the conventional Voronoi approach, can also be considered. For instance, larger substations with greater capacity are more likely to serve businesses, even if they are farther away than a smaller substation.

Specifically, we assume that the probability of an industrial area connecting to its k-th nearest substation follows a geometric distribution P_k, with a sparseness parameter p. This distribution is then weighed by the capacity (apparent power) of the substation C_k, with higher capacity substations given greater likelihood. The weighted probability w_k for an industrial area to connect to the k-th nearest substation is calculated by:

$$P_k = (1-p)^{k-1} \times p \tag{1}$$

$$w_k = \frac{C_k \times P_k}{\sum_{k=1}^{n} C_k \times P_k} \tag{2}$$

Here, p is a parameter between 0 and 1 that controls how likely it is for the industrial area to connect to substations. When p is closer to 1, the distribution becomes less sparse, meaning the probability of connecting to a nearby substation (lower values of k) is relatively high. As p increases, the likelihood of connecting to distant substations (higher values of k) decreases significantly. We set p to 0.8 in this analysis. Figure S3 illustrates an example of constructing this weighted probability distribution. The robustness of the assignment with respect to parameter p is evaluated through a targeted sensitivity analysis (see Supplementary Material Figure S4 and Note 1).

To determine the specific substation an industrial area connects to, we use a cumulative distribution function (CDF) derived from the normalized probabilities for each substation. After calculating the normalized probabilities for all substations, we generate a random seed between 0 and 1. The CDF, which represents the cumulative sum of normalized probabilities, is then used to map the random number to the corresponding substation. If the selected substation fails, as determined by the service status of the substation, the functional status of the industrial area is updated to reflect its disconnection. This probabilistic assignment ensures that the selection of the substation is not deterministic but follows the defined distribution based on distance, substation capacity, and the parameter p.

2.4. Power flow model development

To develop a power flow model, we make use of the pandapower, which is an open-source Python library designed to model, simulate, and analyse power systems [44]. It is particularly useful for performing power flow analysis and solving network optimization problems. Pandapower integrates well with various Python-based scientific tools, making it a versatile choice for power system modelling. A typical power flow model consists of several key components, such as buses, lines, generators, external grid, transformers, and loads, that represent the physical infrastructure of the power grid. Pandapower allows users to define these power grid elements, enabling efficient power flow simulations [44]. Below, we describe each component and how we utilize the simplified power system constructed in Sections 2.1–2.3 to create them.

In a power flow model, *buses* serve as connection points for different network components and represent the electrical nodes where voltage levels are controlled. Each bus corresponds to a specific location in the power grid and is assigned a voltage level. We used the power grid's node information to define each bus in the power flow model. The node locations from the grid are mapped to the buses, and corresponding voltage values are assigned to the buses.

Lines are modelled to represent the flow of electricity across the network. The power grid data provide geography information of transmission lines connecting corresponding buses at both ends. We extracted the connectivity information from the grid, creating the lines between buses. In pandapower, transmission lines are characterized by parameters. However, it is a common challenge in power system modelling where commercial sensitive specification data are often not openly-available. Pandapower provides predefined templates for common line types, we apply standard types to transmission lines based on corresponding voltage level in the absence of official transmission line parameters. Users can also create custom line types based on their data by specifying the line's characteristics, derived from empirical or manufacturer-specific data. The lengths of lines are assigned by the real lengths of transmission lines in the power grid, ensuring that the lines accurately represent the physical characteristics of the network.

Generators supply electrical power to the grid and are connected to specific buses in the power flow model. Each generator has specific parameters, including active and reactive power outputs and capacity. We define generators in the power flow model based on the generator data that include information on the location and generation capacity of each generator. The generators are mapped to the corresponding buses.

Loads represent the power demand in the grid, which are typically defined by their active power consumption. The economic activities data, including the power consumption and geolocations, are defined as loads in the power flow model. Each load is assigned to a specific bus based on the dependencies between economic activities and power grid as described in Section 2.3.2. This ensures that electricity demand is correctly placed within the power flow model, and the simulation could account for the impact of load variations on system performance.

Transformers are essential for voltage regulation and conversion between buses of different voltage levels. In many cases, detailed transformer data may be incomplete, missing, or unavailable. Therefore, we model hypothetical transformers between each pair of buses with different voltage levels using standard two-winding transformer type values provided by pandapower. The ceiling apparent power of transformers is defined as the maximum of the total rated apparent power of the connected lines on either side. The rated apparent power of each line is estimated based on its voltage level and thermal current as defined by standard line types [24]. This approach ensures the integrity of the power flow model while bypassing the need for accurate transformer data.

2.5. Cascading failures analysis under random and targeted disruptions

To evaluate the vulnerability and operational resilience of the developed power system under different disruption scenarios, cascading failures are analysed through a structured framework, including failure initiation, flow-constrained cascading failures, and service disruption assessment. This structure reflects the different mechanisms governing component damage, power redistribution within the network, and the resulting loss of service. The analysis considers two types of disruption scenarios: 1) random disruptions, representing background uncertainty in system vulnerability; and 2) targeted, hazard-induced disruptions based on Typhoon Yagi. In both scenarios, the cascading failures are evaluated using power flow calculations using pandapower.

2.5.1. Failure initiation: stochastic and hazard-driven disruptions

Failure initiation represents the initial loss of system components due to external disturbances. Two distinct initiation mechanisms are considered. First, random disruption scenarios are used to generate failure conditions by randomly removing a fraction of buses from the network. To model large-scale cascading failures, we simulate the random removal of nodes at increasing levels of failure. We define a set of removal fraction values (from 0 % to 30 % of nodes with 1 % at each step) and iterate over a predefined number of simulations (1000 per fraction). In each run, a random subset of buses is removed according to the specified removal fraction by setting their 'in service' status to 'False' to mimic disturbances.

Second, hazard-driven disruption scenarios are modelled using spatial wind fields from Typhoon Yagi (2024) in combination with component-specific fragility functions. More than one-third of Vietnam's electric transmission grids are located in forest areas, making them vulnerable to tree falls during storms [45]. Therefore, we consider Typhoon Yagi, which landed on Vietnam in 2024, as the initial disruption, and simulate the targeted removal of grid components that fail due to strong winds, as determined by corresponding fragility curves. In this case, initial failures are driven by hazard intensity rather than random selection. Transmission lines and buses are assigned local wind speed values based on their spatial exposure, and failure probabilities are computed using corresponding fragility functions. A Monte Carlo simulation is then used to probabilistically determine initial component failures.

The wind field of Yagi (Figure S5) is computed from the CLIMADA tropical cyclone module, an open-source natural hazard risk assessment platform [46]. The track data for Typhoon Yagi (ID 2024246N14125) are retrieved from the International Best Track Archive for Climate Stewardship (IBTrACS) project [47]. To obtain which elements are directly exposed to the wind field of Yagi, we overlay the wind field onto Vietnam's power grid. For point-based assets (e.g., buses), wind speed is sampled directly at the coordinate location. For linear or polygon assets (e.g., transmission lines, loads), zonal statistics are applied to determine the maximum wind speed within each geometry's spatial extent. The wind speed (v_i) assignment process is summarized as:

$$v_i = \begin{cases} v(x_i, y_i), \ \textit{if geometry } i \textit{ is a point} \\ max\ \{v(x,y)\}, \ \textit{if geometry } i \textit{ is a line/polygon over domain } A_i \end{cases} \tag{3}$$

Each asset is thus assigned with a local wind speed value for fragility assessment.

We model asset failures using wind fragility curves for power infrastructure collected by [40], which relate failure probability P_f as a function of wind speed. Assets are categorized into fragility types based on their functional class and voltage level. Failure probabilities are then computed by linearly interpolating the corresponding fragility curves. The fragility curves are shown in Supplementary Material Figure S6.

A Monte Carlo simulation with 1000 iterations is conducted to probabilistically model initial (direct) failures. In each iteration:

1) The wind speed value for each asset is used to compute its failure probability through the corresponding fragility curve.
2) A uniform random variable $U \sim u(0,1)$ is sampled.
3) If $U \leq P_f$, the asset is marked as failed .

The result is a binary failure matrix across iterations:

$$fail_{iterk}(i) = \begin{cases} 1, & \text{if asset } i \text{ failed at iteration } k \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

In both scenarios, assets identified as having failed during the initiation stage are set to “out of service” state and serve as the starting point for subsequent system response evaluation.

2.5.2. Flow-constrained cascading failures

Following initial failures, cascading failures are evaluated under the developed network topology. After each initiation event, power flow calculations are performed to assess whether the remaining network can sustain feasible operation, including network connectivity, line loading levels, and supply-demand balance. The power flow model automatically propagates initial failures to all components electrically dependent on the failed buses.

Cascading propagation is triggered only when operational constraint violations arise under the power flow model, such as transmission lines overloads. When such conditions occur, affected components are disconnected from service and the power flow is recalculated on the updated network. In this way, cascading failures are constrained by network structure and flow redistribution, ensuring physical consistency in the propagation process. This functionality is implemented by pandapower, which tracks power flow and component dependencies through its internal data structures. Only those simulation runs in which power flow converges are retained for further analysis.

2.5.3. Service disruption assessment

Service disruption is assessed by translating cascading failures within the power system into loss of electricity supply at the level of end users. Following the cascading failures analysis, the operational status of each load is evaluated based on its electrical connectivity and feasibility of supply under the post-damage network configuration. A load is considered unserved if it is electrically disconnected from operational generation sources. Conversely, a load is considered served if power flow calculations indicate that its demand can be satisfied under the updated network topology. This binary classification allows the direct mapping of technical system failures to service availability at the demand side.

2.5.4. Performance metrics

The key performance metrics assessed include the failure probability of lines and buses, the overload probability of transmission lines, and the served load ratio, which is the proportion of power demand successfully supplied after the cascading failure process (i.e., the complement of the Energy Not Supplied (ENS) index).

As part of the cascading failure simulations, the power flow rebalances following each element removal. From this, we compute the overload frequency for all remaining operational transmission lines across simulation runs to capture the structural likelihood of line loading. For every iteration, we determine whether each line that is still operational is overloaded. Then the overload frequency is calculated by dividing the number of times a line is overloaded by the total number of successful simulations.

To quantify the severity and spatial extent of service disruptions, we test the functionality in each province by running power flow to evaluate whether each load remains served and calculate two indicators: the served ratio and the failure ratio. The served ratio measures the fraction of total load power demand that remains supplied after cascading failures. It accounts for the magnitude of the demand and thus reflects the severity of supply shortfall in terms of power volume. The failure ratio is defined as the proportion of individual loads that are unserved, which captures the frequency of service interruptions across economic activities. These two indicators highlight both the severity and likelihood of economic activities’ disruptions. For instance, in a province with 10 loads and a total power demand of 500 MW, if 3 loads with total power demand of 200 MW fail, the failed load ratio is 0.3, while the served ratio is 0.6 (300 MW served demand / 500 MW total demand). Service disruption metrics are further aggregated at the provincial level to assess regional impacts on economic activities, enabling to evaluate the spatial heterogeneity in service disruption and to identify regions that are particularly exposed to cascading power outages.

3. Application: power system failures caused by random disruptions and Typhoon Yagi

3.1. Network representation

The representation of the Vietnam high-voltage power grid is

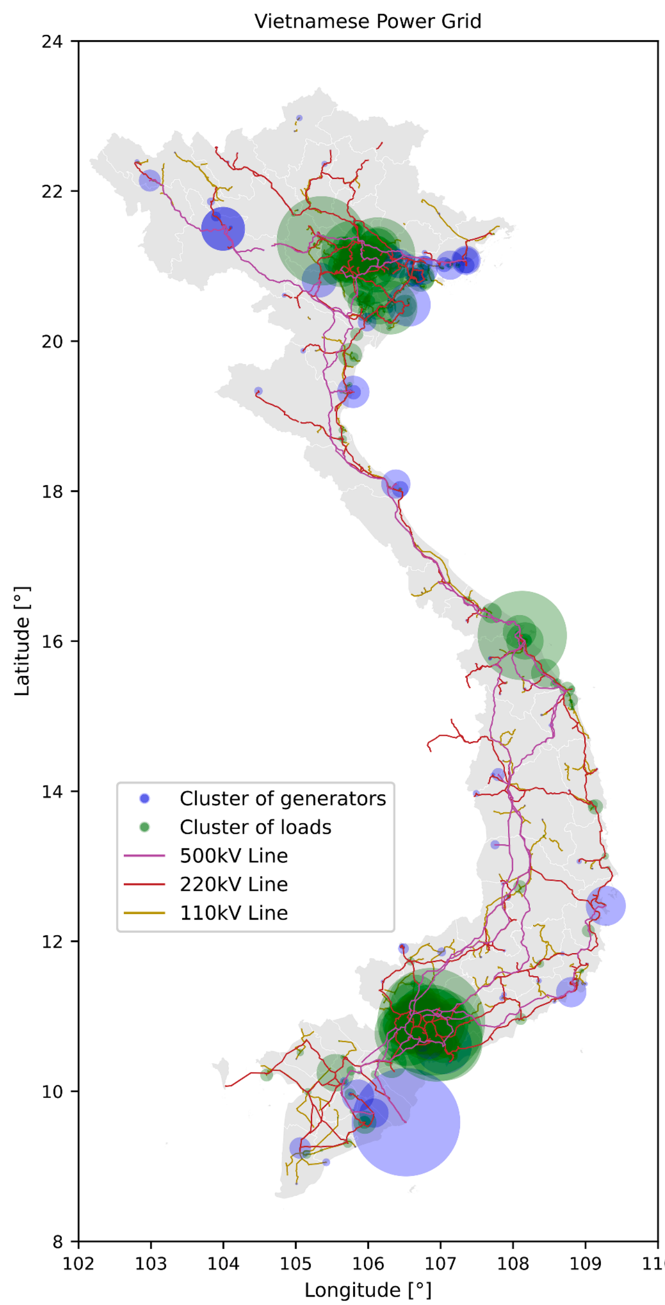

Fig. 3. Vietnam power grid and spatial distribution of load and generation clusters. The size of the cluster is proportional to the normalized total electricity demand or generation at corresponding bus.

constructed based on OSM data retrieved in May 2024. Fig. 3 illustrates the spatial layout of the compiled power system, including transmission lines at 500 kV, 220 kV, and 110 kV voltage levels, as well as clusters of electricity demand and generation. The processed network includes 3692 buses, compromising 818 substations extracted from OSM and 2874 virtual buses, 6799 AC lines, and 235 transformers. Overall, the network spans approximately 46,690 km in route length. All grid components are stored in CSV and GeoPackage files and can be accessed via GitHub repository (see Code and data availability). We also validate the processed data with country-level statistics from [5,48]. The comparison results, presented in Supplementary Material Table S6 and Note 2, indicate that the OSM data coverage of Vietnam's high-voltage power grid is extensive and potentially near-complete.

Fig. 4a presents the number of operational industrial sites retrieved from the Google Places API within OSM-defined industrial areas, categorized by economic subsector. In total, 2064 industrial areas are extracted from OSM, and 47449 sites are retrieved using sector-specific search strings via Google Places API. Among them, 7058 sites locate within the delineated industrial areas, and 6724 of these are operational and considered valid industrial sites for subsequent analysis.

To validate the spatial distribution of processed industrial businesses, we compare our results with the 2023 statistics from the National Statistics Office (NSO) of Vietnam,[2] as shown in Fig. 4b. Since the NSO categorises enterprises based on registered capital size, we focus on enterprises with capital above 10 billion Vietnamese dong (VND). Both datasets are normalised by their respective total number of businesses. Overall, the normalised proportions are generally aligned, with the highest concentrations consistently observed in the Southeast and Red River Delta regions. Importantly, the relative difference between processed results and statistics in these two economic hubs is within ± 3 %. Larger discrepancies in other regions may reflect data collection biases or regional disparities in data quality. These results support the use of Google Places data as a reasonable proxy for estimating the business activity in data-scarce contexts.

Figure S7 shows a spatial overview of industrial activities by economic subsector across Vietnam. The maps reveals a high concentration of industrial activities in the Southeast and Red River Delta regions, with sectoral variation across provinces. High electricity demand is mainly concentrated in these two economic centres. In 2015, these two regions accounted for 25 % and 46 % of Vietnam's gross industrial output, respectively [49]. Given this spatial pattern, network defragmentation is performed on the entire network, but the subsequent results focus primarily on these two regions that have larger exposure due to high demand concentration and network density.

3.2. The impacts of random disruptions

To evaluate the structural robustness of the power system under random disruptions, we analyse the size of the giant component (i.e., largest connected subnetwork) based on topological properties (Figure S8).This metric reflects the number of buses that remain mutually connected after a fraction of buses has been randomly removed. A decreasing giant component fraction indicates progressive fragmentation of the network. The rate of this decline provides insight into the overall connectivity structure: the sharp decrease, as observed in the first 0–5 % of bus removal, suggests that the network is structurally vulnerable, with limited redundancy and a strong dependence on a relatively small set of nodes to maintain global connectivity.

The spatial maps of failure probabilities of lines and nodes shown in Fig. 5 are generated from random percolation analysis with 1000 iterations. In each iteration, a random fraction of buses (from 0 % to 30 %, in 1 % steps) is removed, and the resulting cascading failures are simulated based on a power flow analysis that accounts for both network topology and redistribution. The failure probability of each line or bus is calculated as the percentage of iterations in which that specific component becomes non-operational (i.e., goes out of service due to direct removal or cascading effects). The Red River Delta and the Southeast region exhibit several high-risk corridors with elements in the top 5 % failure probability, likely due to dense interconnectivity and load centralization. These regions are critical for national grid stability, and the identification of high-probability failure links can inform targeted grid strengthening.

Fig. 5 also presents the spatial distribution of transmission line overload probabilities in the Red River Delta and the Southeast region. While the majority of lines in both regions have relatively low overload probabilities (i.e., typically below 0.1), indicating relatively low risk of overload under most disruption scenarios, several critical corridors show distinctly higher probabilities. These are particularly concentrated within the dense transmission network around Ha Noi and Ho Chi Minh City, where localized clusters of lines exhibit higher overload probabilities exceeding 0.5.

The histogram in Figure S9 illustrates the proportion of lines falling into different overload probability intervals. A majority of transmission lines (>95 %) have zero overload probability, indicating that they are not affected by flow redistribution and the network has slack capacity under a wide range of disturbance scenarios. Most lines have very low overload probabilities, with 2.84 % of lines falling into the 0 – 0.1 range. Only a small percentage of lines experience relatively higher overload risk. While this aggregated view provides a general picture of line vulnerability, further breakdown by removal fraction reveals how overload patterns evolve under increasing stress levels.

To evaluate the spatial disparity in load service resilience under failures, we analysed the served load ratio across provinces under a bus removal fraction of 5 %. Fig. 6 shows box plots of the served ratio across all iterations with random removal of 5 % buses. Provinces are sorted by the median served ratio. For each province, the median served ratio is shown along with the interquartile range represented by the box, and the 5th and 95th percentiles indicated by the whiskers. Additionally, a secondary Y-axis shows the number of loads per province. The results indicate that 42 % of provinces (24 out of 57) exhibit a high median served ratio above 0.85, although a wide range of variability remains observable. Provinces with the largest variability in served ratio (i.e., the most unstable) are highlighted in bold on the X-axis, with the top five being Da Nang city, Binh Phuoc, Bac Giang, Quang Nam, and Vinh Long (listed in Table S7). The provinces labelled in green are the most stable (low standard deviation in served ratio), suggesting relatively robust system performance under components removal. It should be noted that robustness does not necessarily imply a high served ratio. For example, Ha Giang ranks among the top five most stable provinces (Table S7), yet the corresponding median served ratio is relatively low at 0.82 compared to other provinces. In contrast, Quang Nam exhibits a high median served ratio at 0.87 but shows large variability across simulations, placing it among the most unstable provinces.

The averaged served ratio is then mapped to a geographic visualization in Figure S10, highlighting the spatial patterns of load resilience. Loads with higher served load ratios are considered more robust, while those with lower served ratios are highly vulnerable to the cascading failures of power grids.

To explore the relationship between topological integrity and operational functionality of the power grid under disruption, we analyse how the size of the giant component correlates with the served load ratio. Figure S11 plots this relationship across all iterations with various removal fractions, in which a non-linear upward trend is observed. As the giant component fraction increases, the served load ratio also increases, which confirms that larger network connectivity enables more effective redistribution of power flow. When the giant component fraction is below 0.4, there is a fast drop of served load ratio, indicating that once the grid is significantly fragmented, the network is insufficient

[2] Source: https://www.nso.gov.vn/en/px-web/?pxid=E0529&theme=Enterprise

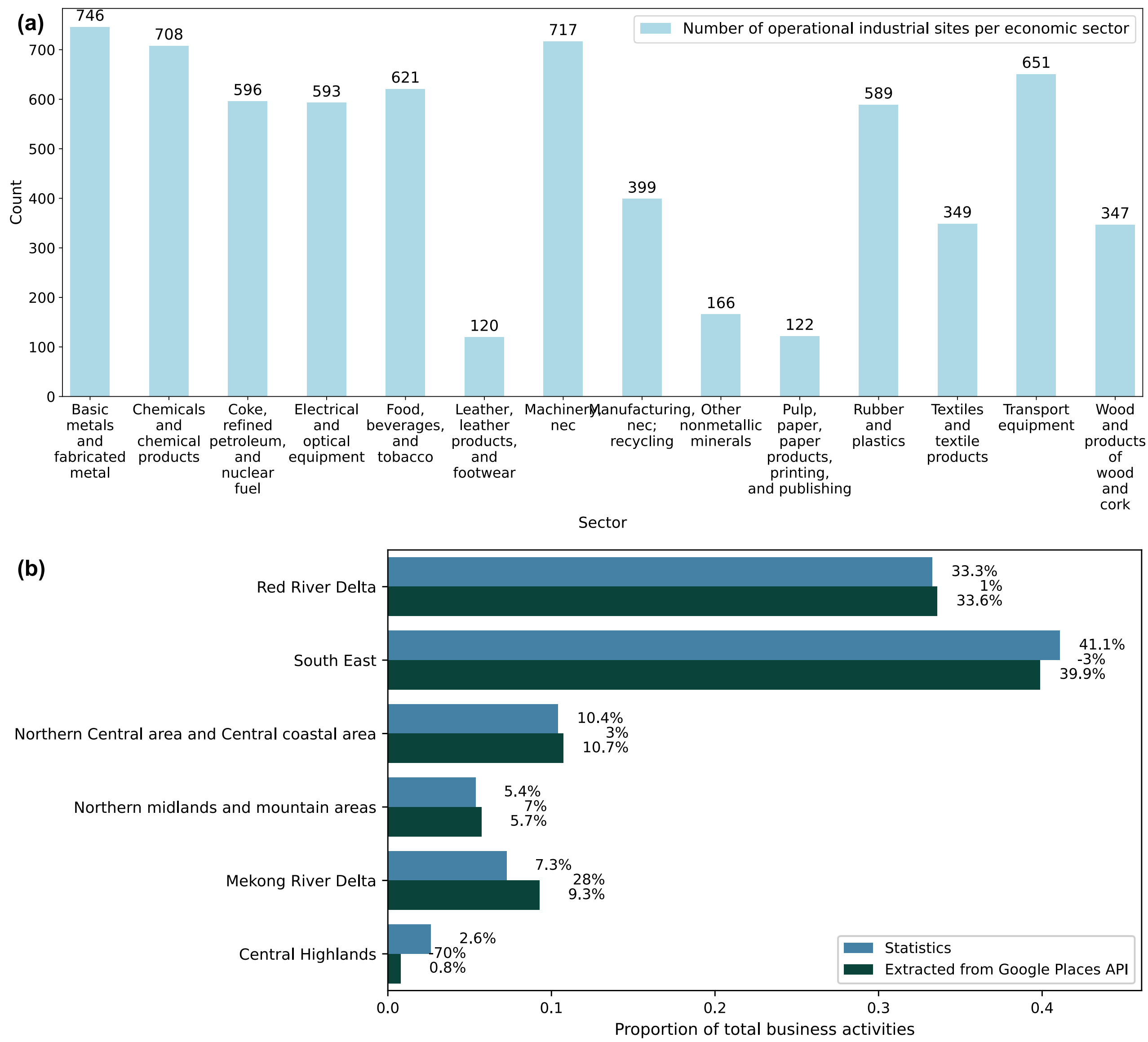

Fig. 4. Panel (a): Number of industrial sites per economic sector. Panel (b): Comparison of the proportion of business activities across Vietnam regions between statistics and data extracted from Google Places API. For each region, bars represent the share of businesses relative to the national total count, as reported in official statistics and estimated in this study. Values at the end of each bar indicate the relative deviation between the official statistics and estimations.

to support demand due to disconnection from the whole network. As the giant component fraction exceeds 0.6, the served load ratio begins to plateau near 0.9–1.0, suggesting that once a critical portion of the grid is preserved, additional improvements in topological connectivity yields limited marginal benefits in service delivery.

3.3. The impacts of Typhoon Yagi

This section presents how both physical damage (under scenario of historical Typhoon Yagi) and systemic vulnerabilities (cascading failures and overloads) contribute to widespread service disruptions, particularly in provinces located along the cyclone's trajectory in the northern region of Vietnam. Fig. 7 illustrates the spatial distribution and probability of direct and indirect failures of lines and buses under the strong wind scenario of Typhoon Yagi, aggregated across 1000 iterations. The top row shows the most frequent failure status observed in the simulations per line and bus, indicating whether each component is directly failed, indirectly failed, or remained functional in the majority of iterations. The middle and bottom rows present the corresponding probabilities of direct and indirect failures, respectively. High direct failure probabilities for both lines and buses are concentrated along the east-west-oriented transmission corridors in the northern part of the study area, aligning with the regions of highest wind exposure. Indirect failures are more spatially dispersed and represent cascading impacts that extend westward and southward beyond the high-wind zones, highlighting the role of network interdependencies in amplifying disruptions beyond the initial impact areas.

The overall failure impact is summarized in Table 1. Across the entire transmission network, a total of 2440.9 km of lines (5.2 %) experienced direct failures caused by strong winds of Typhoon Yagi, while an additional 1414.6 km (3.0 %) were affected indirectly. For buses, 248 (6.7 %) failed directly and 83 (2.3 %) failed indirectly. These results highlight the importance of both physical exposure and topological interdependencies in shaping the overall disruption patterns of power grid during extreme wind events.

Our method incorporates power flow dynamics, allowing to reflect the redistribution of electricity after initial failures, which is difficult to capture by traditional topological network modelling. Fig. 8 and

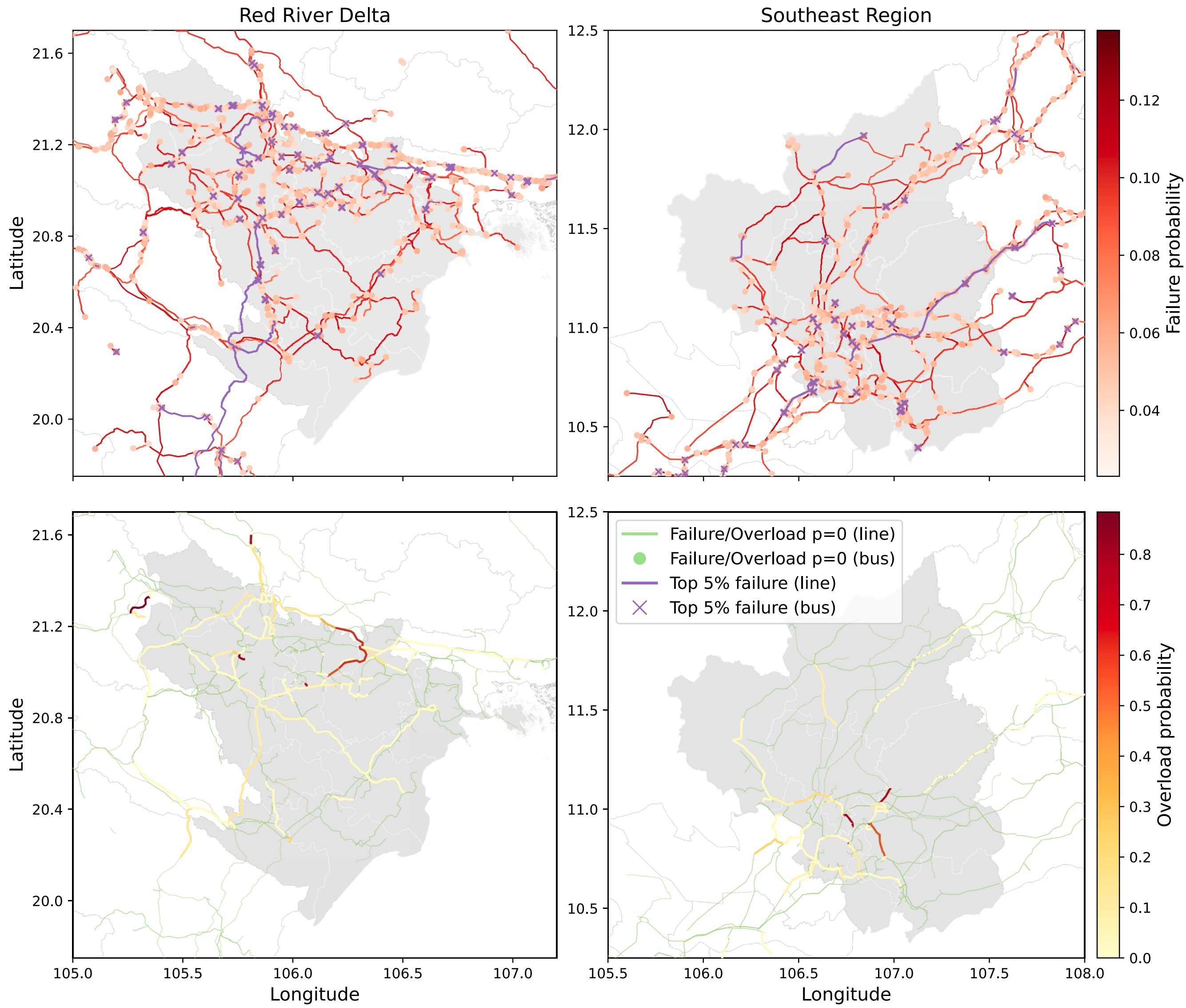

Fig. 5. Spatial distribution of buses and lines failure and overload probabilities in the Red River Delta and Southeast region under 5 % random bus removal simulations. Purple buses and lines in failure probability heatmaps are at top 5 % failure frequency.

Figure S12 present the spatial distribution and statistics of overload probability of transmission lines under the Typhoon Yagi scenario. Out of 6799 transmission lines assessed, 391 (5.75 %) exhibit non-zero overload probability across 1000 iterations. Most lines exhibit very low overload probabilities, with 2.54 % falling within the 0 – 0.1 range, while around 0.04 % lines experience relatively higher overload risk exceeding 0.5. Ten lines recorded overload probabilities exceeding 0.95, including seven with values of 1.0.

Fig. 9 and Table S8 present the distribution of load served ratio across 17 provinces affected by Typhoon Yagi. The overall average served ratio across all loads is approximately 0.91. Results reveal clear and considerable provincial variation. Thai Nguyen, Thanh Hoa, Ninh Binh, and Nam Dinh show the highest served ratio, each with a median of 1.0. In contrast, Vinh Phuc reports the lowest median served ratio at 0.036, followed by Bac Ninh at 0.365, while the remaining provinces maintain medians above 0.9. We also compute the standard deviation of load served ratios per province to assess the consistency of power supply. Provinces with high variability in served ratio are highlighted in bold on the X-axis. The three most unstable provinces are Hoa Binh, Quang Ninh, and Hung Yen, indicating large fluctuations in served ratios across iterations. Provinces labelled in green (i.e., Thanh Hoa, Ninh Binh, and Nam Dinh) show the lowest standard deviations, suggesting relatively robust and consistent system performance under component failures.

The averaged served ratio across all 1000 iterations is visualised in Figure S13, offering a detailed spatial representation of load-level resilience under the Typhoon Yagi scenario. The average served ratios across all loads is 0.91. Loads with higher served ratio are considered more resilient, as they consistently receive electricity despite disruptions. In contrast, lower served ratios indicate high vulnerability to cascading failures. Spatially, vulnerable loads are clustered in the northern part of the study area, particularly in provinces such as Vinh Phuc and Bac Ninh, where median served ratios are low and system performance is unstable (Table S7). This map highlights the heterogeneity in resilience of the power system, shaped by both network topology and exposure to initial disruption. The average load failure probability of each load across 1000 iterations under Typhoon Yagi scenario is presented in Figure S14. A total of 1378 loads were assessed, of which 358 (26.0 %) experienced service loss in at least one iteration. The mean failure probability among failed loads is 0.32, with a median of 0.10. In particular, 24 loads showed a loss-of-service probability of 1.0, corresponding to complete loss of service across all iterations.

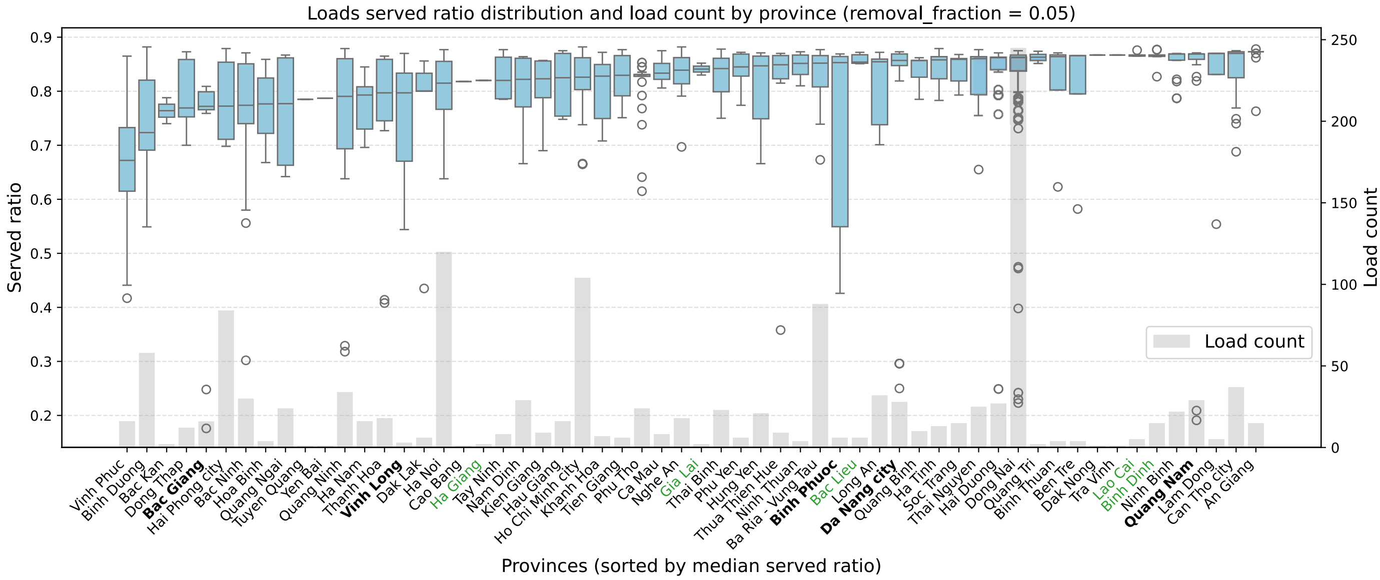

Fig. 6. Boxplot of served load ratio by province with removal of 5 % nodes. Each box represents the distribution of served ratios across loads in each province, sorted by median value. Province labels are styled based on the standard deviation of served ratios: bold indicates top five high standard deviation (i.e., least stable), while green indicates the top five provinces with the lowest standard deviations (i.e., most stable) among the affected provinces.

To ensure the adequacy of the Monte Carlo sample size, convergence of key indicators was evaluated by analysing running means and confidence intervals of the served load ratio and the number of overloaded lines. Results indicate that both indicators stabilise well before 1000 iterations (see Supplementary Material Figure S15 and Note 3).

4. Discussion

In this study, we contribute to the growing body of work that builds power system model using open-access data and open-source tools. This modelling approach enables transparency, reproducibility, and adaptability, which are particularly valuable in data-scarce environments. Commonly used tools in this field include pandapower [23,50], PyPSA [24], MATPOWER [25], etc. We adopted pandapower to construct a power system model that captures the structure and operation of the power system. The framework enables power flow simulations under different grid configurations and stress conditions, while also allowing the modelling of cascading failures. It allows to trace how failures propagate through the system to end users based on their interdependencies, and to quantify the extent of disruption under different scenarios. In addition, the model outputs include loading levels, which helps identify overload elements and potential capacity constraints. This is useful for analysing network stress and highlighting areas where capacity limits may be exceeded under certain conditions.

This study also proposes a novel probabilistic method for linking economic activities to power grid, which improves on spatial heuristics by incorporating both distance and substation capacity into the assignment process. The approach addresses the limitations of the traditional Voronoi decomposition, which rigidly assumes that all assets or population are served by their nearest substations, neglecting real-world complexities such as load capacity, infrastructure constraints, and institutional boundaries [12,13,23,51–53]. This improvement is useful for power flow simulations. By setting higher weights to substations with more power capacity, the model supports more realistic load distributions, which may improve the validity of power flow simulations.

This study simulates transmission network performance under strong wind during Typhoon Yagi and estimates resulting disruptions to electricity demand (i.e., economic activities). The model identifies 2441 km of direct line failure and 1415 km of indirect failure, along with 248 directly and 83 indirectly affected nodes, corresponding to 5.2 % and 3.0 % of the modelled transmission network length, and 6.7 % and 2.3 % of nodes, respectively (Table 1). According to post-disaster records from the Vietnam Multi-Sector Assessment Report [5], 14 and 40 line failures are reported at the 500 kV and 220 kV voltage levels, and 190 line incidents at 110 kV. However, this comparison must be interpreted with caution due to several limitations. First, the definition of failure differs: the model includes both physical wind-induced failures and cascading disconnections, while official reports document only physical asset damage. Second, due to network simplification in the modelling process, the model aggregates multiple physical lines, reporting failures by total length rather than line count, which complicates direct comparison. Third, the model considers only one static operating state during peak wind exposure, without modelling time-varying grid operations, load changes, or system adaptations. As a result, this snapshot represents a worst-case estimate of disruption. Moreover, key impacts reported in [5], including damage to hydropower plants due to flooding, are outside the scope of this simulation and may lead to an underestimation of the damage.

The model captures important aspects of system behaviour under climate-related hazards by characterising exposure-driven failures and their cascading impacts across the transmission network. While adaptive responses such as demand-side flexibility, redundancy, or post-failure reconfiguration are not explicitly modelled, the framework is designed to identify vulnerable system components, the spatial propagation of service disruptions, and industrial areas most exposed to power outages. Component failure is represented probabilistically through fragility functions, consistent with existing studies [12], rather than through deterministic exposure-based failure [11]. In line with climate extreme-induced infrastructure risk studies, conservative interpretations of failure states are commonly adopted to represent direct impacts and immediate loss of service when detailed information on system operations is unavailable or highly uncertain [12,54]. Accordingly, loads are assumed to be interrupted immediately once their associated system component becomes inoperative. In this context, the resulting estimates should therefore be interpreted as upper-bound impact that highlight exposure patterns and critical system dependencies.

While the current method is a large step forward, several directions remain open for further work. Future research could improve realism by integrating generation and consumption dynamics, modelling system

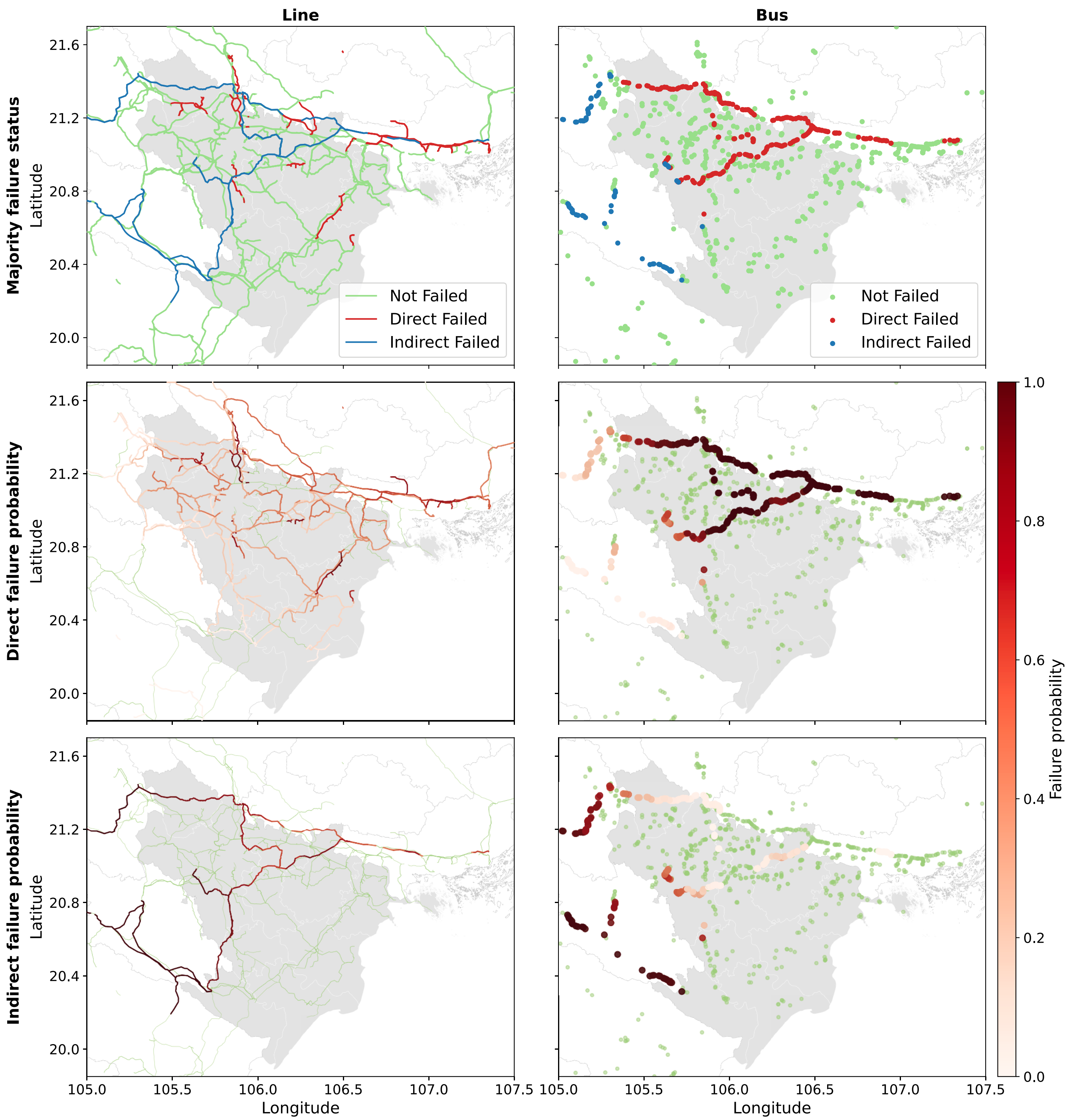

Fig. 7. Failure probability of direct damage under Typhoon Yagi.

Table 1
Summary of direct and indirect failures in lines and buses for 1000 iterations.

Lines				Buses			
Direct failure length (km)	Indirect failure length (km)	Direct failure ratio	Indirect failure ratio	Direct failure count	Indirect failure count	Direct failure ratio	Indirect failure ratio
2440.9	1414.6	5.2 %	3.0 %	248	83	6.7 %	2.3 %

behaviour over multiple time steps, and incorporating recovery processes. These extensions would help to capture how the system evolves after disruptions, show cascading failures over time, and explore possible strategies for restoring electricity for critical areas, which would support a more comprehensive understanding of cascading risks and recovery strategies in evolving hazard conditions. However, implementing these functions would also require more detailed and higher-resolution data, such as time-dependent load profiles, generation

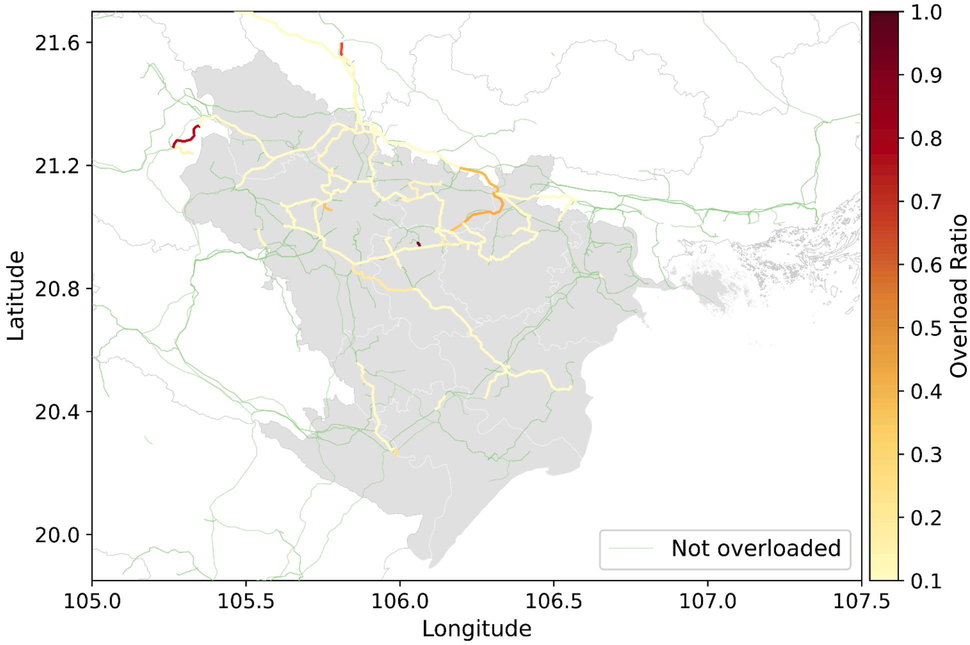

Fig. 8. Probability of line overloads after initial damage.

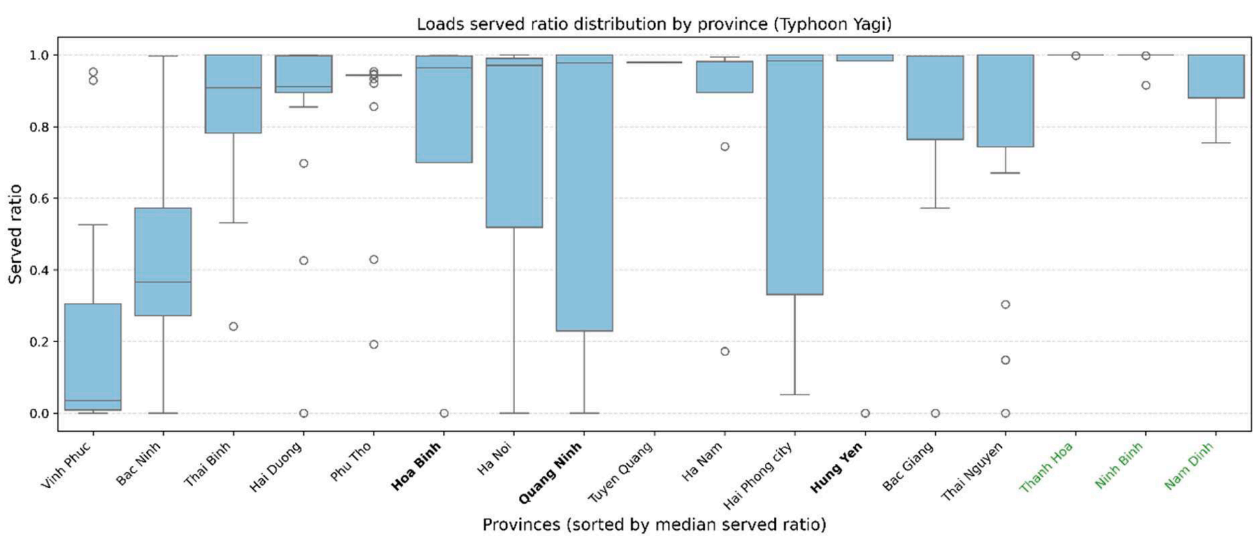

Fig. 9. Boxplot of served load ratio by province. Each box represents the distribution of served ratios across loads in each province, sorted by median value. Province labels are styled based on the standard deviation of served ratios: bold indicates top three high standard deviation (i.e., least stable), while green indicates the top three provinces with the lowest standard deviations (i.e., most stable) among the affected provinces.

availability, and infrastructure restoration rates, which may not always be accessible in data-scarce contexts.

Real-world substation to demand connections are often shaped by utility service territories or administrative boundaries, which strongly influence connection likelihood. For example, a business that is located on the border between two regional utility providers may be required to connect to a substation under its designated provider, even if another substation operated by another provider is physically closer. Future models could integrate this constraint condition to reflect such institutional rules. In some regions, such as Texas in the United States, which operates its own independent power grid (ERCOT), there are only a few connections between ERCOT and grids in the rest of the country, so power cannot flow freely across these boundaries. These topological constraints should be introduced into both dependency heuristics and power flow modelling to prevent unrealistic linkages and improve applicability across national and regional contexts.

Due to the lack of access to real power grid data from utility providers, this study uses OSM data and Google Places API to model the power grid and utility end users. While these data provide a valuable open-source alternative, they also have limitations, as data quality and completeness vary across regions [55]. Compared to infrastructure types like roads [56] and building footprints [57], electricity infrastructure in OSM is less comprehensively mapped. In addition, many power grid assets in OSM lack technical attributes needed for power system modelling. Another important limitation is related to data granularity. The current model aggregates electricity demand at the level of industrial areas without distinguishing between economic subsectors, making it challenging to assess how the grid operational failures impact specific sectors and to quantify the resulting indirect losses through supply chains. Addressing these data gaps is essential, and improving the availability and quality of power infrastructure data should be a priority for future research on power system risk assessment.

5. Conclusion

This paper aims to comprehensively address the challenges of analysing how power grids perform under cascading failures and how such failures lead to service disruptions by introducing an integrated power flow-based modelling framework. A capacity and distance-informed assignment is used to provide a more realistic and generic representation of the connection between economic activities and substations.

Using open-access OSM data as a transparent and reproducible foundation, we develop a framework to simulate both direct and indirect failures within power grids and service disruptions, under random and hazard-driven disruptions, demonstrated through a case study of Typhoon Yagi in Vietnam. Results show that the transmission network is structurally vulnerable, with the giant component fragmenting rapidly under low levels of random disturbances, indicating limited redundancy and a strong dependence on a small subset of components to maintain connectivity. While >95 % of transmission lines do not experience overloads under random disruptions, a small number of critical corridors, primarily located around major urban and industrial centres, exhibit extremely higher overload probabilities.

Under the Typhoon Yagi scenario, both direct and indirect failures contribute substantially to overall system disruption. Approximately 5.2 % of transmission line length and 6.7 % of buses fail directly due to wind exposure, while an additional 3.0 % of line length and 2.3 % of buses fail indirectly through cascading effects. Although the average served load ratio across affected provinces remains relatively high at around 0.91, pronounced spatial disparities are observed. Several provinces experience consistently low or highly unstable service levels, and 26 % of loads loss service in at least one simulation, with a small subset experiencing complete loss of supply across all simulations.

By capturing both physical damages from wind exposures and system-level interactions through power flow modelling, the proposed framework highlights how climate-induced infrastructure failures can propagate beyond initial impact zones and lead to uneven service disruptions. Its transparent and scalable design makes it particularly suitable for risk assessment, resilience planning, and targeted investment decisions, especially in data-scarce environments where open-access data can provide a valuable starting point for modelling.

Code and data availability

Google Places API: Restrictions apply to the availability of these data, which were used under license for this study. The code and exported dataset are available in the GitHub repository “PowerFlowCascade” (https://github.com/Mengqi-Ye/PowerFlowCascade). This repository includes the complete source code required to reproduce all findings and figures presented in this study.

CRediT authorship contribution statement

Mengqi Ye: Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Surender Raj Vanniya Perumal:** Writing – review & editing, Methodology, Investigation. **Miloš Pantoš:** Writing – review & editing, Methodology. **Sebastijan Ribnikar Cimerman:** Writing – review & editing, Methodology. **Philip J Ward:** Writing – review & editing, Supervision. **Elco E Koks:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

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

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ress.2026.112575.

Data availability

I have shared the link to my data and code at the end of the manuscript.

References

[1] Shamsi N, Helmrich A. Interdependency classification: a framework for infrastructure resilience. Env Res: Infrastruct Sustain 2025;5:015009. https://doi.org/10.1088/2634-4505/adac89.
[2] Rinaldi SM, Peerenboom JP, Kelly TK. Identifying, understanding, and analyzing critical infrastructure interdependencies. IEEE Control Syst Mag 2001;21:11–25. https://doi.org/10.1109/37.969131.
[3] Bennett JA, Trevisan CN, DeCarolis JF, Ortiz-García C, Pérez-Lugo M, Etienne BT, et al. Extending energy system modelling to include extreme weather risks and application to hurricane events in Puerto Rico. Nat Energy 2021;6:240–9. https://doi.org/10.1038/s41560-020-00758-6.
[4] Busby JW, Baker K, Bazilian MD, Gilbert AQ, Grubert E, Rai V, et al. Cascading risks: understanding the 2021 winter blackout in Texas. Energy Res Soc Sci 2021;77:102106. https://doi.org/10.1016/j.erss.2021.102106.
[5] United Nations Vietnam. Viet Nam multi-sector assessment (VMSA) report for Typhoon Yagi recovery. 2024.
[6] Watson EB, Etemadi AH. Modeling electrical grid resilience under Hurricane wind conditions with increased solar and wind power generation. IEEE T Power Syst 2020;35:929–37. https://doi.org/10.1109/TPWRS.2019.2942279.
[7] Montoya-Rincon JP, Mejia-Manrique SA, Azad S, Ghandehari M, Harmsen EW, Khanbilvardi R, et al. A socio-technical approach for the assessment of critical infrastructure system vulnerability in extreme weather events. Nat Energy 2023;8:1002–12. https://doi.org/10.1038/s41560-023-01315-7.
[8] Singh EA. Compounding impacts of lifeline infrastructure failure during natural hazard events. In: Karácsonyi D, Taylor A, Bird D, editors. The demography of disasters: impacts for population and place. Cham: Springer International Publishing; 2021. p. 189–210. https://doi.org/10.1007/978-3-030-49920-4_10.
[9] Panteli M, Mancarella P. The grid: stronger, bigger, smarter?: presenting a conceptual framework of power system resilience. Power and Energy Magazine. IEEE 2015;13:58–66. https://doi.org/10.1109/MPE.2015.2397334.
[10] Intergovernmental Panel on Climate Change (IPCC). Climate Change 2021: the physical science basis: working Group I. Contribution to the sixth assessment report of the intergovernmental panel on climate change. Cambridge: Cambridge University Press; 2023. https://doi.org/10.1017/9781009157896.
[11] Brunner LG, Peer RAM, Zorn C, Paulik R, Logan TM. Understanding cascading risks through real-world interdependent urban infrastructure. Reliab Eng Syst Saf 2024;241:109653. https://doi.org/10.1016/j.ress.2023.109653.
[12] Muehlhofer E, Koks EE, Kropf CM, Sansavini G, Bresch DN. A generalized natural hazard risk modelling framework for infrastructure failure cascades. Reliab Eng Syst Saf 2023;234. https://doi.org/10.1016/j.ress.2023.109194.
[13] Koks E, Pant R, Thacker S, Hall JW. Understanding business disruption and economic losses due to electricity failures and flooding. Int J Disast Risk *Sc* 2019;10:421–38. https://doi.org/10.1007/s13753-019-00236-y.
[14] Mylonas C, Mitsakis E, Kepaptsoglou K. Criticality analysis in road networks with graph-theoretic measures, traffic assignment, and simulation. Phys A: Stat Mech Appl 2023;629:129197. https://doi.org/10.1016/j.physa.2023.129197.
[15] Thacker S, Pant R, Hall JW. System-of-systems formulation and disruption analysis for multi-scale critical national infrastructures. Reliab Eng Syst Safe 2017;167:30–41. https://doi.org/10.1016/j.ress.2017.04.023.
[16] Koks E, Rozenberg J, Tariverdi M, Dickens B, Fox C, van Ginkel K, et al. A global assessment of national road network vulnerability. Env Res: Infrastruct Sustain 2023. https://doi.org/10.1088/2634-4505/acd1aa.
[17] Abdulla B, Birgisson B. Characterization of vulnerability of road networks to random and nonrandom disruptions using network percolation approach. J Comput Civ Eng 2021;35:04020054. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000938.
[18] Dong S, Yu T, Farahmand H, Mostafavi A. Probabilistic modeling of cascading failure risk in interdependent channel and road networks in urban flooding. Sustain Cities Soc 2020;62:102398. https://doi.org/10.1016/j.scs.2020.102398.
[19] Wang W, Karimi F, Khalilpour K, Green D, Varvarigos M. Robustness analysis of electricity networks against failure or attack: the case of the Australian National Electricity Market (NEM). Int J Crit Infrastruct Prot 2023;41. https://doi.org/10.1016/j.ijcip.2023.100600.
[20] Hamedmoghadam H, Jalili M, Vu HL, Stone L. Percolation of heterogeneous flows uncovers the bottlenecks of infrastructure networks. Nat Commun 2021;12:1254. https://doi.org/10.1038/s41467-021-21483-y.
[21] Mühlhofer E, Bresch DN, Koks EE. Infrastructure failure cascades quintuple risk of storm and flood-induced service disruptions across the globe. One Earth 2024;7:714–29. https://doi.org/10.1016/j.oneear.2024.03.010.
[22] MdS Alam, Kim K, Horner MW, Alisan O, Antwi R, Ozguven EE. Large-scale modeling of hurricane flooding and disrupted infrastructure impacts on

accessibility to critical facilities. J Transp Geogr 2024;116:103852. https://doi.org/10.1016/j.jtrangeo.2024.103852.
[23] Banze T, Kneiske TM. Open data for energy networks: introducing DAVE—A data fusion tool for automated network generation. Sci Rep 2024;14:1938. https://doi.org/10.1038/s41598-024-52199-w.
[24] Xiong B, Fioriti D, Neumann F, Riepin I, Brown T. Modelling the high-voltage grid using open data for. Eur beyond Sci Data 2025;12:277. https://doi.org/10.1038/s41597-025-04550-7.
[25] Bollinger LA, Dijkema GPJ. Evaluating infrastructure resilience to extreme weather – the case of the Dutch electricity transmission network. Eur J Transp Infrastruct Res 2016;16. https://doi.org/10.18757/ejtir.2016.16.1.3122.
[26] Alhamwi A, Medjroubi W, Vogt T, Agert C. OpenStreetMap data in modelling the urban energy infrastructure: a first assessment and analysis. Energy Procedia 2017;142:1968–76. https://doi.org/10.1016/j.egypro.2017.12.397.
[27] Wiegmans B. Improving the topology of an electric network model based on open data. University of Groningen; 2016.
[28] Medjroubi W, Müller UP, Scharf M, Matke C, Kleinhans D. Open data in power grid modelling: new approaches towards transparent grid models. Energy Rep 2017;3:14–21. https://doi.org/10.1016/j.egyr.2016.12.001.
[29] Gaugl R, Wogrin S, Bachhiesl U, Frauenlob L. GridTool: an open-source tool to convert electricity grid data. SoftwareX 2023;21:101314. https://doi.org/10.1016/j.softx.2023.101314.
[30] Verschuur J, Balakrishnan S. Brief communication: evidence of the impacts of climate extremes on power system outages in India. EGUsphere 2024:1–14. https://doi.org/10.5194/egusphere-2024-3176.
[31] Thacker S, Barr S, Pant R, Hall JW, Alderson D. Geographic hotspots of critical national infrastructure: geographic hotspots of critical national infrastructure. Risk Anal 2017;37:2490–505. https://doi.org/10.1111/risa.12840.
[32] Johansson J, Hassel H. An approach for modelling interdependent infrastructures in the context of vulnerability analysis. Reliab Eng Syst Saf 2010;95:1335–44. https://doi.org/10.1016/j.ress.2010.06.010.
[33] Zorn C, Pant R, Thacker S, Shamseldin AY. Evaluating the magnitude and spatial extent of disruptions across interdependent national infrastructure networks. ASCE-ASME J Risk Uncert Engrg Sys B Mech Engrg 2020;6. https://doi.org/10.1115/1.4046327.
[34] Xie X, Zhou Y, Xu Y, Hu Y, Wu C. OpenStreetMap data quality assessment via Deep learning and Remote sensing imagery. IEEE Access 2019;7. https://doi.org/10.1109/ACCESS.2019.2957825. 1–1.
[35] Qi Z, Zhang Y, Chang K, Brovelli M. Assessing OSM building completeness for almost 13,000 cities globally. Int J Digit Earth 2022;15:2400–21. https://doi.org/10.1080/17538947.2022.2159550.
[36] Ullah T, Lautenbach S, Herfort B, Reinmuth M, Schorlemmer D. Assessing completeness of OpenStreetMap building footprints using MapSwipe. ISPRS Int J Geoinf 2023;12:143. https://doi.org/10.3390/ijgi12040143.
[37] Alhamwi A, Medjroubi W, Vogt T, Agert C. GIS-based urban energy systems models and tools: introducing a model for the optimisation of flexibilisation technologies in urban areas. Appl Energy 2017;191:1–9. https://doi.org/10.1016/j.apenergy.2017.01.048.
[38] Braese J, De Vries Robbe SA, Rentschler J. Coastal development between opportunity and disaster risk: a multisectoral risk assessment for vietnam. World Bank Group; 2020.
[39] van Ginkel KCH, Dottori F, Alfieri L, Feyen L, Koks EE. Flood risk assessment of the European road network. Nat Hazards Earth Syst Sci 2021;21:1011–27. https://doi.org/10.5194/nhess-21-1011-2021.
[40] Ye M, Ward PJ, Bloemendaal N, Nirandjan S, Koks EE. Risk of tropical cyclones and floods to power grids in Southeast and East Asia. Int J Disaster Risk Sci 2024;15:494–507. https://doi.org/10.1007/s13753-024-00573-7.
[41] Koks E., Mühlhofer E., Riedel L., Kropf C.M. OSM-flex: a Python package for flexible data extraction from OpenStreetMap 2023. https://doi.org/10.5281/zenodo.10204123.
[42] Mühlhofer E, Kropf C, Bresch D, Koks E. OpenStreetMap for multi-faceted climate risk assessments. Environ Sci; 2023. https://doi.org/10.31223/X5SQ2J.
[43] National Energy Effeciency Programme. Vietnam Energy Statistics. Institute of Energy; 2021.
[44] Thurner L, Scheidler A, Schäfer F, Menke J-H, Dollichon J, Meier F, et al. pandapower - an open source Python tool for convenient modeling, analysis and optimization of Electric Power systems. IEEE Trans Power Syst 2018;33:6510–21. https://doi.org/10.1109/TPWRS.2018.2829021.
[45] de Vries Robbé S, Rentschler J, Braese J, Nguyen DH, van Ledden M, Pozueta Mayo B. Resilient shores: vietnam's coastal development between opportunity and disaster risk. World Bank; 2020.
[46] Bresch DN, Aznar-Siguan G. CLIMADA v1.4.1: towards a globally consistent adaptation options appraisal tool. Geosci Model Dev 2021;14:351–63. https://doi.org/10.5194/gmd-14-351-2021.
[47] Knapp KR, Kruk MC, Levinson DH, Diamond HJ, Neumann CJ. The International Best track Archive for Climate stewardship (IBTrACS): unifying tropical cyclone data. Bull Am Meteorol Soc 2010;91:363–76. https://doi.org/10.1175/2009BAMS2755.1.
[48] EVN. EVN Annual Report 2022-2023. 2024.
[49] Wrana J., Nguyen T. "Strategic coupling" and regional development in a transition economy: what can we learn from Vietnam? 2019. https://doi.org/10.1080/23792949.2019.1608838.
[50] Zorn C., Koks E., Eggimann S. Demand side management impacts on electricity network vulnerability 2020. https://doi.org/10.48550/arXiv.2008.04954.
[51] Asaridis P, Molinari D, Di Maio F, Ballio F, Zio E. A probabilistic modeling and simulation framework for power grid flood risk assessment. Int J Disaster Risk Reduct 2025;120:105353. https://doi.org/10.1016/j.ijdrr.2025.105353.
[52] Karakoc DB, Barker K, Zobel CW, Almoghathawi Y. Social vulnerability and equity perspectives on interdependent infrastructure network component importance. Sustain Cities Soc 2020;57:102072. https://doi.org/10.1016/j.scs.2020.102072.
[53] Pant R, Thacker S, Hall JW, Alderson D, Barr S. Critical infrastructure impact assessment due to flood exposure: critical infrastructure impact assessment due to flood exposure. J Flood Risk Manag 2018;11:22–33. https://doi.org/10.1111/jfr3.12288.
[54] Ouyang M. Review on modeling and simulation of interdependent critical infrastructure systems. Reliab Eng Syst Saf 2014;121:43–60. https://doi.org/10.1016/j.ress.2013.06.040.
[55] Senaratne H, Mobasheri A, Ali AL, Capineri C, Haklay M (Muki). A review of volunteered geographic information quality assessment methods. Int J Geogr Inf Sci 2017;31:139–67. https://doi.org/10.1080/13658816.2016.1189556.
[56] Barrington-Leigh C, Millard-Ball A. The world's user-generated road map is >80% complete. PLOS ONE 2017;12:e0180698. https://doi.org/10.1371/journal.pone.0180698.
[57] Herfort B, Lautenbach S, Porto de Albuquerque J, Anderson J, Zipf A. A spatio-temporal analysis investigating completeness and inequalities of global urban building data in OpenStreetMap. Nat Commun 2023;14:3985. https://doi.org/10.1038/s41467-023-39698-6.