

## Article

# An Integrated Geotechnical Ground–HAZUS Framework for Urban Seismic Vulnerability Assessment in Seoul, Korea

Han-Saem Kim <sup>1</sup>  and Ju-Hyung Lee <sup>2,\*</sup><sup>1</sup> Department of Civil and Environmental Engineering, Dongguk University, Seoul 04620, Republic of Korea; hansaem@dgu.ac.kr<sup>2</sup> Department of Geotechnical Engineering Research, Korea Institute of Civil Engineering and Building Technology, Goyang 10223, Republic of Korea

\* Correspondence: leejh73@kict.re.kr

## Abstract

This study presents an integrated framework that couples three-dimensional geotechnical ground modeling with a HAZUS-based urban seismic vulnerability assessment for Seoul, Korea. Over 63,000 boreholes, in situ seismic tests, and building inventory records were compiled into a unified relational database following rigorous multi-stage quality control. A multi-parameter  $N$ – $V_s$  regression model was calibrated to supplement missing shear-wave velocity ( $V_s$ ) data, reducing prediction errors by more than 20% relative to conventional empirical equations. Based on the quality-controlled  $V_s$  dataset, a high-resolution three-dimensional  $V_s$ –ground model was constructed to represent subsurface heterogeneity and associated uncertainty across the metropolitan area. The building inventory, comprising approximately 700,000 structures, was standardized according to the HAZUS structural taxonomy and mapped to Korean seismic design eras, enabling a Seoul-adapted vulnerability assessment in which exposure characterization and seismic demand are localized. Site-specific ground-motion amplification and response spectra derived from the 3D ground model were used to modify the spectral acceleration input to the HAZUS fragility functions. Results reveal pronounced spatial variability in site conditions, with northern mountainous zones corresponding primarily to NEHRP Site Class B, central districts to Class C, and southern alluvial basins to Classes D–E, producing amplification differences of up to 1.7 under identical input spectral accelerations. High-risk zones such as Gangnam, Songpa, and Yeouido exhibit concentrated expected damage where thick alluvial deposits coincide with dense stocks of mid-rise reinforced-concrete buildings. Overall, the study demonstrates that integrating high-resolution 3D geotechnical ground models with HAZUS-based fragility analysis provides a physically consistent and data-driven basis for urban-scale seismic risk assessment and resilience planning.



Academic Editor: Tiago Filipe da Silva Miranda

Received: 14 December 2025

Revised: 12 January 2026

Accepted: 14 January 2026

Published: 29 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

**Keywords:** geotechnical database;  $V_{s30}$ ;  $N$ – $V_s$  regression; 3D geotechnical ground model; kriging; sequential Gaussian simulation; HAZUS; urban seismic risk

## 1. Introduction

Seoul is one of the most densely populated metropolitan regions in Korea and accommodates critical administrative, economic, and transportation infrastructures. Although it has traditionally been regarded as a region of moderate seismicity, recent inland earthquakes—such as the 2016 Gyeongju (M5.8) and 2017 Pohang (M5.4) events—have clearly shown that even moderate-magnitude earthquakes can lead to substantial localized

damage when unfavorable site conditions are present [1,2]. In particular, the Pohang earthquake revealed pronounced liquefaction, ground deformation, and structural damage in reclaimed and alluvial areas, underscoring that urban-scale seismic risk cannot be assessed solely in terms of regional ground-shaking intensity. Realistic risk evaluation must explicitly account for soil amplification, liquefaction potential, basin and slope effects, and the vulnerability of buildings and lifelines that are concentrated in limited areas of a city [3–5].

Recent inland earthquakes—such as the 2016 Gyeongju (M5.8) and 2017 Pohang (M5.4) events—have demonstrated that even moderate-magnitude earthquakes can cause substantial localized damage when unfavorable site conditions are present. In particular, the Pohang earthquake highlighted the vulnerability of reclaimed and alluvial areas through manifestations such as liquefaction and ground deformation, underscoring the broader importance of site-specific geotechnical effects in urban seismic risk assessment. In this study, liquefaction is referenced as a motivating example of the limitations inherent in regionally uniform or simplified seismic assessment practices, rather than as an analytical target. The scope of the present work is focused on ground-motion amplification, site response, and resulting structural vulnerability under seismic loading, which represent the dominant contributors to urban-scale seismic risk in densely built metropolitan environments.

The subsurface environment of Seoul is characterized by marked spatial heterogeneity. Northern districts are generally underlain by shallow, stiff granite and gneissic bedrock, while the Han River alluvial plain and reclaimed zones are composed of thick, soft deposits with low shear-wave velocities. As later demonstrated in this study through three-dimensional mapping of bedrock depth,  $V_{s30}$ , the average shear-wave velocity of sediments above engineering bedrock ( $V_{S,soil}$ ), and the fundamental site period  $T_G$ , this contrast produces substantial variations in site amplification and predominant periods over relatively short distances. Nevertheless, current domestic seismic assessment practices remain largely fragmented. They tend to rely on simple indicators such as compliance with structural design codes or building age, and they often neglect the systematic integration of geotechnical, structural, and socioeconomic factors. Existing national-scale or city-scale seismic hazard and risk maps for Seoul therefore employ coarse soil classifications or district-averaged site classes [6–8], and may not reproduce the spatial patterns of damage that would be expected given the actual subsurface structure and building stock.

Internationally, the HAZUS-MH framework developed by the Federal Emergency Management Agency (FEMA) has become one of the most widely used platforms for multi-hazard loss estimation at national and metropolitan scales [9]. HAZUS integrates ground-motion parameters, building inventories, fragility functions, and socioeconomic indicators within a unified GIS-based system, enabling scenario-based assessments of structural damage, economic loss, casualties, and lifeline disruption. Japan has combined HAZUS-type approaches with dense geotechnical and seismological networks such as KiK-net and the J-SHIS system to derive site-specific amplification characteristics and urban-scale seismic hazard maps [10,11]. Several European countries, including Italy and Greece, have localized fragility functions through regional calibration with observed damage and national building typologies, and have implemented GIS-based platforms that explicitly treat basin effects and building stock variability [12–14]. These efforts collectively demonstrate that physically consistent, data-rich subsurface and building information is essential for realistic urban seismic risk assessment.

Despite this international progress, HAZUS applications in Korea remain limited. Existing studies have mainly focused on pilot implementations for a small number of provincial cities, such as Pohang, Ulsan, and Daegu, often employing simplified soil classifications, generic site coefficients, or coarse representations of ground conditions [15–17]. Although these works are valuable as initial demonstrations of HAZUS-based loss es-

timation, they generally do not incorporate high-resolution three-dimensional ground models or large, integrated geotechnical databases. As a result, their capability to represent the complex subsurface and heterogeneous building stock of large metropolitan areas is restricted. For a megacity such as Seoul—where high-rise buildings, underground spaces, and critical infrastructures are densely concentrated—direct adoption of default HAZUS modules without localization of geotechnical and structural parameters is unlikely to capture key effects such as strong site amplification, deep-basin resonance, or soil–structure interaction [18,19].

In parallel with the development of loss estimation frameworks, numerous studies have addressed the characterization and modeling of shear-wave velocity ( $V_s$ ) and its correlation with in situ penetration tests. Classical empirical correlations between  $V_s$  and the standard penetration test blow count ( $N$ ) have been widely developed and applied, especially in Japan and other seismically active regions [20–22]. Subsequent works refined these relationships using larger datasets and examined regional variability in the  $V_s$ – $N$  correlation, as well as the influence of soil type and geological age [23,24]. In Korea, several studies have derived correlations between  $V_s$  and various geotechnical in situ test data and synthesized seismic test results to define representative  $V_s$  profiles for major geotechnical layers and geomorphological units [25–28]. More recent research has introduced multi-parameter or machine-learning-based  $V_s$  prediction models that incorporate soil classification, stratigraphy, and geomorphologic indicators, and has developed regional and national  $V_{s30}$  models by combining SPT data with broadband seismic observations [29–31].

While these prior studies provide important empirical and methodological foundations for  $V_s$  prediction and  $V_{s30}$  mapping, most of them focus on one-dimensional site characterization or broad regional-scale models and do not embed the results within a three-dimensional, city-wide ground model that is directly linked to HAZUS-type vulnerability assessments. Similarly, existing Korean seismic risk studies have typically emphasized either structural vulnerability or ground-motion hazard, but rarely their fully integrated interaction in a metropolitan context [15,18,32]. The lack of a unified geotechnical database and the absence of a 3D  $V_s$ -based ground model that can be directly coupled with HAZUS have therefore been major obstacles to establishing a realistic, Seoul-specific seismic risk assessment framework.

While several previous studies have applied HAZUS or HAZUS-type methodologies to urban seismic risk assessment, most applications rely either on the default HAZUS software environment or on simplified representations of subsurface conditions, typically through uniform site classes or coarse  $V_{s30}$ -based zonations. In such cases, site effects are treated as predefined modifiers rather than as spatially continuous, physically derived parameters, and subsurface uncertainty is rarely quantified explicitly.

In contrast, the present study introduces an integrated workflow in which a high-resolution, city-scale three-dimensional geotechnical ground model directly governs site response, amplification, and seismic demand prior to fragility evaluation. The proposed framework differs fundamentally from direct use of HAZUS software in that it replaces generic site-condition inputs with a locally calibrated 3D  $V_s$ –ground model constructed from more than 63,000 boreholes and in situ seismic tests, supplemented by a multi-parameter  $N$ – $V_s$  regression tailored to the geological and depositional conditions of Seoul.

Compared with previous international urban-scale studies, which often rely on sparse geotechnical data, topographic proxies, or two-dimensional interpolations, this framework provides a spatially continuous, voxel-based representation of subsurface stiffness and its associated uncertainty. Ordinary kriging and sequential Gaussian simulation are used not only to interpolate  $V_s$  values, but also to quantify spatial variability, allowing geotechnical

uncertainty to be explicitly propagated into site-response characterization and subsequent vulnerability assessment.

As a result, the proposed approach bridges the gap between advanced geotechnical ground modeling and standardized fragility-based loss estimation. Rather than treating ground conditions and structural vulnerability as loosely coupled components, the framework establishes a physically consistent linkage between subsurface heterogeneity, site-specific seismic demand, and building damage probabilities. This integrated structure represents a key methodological advance over conventional HAZUS applications and provides a transferable template for seismic risk assessment in other data-rich metropolitan regions. In this study, “adaptation” refers to the localization of exposure and seismic demand using the Seoul building inventory and the 3D ground model; the fragility functions are HAZUS-based and are not calibrated using Seoul-specific damage observations, which is discussed as a limitation and a priority for future work.

## 2. Materials and Methods

### 2.1. Construction of the Multi-Source Geotechnical Database

To support the development of an urban-scale seismic vulnerability assessment framework, a comprehensive multi-source geotechnical database was constructed for Seoul and its surrounding region. The target area covers approximately 1000 km<sup>2</sup>, including the entire administrative area of Seoul (about 605 km<sup>2</sup>) and adjacent portions of Gyeonggi Province (about 420 km<sup>2</sup>). Extending the domain beyond the administrative boundary is essential because the northern mountains, the central basin, and the southern alluvial plain share a physically continuous wave-propagation environment despite their contrasting geological characteristics [32]. Within this domain, more than 63,000 boreholes and associated in situ tests were compiled and standardized to form an integrated geotechnical information system (Figure 1a).

Geotechnical data were collected from multiple institutional sources, including the National Geotechnical Information Infrastructure (NGII GeoInfo), the Seoul Metropolitan Government and its Urban Infrastructure Headquarters, the Seoul Institute, Dongguk University research archives, the Korea Institute of Geoscience and Mineral Resources (KIGAM), the Korea Institute of Civil Engineering and Building Technology (KICT), and the National Disaster Management Research Institute (NDMI). The main data types consist of (i) borehole logs containing stratigraphic descriptions, depth intervals, sampling information, SPT-*N* values, and groundwater levels; (ii) geophysical test results such as downhole tests, PS-logging, S-wave seismic surveys, and surface-wave methods (e.g., MASW); and (iii) 1:25,000 geological maps, a 5 m-resolution digital elevation model (DEM), and auxiliary field investigation records from major construction and urban redevelopment sites [2,23,25,32]. The compiled dataset spans investigation years from 2003 to 2025, with a notable increase in the proportion of SPT and *V<sub>s</sub>* measurements after 2010, reflecting heightened demand for quantitative seismic hazard and vulnerability evaluations in Korea [2,23,24,31]. Figure 1a illustrates the spatial distribution of the compiled boreholes overlaid on the river network and urban base map, whereas Figure 1a,b presents the regional topography and generalized geology of the Seoul metropolitan area, respectively. The DEM clearly delineates the high-relief granitic mountains in the north and south and the low-lying Han River alluvial plain, while the geological map emphasizes the contrast between mountainous bedrock zones and basin-fill deposits along the river corridor [32–35].

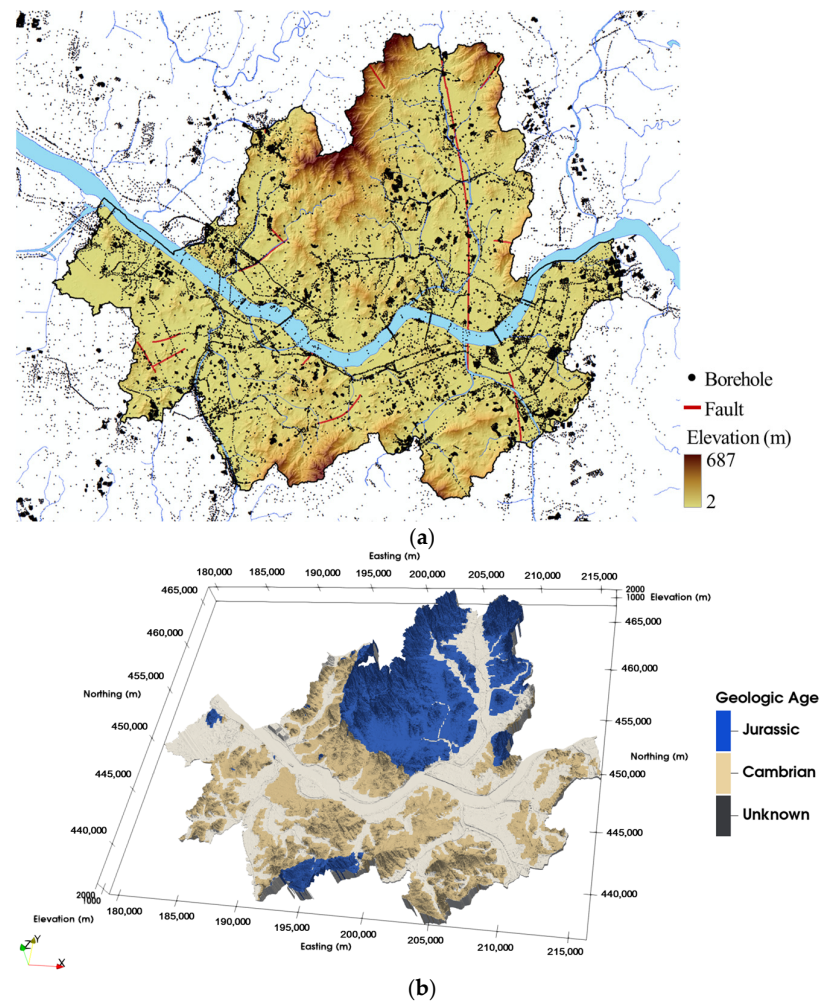


**Figure 1.** (a) Integrated geotechnical borehole database of the Seoul metropolitan area, comprising approximately 63,000 records, overlaid with elevation and major fault lines; (b) updated geological map of the Seoul metropolitan area, revised based on regionally validated geological interpretations from previous studies [23].

Because the original data were produced by different agencies for heterogeneous project objectives, substantial discrepancies existed in coordinate systems (WGS84, UTM-K, TM), depth reference levels (ground surface versus mean sea level), stratigraphic naming conventions, and engineering units. A unified data-processing pipeline was therefore established. First, all spatial coordinates were transformed to the Korea 2000 national grid (EPSG:5179), and depths were standardized to a ground-surface reference; units were converted to a consistent SI system (m, m/s, kPa). Second, stratigraphic descriptions were harmonized into a common engineering classification (e.g., fill, alluvium, weathered soil, weathered rock, soft rock, hard rock), enabling consistent grouping of layers with similar mechanical behavior, following standard soil classification practice [22,36,37]. Third, essential site-response parameters such as  $V_{s30}$ , depth to engineering bedrock (H), and fundamental site period ( $T_G$ ) were derived for each borehole location using standardized procedures consistent with NEHRP and HAZUS guidelines [38–46]. Finally, all processed information was stored in a relational database structure and visualized in a GIS environment, providing an internally consistent geospatial dataset that can be directly linked to building and underground-utility databases [2,23,25,32].

Operational definition of engineering bedrock depth (H). In this study, engineering bedrock depth H is defined as the depth from the ground surface to the first depth at

which  $V_s \geq 760$  m/s, following the definition of bedrock/engineering bedrock adopted in the Korean seismic design standard (KDS 17 10 00), where bedrock is defined as a stratum with shear-wave velocity of 760 m/s or greater. This threshold is also consistent with the commonly used NEHRP/HAZUS stiffness boundary adopted in international site-classification practice.

To facilitate multi-layered data interoperability and future extension to other hazards, the database was implemented in a relational database management system benchmarked against internationally recognized geotechnical and ground-motion databases, including the PEER NGA-West2 database, the Next Generation Liquefaction (NGL) database, the British Geological Survey GeoIndex, and the New Zealand Geotechnical Database (NZGD) [47–50]. The core schema consists of a project table (metadata), a borehole table storing coordinates, drilling attributes, investigation year, and groundwater information, a geolayer table describing stratigraphic units and thicknesses, a  $\text{test}_{\text{result}}$  table containing  $V_s$ ,  $V_p$ , SPT- $N$ , strength, and stiffness parameters (including regressed  $V_s$  values), and a  $\text{QC}_{\text{record}}$  table that logs the results of subsequent quality-control procedures [25,32]. Each table is linked through unique identifiers and geometry fields, allowing efficient spatial queries and direct connection to HAZUS-compatible input variables such as site class,  $V_{s30}$ , and  $T_G$  [21,28,41–44]. This standardized geotechnical database forms the foundational input for the  $N$ – $V_s$  regression modeling, three-dimensional  $V_s$ –ground modeling, and hybrid seismic vulnerability assessment described in the following sections.

## 2.2. Multi-Step Quality Control and Outlier Validation

The reliability of the integrated geotechnical database directly governs the credibility of subsequent site-response analyses, liquefaction evaluation, and urban-scale seismic vulnerability assessment. For this reason, a multi-step quality control (QC) and outlier validation procedure was developed by adapting the principles of international standards (e.g., ISO/TC182, ASTM D420, BS EN ISO 14688) and the QC practices used in large geotechnical databases worldwide [22,37,47–50]. The conceptual workflow is summarized in Figure 2. The left-hand side of the diagram represents the ingestion of borehole, geo-layer, and in situ test information (SPT, CPT, and seismic tests) into the relational database. The central block denotes the construction of the geotechnical database and the subsequent determination of multi-site-response parameters such as  $V_{s30}$ , depth to engineering bedrock (H), and fundamental site period  $T_G$ . The right-hand side depicts three tiers of automated checks—mandatory-field and input-error checks, spatial coincidence and layer-consistency checks, and test-result checks—each feeding into a four-step outlier verification module. Data that pass all checks are labeled as “OK” and become part of the final database used for liquefaction and seismic-response analyses.

The first tier comprises format and logical consistency checks. Missing coordinates, inconsistent units (e.g., depth erroneously stored in feet), and obvious typographical errors are automatically detected and corrected using rule-based scripts. Depth-order inconsistencies (e.g., records where  $\text{Depth}_{\text{from}}$  exceeds  $\text{Depth}_{\text{to}}$ ), duplicated boreholes within a  $\pm 5$  m radius and within a two-year investigation window, and conflicting stratigraphic labels are identified and either reconciled or removed. In addition, basic physical plausibility is examined by comparing paired parameters such as  $V_s$  and SPT- $N$  at the same depth: for instance, records with  $N = 0$  but  $V_s > 500$  m/s are treated as erroneous entries. Spatial validation is carried out in a GIS environment to detect duplicated coordinates, overlapping boreholes, and unrealistic locations such as boreholes located within the Han River channel. Through this automated pre-QC process, approximately 5.8% of the original entries were identified as duplicates or inconsistent records and were either merged or discarded, resulting in a spatially coherent, format-harmonized dataset ready for quantitative outlier analysis.

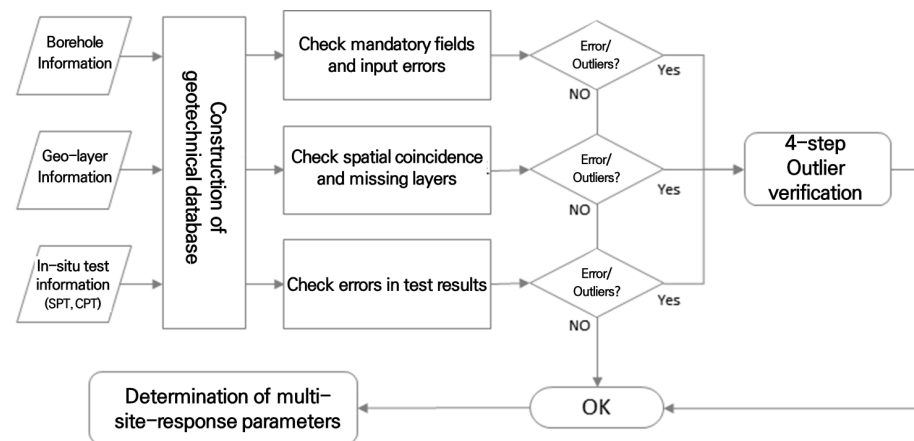


**Figure 2.** Conceptual workflow of multi-step quality control and data validation for liquefaction evaluation using the geotechnical database.

Beyond these preliminary checks, a four-stage hybrid outlier validation scheme was implemented to secure the physical, statistical, and spatial integrity of key geotechnical parameters ( $V_s$ , SPT- $N$ , and  $H$ ). The multi-stage framework, illustrated conceptually in Figure 3, combines (a) a classical 3-sigma rule, (b) generalized extreme value (GEV) statistics derived from kriging-based cross-validation, (c) spatial-outlier detection using Moran's  $I$ , and (d) cluster-based validation using K-means clustering [27,38–40]. Each method captures a different aspect of anomaly: univariate extremes in the marginal distributions, inconsistency with local spatial trends, disruption of spatial continuity, and atypical behavior relative to peer groups, respectively. Bivariate correlations among geotechnical properties and hazard-related indices ( $V_s$ ,  $N$ ,  $H$ ,  $V_{s30}$ , and liquefaction-related proxies) were used to guide the choice of thresholds and to ensure that the outlier detection focuses on parameters that are most influential for seismic and liquefaction hazard assessment [2,23,31].

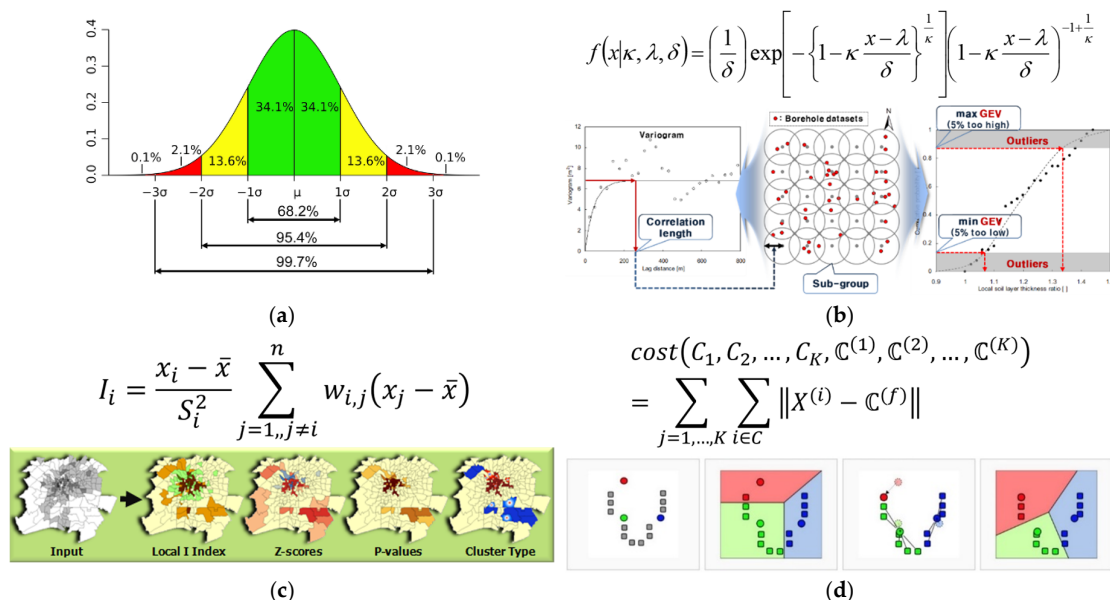


**Figure 3.** Multi-stage hybrid outlier detection framework: (a) 3-Sigma-rule; (b) Generalized Extreme Value Distribution; (c) Moran's outlier (blue zone: Moran's  $I \geq 0.75$ ; blue zone:  $0.75 > \text{Moran's } I \geq 0.5$ ; blue zone:  $0.5 > \text{Moran's } I \geq 0.25$ ; blue zone:  $\text{Moran's } I < 0.25$ ); (d) K-mean clustering.

In the first stage, depth-wise distributions of  $V_s$  and SPT- $N$  were analyzed under an approximate normality assumption. For each depth interval, the mean  $\mu$  and standard

deviation  $\sigma$  were calculated, and values lying outside the range  $\mu \pm 3\sigma$  were flagged as provisional outliers. As a representative example, at a depth of 10 m the mean  $V_s$  was about 250 m/s and the standard deviation was about 60 m/s; thus, measurements below roughly 70 m/s or above about 430 m/s were classified as first-stage anomalies. This step removed about 4.2% of the raw samples, mainly extreme values that were inconsistent with the overall depth-wise trends. Figure 3a schematically illustrates this 3-sigma rule, with outliers populating the tails of the distribution beyond  $\pm 3\sigma$ .

The second stage employs kriging-based cross-validation to evaluate local spatial consistency. Using the experimental variogram of  $V_s$ , ordinary kriging was performed to predict  $V_s$  at each borehole location from neighboring boreholes [39,40]. The relative root-mean-square error (RRMSE) was then computed as

$$\text{RRMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( \frac{V_{s,i}^{\text{meas}} - V_{s,i}^{\text{pred}}}{V_{s,i}^{\text{meas}}} \right)^2}, \quad (1)$$

where  $V_{s,i}^{\text{meas}}$  and  $V_{s,i}^{\text{pred}}$  denote the measured and kriging-predicted shear-wave velocities at location  $i$ , and  $n$  is the number of control points. Records with  $\text{RRMSE} > 0.25$  were regarded as statistically unstable in the context of their local neighborhood, suggesting either measurement error or extremely localized heterogeneity. The distribution of GEV-based tail probabilities for these RRMSE values, shown conceptually in Figure 3b, was used to confirm that the chosen threshold isolates only the most inconsistent points.

The third stage focuses on the spatial continuity of  $V_s$  using Moran's I statistic [38].  $V_s$  values were aggregated to a 100 m  $\times$  100 m grid covering the study area, and global as well as local Moran's I indices were computed. The overall global Moran's I for the quality-controlled dataset was about 0.61, indicating moderate-to-strong positive spatial autocorrelation. However, certain alluvial and reclaimed zones—particularly Yeouido, parts of Yeongdeungpo, and sections of the Gangnam alluvial plain—exhibited local Moran's I values below 0.3, revealing highly heterogeneous subsurface conditions. Within these low-autocorrelation zones, individual grid cells whose local I and associated z-scores deviated markedly from their neighbors were flagged as spatial outliers. Figure 3c schematically depicts the transformation from the raw  $V_s$  map to local Moran's I, z-score, and cluster-type maps, which together highlight anomalous hot and cold spots.

In the fourth stage, unsupervised clustering was adopted to identify observations whose combined  $V_s$ – $N$ –depth characteristics deviate from those of similar sites. Normalized  $V_s$ ,  $N$ , and depth values were used as feature variables in a K-means clustering analysis with  $k = 5$ , resulting in clusters that can be interpreted in terms of representative geotechnical environments such as shallow bedrock, weathered rock, central basin deposits, riverine alluvium, and artificial fill. For each cluster, the centroid behavior of  $V_s$  versus depth and  $V_s$  versus  $N$  was established, and points with coefficient-of-determination deviations exceeding  $R_{\text{dev}}^2 > 0.15$  from the centroid trend were identified as candidate outliers. Clusters dominated by riverine alluvium and artificial fill exhibited particularly wide residual distributions, reflecting their inherently heterogeneous composition. The conceptual behavior of the K-means clustering and its decision boundaries is shown in Figure 3d.

An ensemble decision rule was applied to integrate the results of the four stages. Only records flagged as anomalous by at least three of the four methods—3-sigma, kriging-based cross-validation, Moran's I spatial analysis, and K-means clustering—were classified as true outliers and removed from the database. This ensemble-based approach leverages the complementary strengths of statistical, spatial, and cluster-based diagnostics while minimizing the risk of discarding physically meaningful but locally extreme values. After

applying this rule, approximately 7.5% of the borehole records (about 4720 entries) were eliminated. Most of the removed records were concentrated in the Han River alluvial belt, reclaimed urban districts, and low-lying river terraces, where strong lateral variability of geomaterials and groundwater conditions is expected.

Post-QC evaluation confirmed that the hybrid procedure substantially improved the internal consistency of the database. The global RRMSE between measured and kriging-predicted  $V_s$  values decreased from 0.28 to 0.17, reflecting enhanced spatial predictability, while the global Moran's  $I$  increased from 0.61 to 0.73, indicating stronger spatial clustering of similar  $V_s$  values. The  $V_{s30}$  distribution also became more stable, with a mean of approximately 342 m/s and a standard deviation of 82 m/s, which is broadly compatible with NEHRP Site Class D for large portions of the central and southern basin [41–44]. These improvements provide a robust foundation for subsequent development of the  $N$ – $V_s$  regression model, three-dimensional  $V_s$ –ground model, and HAZUS-based hybrid seismic vulnerability assessment.

### 2.3. Development of the Multi-Parameter $N$ – $V_s$ Regression Model

A reliable correlation between standard penetration resistance (SPT- $N$ ) and shear-wave velocity  $V_s$  is essential for extending limited seismic test data to the broader borehole database and for estimating  $V_{s30}$  at locations where direct measurements are unavailable. Conventional empirical  $N$ – $V_s$  equations, such as those proposed by Ohta and Goto, Imai and Tonouchi, and Hasancebi and Ulusay [34–36], were developed primarily from Japanese and Turkish datasets with geological and depositional conditions that differ substantially from those of the Seoul metropolitan alluvial basin. When these equations were applied directly to the present database, the mean absolute error (MAE) between predicted and measured  $V_s$  values typically ranged from about 100 to 150 m/s, and in some cases exceeded 250 m/s, as revealed by a comparative evaluation of 45 published correlations. These levels of misfit are unacceptable for urban-scale seismic-response and liquefaction analyses, and they clearly indicate that a locally calibrated  $N$ – $V_s$  model is required for Seoul.

To construct such a model, a dedicated regression dataset was assembled from approximately 1200 locations where both SPT and seismic tests (downhole, crosshole, PS logging, or MASW) had been performed within a 50 m radius. For each location, depth-matched pairs of  $N$  and  $V_s$  were extracted from the quality-controlled database and annotated with (i) soil type (sand, clay/silt, gravel, weathered rock), (ii) geological age (Cenozoic, Mesozoic, Paleozoic; with particular emphasis on Quaternary deposits), and (iii) depth and overburden stress. This structure allows the regression to capture not only the primary dependence of  $V_s$  on penetration resistance, but also secondary dependencies related to material type and aging effects. Because the selection rule for the calibration dataset is defined explicitly by the co-location of SPT and seismic tests within a 50 m radius and the availability of both measurements, the same subset of the unified database can be reconstructed by other users, which facilitates independent recalibration or regional adaptation of the  $N$ – $V_s$  relationships.

The general form of the regression model adopted in this study is

$$V_s = a N^b (1 + c A_g), \quad (2)$$

where  $N$  is the SPT blow count,  $A_g$  is a geological-age factor, and  $a$ ,  $b$ , and  $c$  are regression coefficients calibrated for each soil group and age class. For Quaternary deposits, the representative best-fit equations are summarized as follows: for all soil types combined,

$$V_s = 360 N^{0.275} (R^2 = 0.68); \quad (3)$$

for clays and silts,

$$V_s = 260 N^{0.317} (R^2 = 0.72); \quad (4)$$

for sands,

$$V_s = 430 N^{0.239} (R^2 = 0.74); \quad (5)$$

and for gravels and weathered rock,

$$V_s = 600 N^{0.178} (R^2 = 0.64). \quad (6)$$

The age-scaling factor between Holocene and Pleistocene units falls in the range of approximately 1.12–1.18, reflecting the systematically higher stiffness of older, more consolidated deposits. These coefficients were obtained from a multi-parameter nonlinear regression in which soil type and age were treated as categorical variables and depth and overburden stress were included as continuous covariates; the resulting coefficients of determination ( $R^2 \approx 0.64$ – $0.76$ ) exceed those of the global empirical equations when evaluated on the same dataset. The functional form of the model and the calibrated coefficients for the main soil groups and age classes are thus fully specified in the text, allowing the  $N$ – $V_s$  relationships to be directly reimplemented for the Seoul basin or adapted to comparable urban alluvial settings.

The characteristics of the multi-parameter correlation are illustrated in Figure 4. Figure 4a plots measured  $V_s$  versus  $N$  for different soil types, together with the fitted regression curves and representative global equations. For a given  $N$ , gravelly and weathered-rock layers exhibit  $V_s$  values exceeding 1000 m/s, whereas clayey and silty soils commonly show  $V_s$  below 300 m/s. The locally calibrated curves capture these contrasts more faithfully than the generic correlations, which tend to overestimate  $V_s$  in fine-grained soils and underestimate it in coarse or partially weathered units. Figure 4b reorganizes the same data by geological age, revealing clear aging effects: Cenozoic deposits have relatively low  $V_s$  for a given  $N$ , whereas Paleozoic units display markedly higher velocities, consistent with long-term cementation and stress history. Again, the age-dependent regression curves track the observed trends more closely than single-equation models that ignore stratigraphic age.

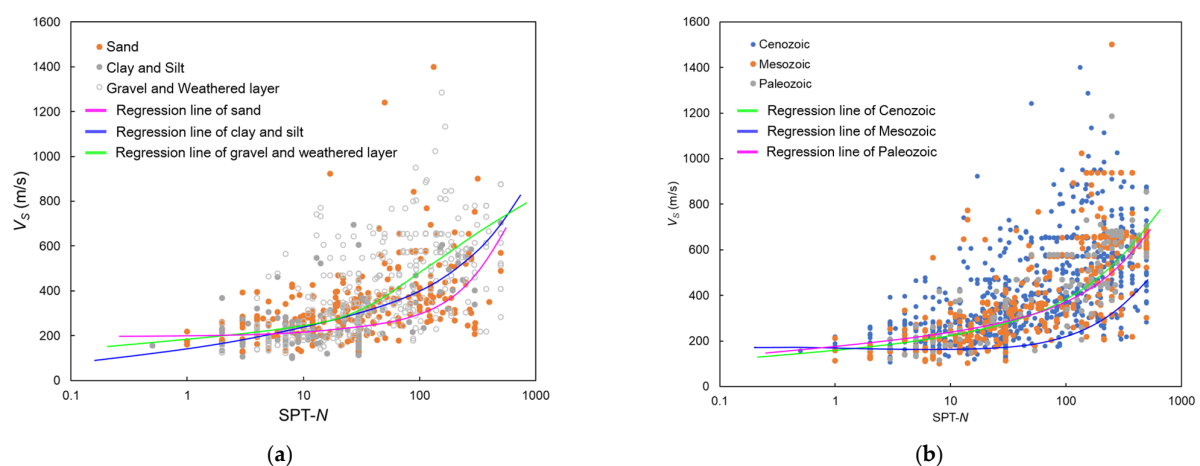


**Figure 4.** Example of multi-parameter  $N$ – $V_s$  regression calibrated for the Seoul geotechnical database: (a) Shear-wave velocity  $V_s$  versus SPT- $N$  by soil type with fitted multi-parameter regression curves; (b) Shear-wave velocity  $V_s$  versus SPT- $N$  by geological age with age-dependent regression curves.

The calibrated  $N$ – $V_s$  relations were then used to supplement missing  $V_s$  values at untested depths and locations in the integrated database. In particular, they played a critical role in interpolating  $V_s$  within soft alluvial sequences and in shallow weathered

rock where seismic testing is sparse. To evaluate predictive performance, regressed  $V_s$  profiles were compared with measured vertical seismic profile (VSP) and downhole logs at representative sites distributed across the three main geotechnical zones of Seoul. Over the 0–30 m depth interval, the mean relative difference between predicted and measured  $V_s$  was generally less than 10%, and the regression reproduced the depth of the engineering bedrock transition (defined approximately by  $V_s > 760$  m/s) with good fidelity. For upper soft layers above the bedrock interface, the coefficient of variation of  $V_s$  was approximately 90 m/s, whereas for deeper rock layers it was on the order of 160 m/s, reflecting the increasing influence of irregular bedrock geometry and test-method limitations at depth. Because the model structure, calibration-dataset definition, and representative coefficients are explicitly documented, the  $N$ – $V_s$  relationships can be readily reproduced or recalibrated by other investigators working with similar regional geotechnical databases.

Figure 5 presents mean  $V_s$  profiles and their  $\pm 1$  standard-deviation envelopes for NEHRP site classes B, C, and D derived from the Seoul database using the multi-parameter  $N$ – $V_s$  model and available seismic measurements [41,43,44]. Class B sites, predominantly in the northern mountainous areas, exhibit  $V_s$  values exceeding 700 m/s over most of the upper 30 m; Class C sites in the central basin show intermediate stiffness with  $V_s$  between about 400 and 600 m/s; and Class D sites in the southern alluvial plain display soft-soil behavior with  $V_s$  commonly in the 200–350 m/s range. The overall shapes of these average profiles are comparable to typical NEHRP reference models, confirming that the Seoul-specific correlations yield physically plausible stratified velocity structures while preserving regional particularities such as thick alluvial deposits along the Han River. The locally calibrated  $N$ – $V_s$  model thus serves two complementary purposes in the proposed framework. First, it reduces prediction errors by 20–40% relative to widely used global correlations when evaluated against the same set of measured  $V_s$  profiles, thereby providing more reliable estimates of  $V_{S30}$  and depth-dependent stiffness for ground-response analysis. Second, by enabling consistent estimation of  $V_s$  at unmeasured locations and depths, it ensures that the three-dimensional  $V_s$ –ground model and subsequent HAZUS-based site classification and amplification modeling can be performed on a spatially continuous, physically coherent basis across the entire Seoul metropolitan area.

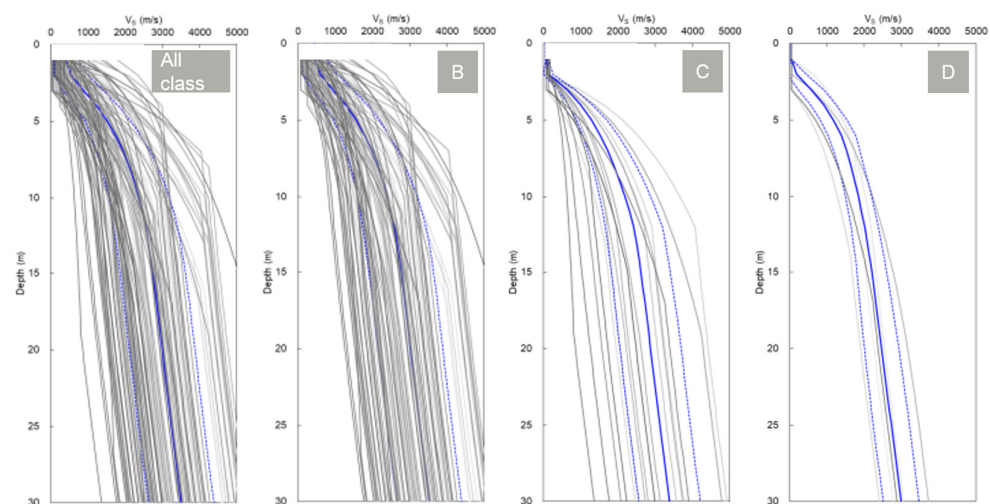


**Figure 5.** Mean shear-wave velocity profiles for NEHRP site classes B–D derived from the Seoul geotechnical database, showing average profiles (solid lines) and  $\pm 1$  standard-deviation envelopes (dashed lines).

In addition to reporting coefficients of determination, the regression performance was examined using standard diagnostic checks to detect potential bias, heteroscedasticity,

and systematic trends associated with depth and soil class. Residuals were defined as  $\varepsilon = V_{S,meas} - V_{S,pred}$ , where  $V_{S,meas}$  denotes the measured shear-wave velocity from down-hole/VSP/MASW logs and  $V_{S,pred}$  denotes the velocity predicted by the locally calibrated multi-parameter model. Three complementary diagnostics were performed: (i) residuals versus fitted values to evaluate heteroscedasticity and global bias, (ii) residuals versus depth to detect depth-dependent trends, and (iii) residual distributions stratified by soil type (and geological age where applicable) to verify that the model does not systematically overestimate or underestimate specific material groups. Quantitative metrics were reported together with  $R^2$ , including the mean absolute error (MAE), root-mean-square error (RMSE), and mean bias error (MBE). These diagnostics confirmed that residuals are approximately centered around zero without pronounced monotonic trends with depth, while moderate heteroscedasticity is observed at higher  $V_s$  levels, which is consistent with the increasing influence of geological heterogeneity and measurement uncertainty in stiff or partially weathered units. The diagnostic plots and metrics are provided in Figure 6 for transparency and reproducibility.

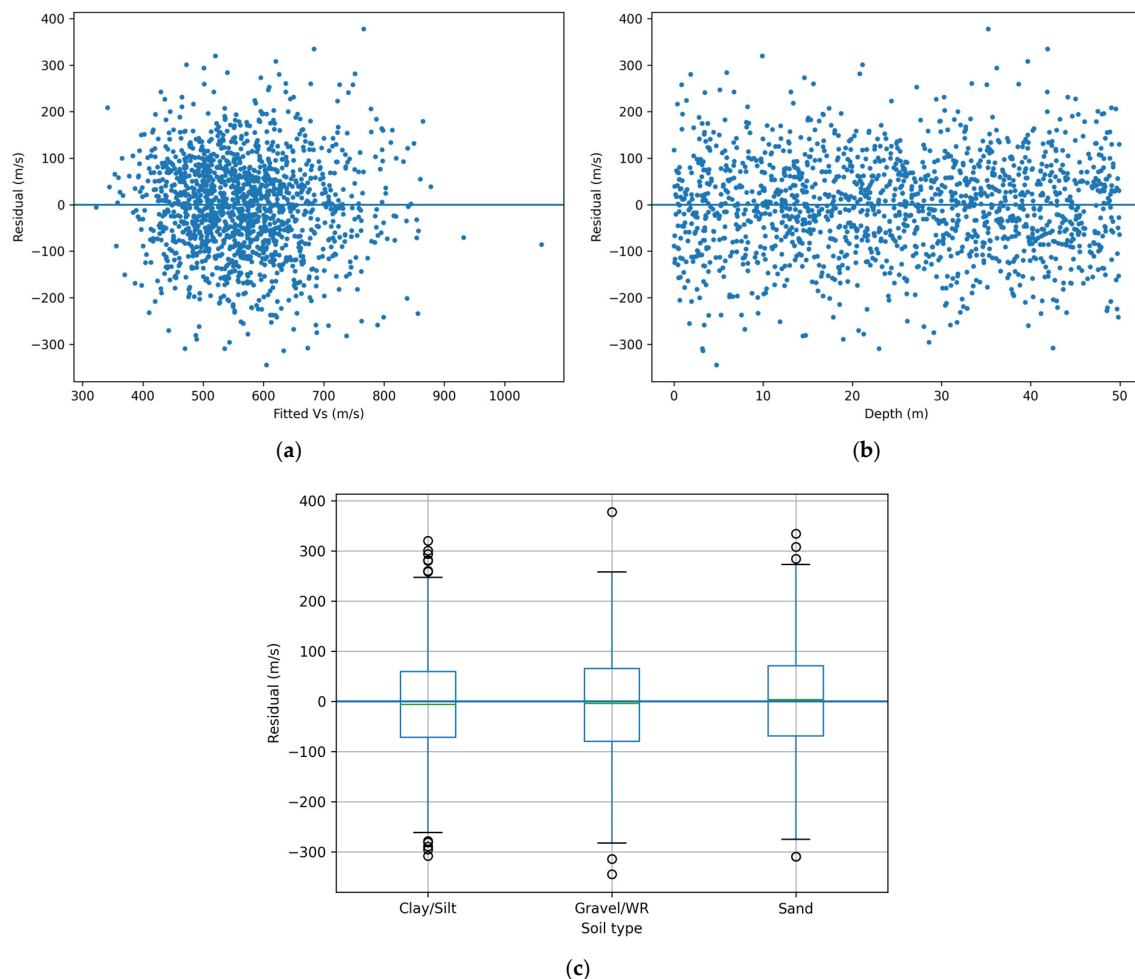


**Figure 6.** Diagnostic evaluation of the locally calibrated  $N$ - $V_s$  regression model: (a) residuals versus fitted value, (b) residuals versus depth, (c) residuals versus soil type (box-plot).

Residual diagnostics for the multi-parameter  $N$ - $V_s$  regression model, illustrating (a) residuals versus fitted shear-wave velocity to assess global bias and heteroscedasticity, (b) residuals versus depth to identify potential depth-dependent trends, and (c) residual distributions stratified by soil type to examine systematic over- or underestimation for specific material groups (Figure 6). Residuals are defined as  $\varepsilon = V_{S,meas} - V_{S,pred}$ . The

diagnostics indicate that residuals are generally centered around zero without pronounced monotonic trends with depth, while moderate heteroscedasticity at higher velocity levels reflects increasing geological heterogeneity and measurement uncertainty in stiff or partially weathered units.

#### 2.4. Spatial Modeling: 3D $V_s$ –Ground Model

The quality-controlled  $V_s$  dataset, consisting of one-dimensional velocity logs from approximately 63,000 boreholes, was transformed into a three-dimensional  $V_s$ –ground model by integrating geostatistical interpolation and stochastic simulation. Each  $V_s$  log was stored with full spatial coordinates ( $X, Y, Z$ ), enabling the spatial continuity of  $V_s$  to be quantified in both the horizontal (plan) and vertical (depth) directions using experimental variograms [39]. Directional variograms were computed separately for the horizontal plane and the vertical axis to capture anisotropy associated with laterally variable basin deposits and vertically stratified profiles. The experimental semivariances were fitted with standard covariance models following Journel and Huijbregts [39] and Deutsch and Journel [40], and the fitted variogram parameters were consistently adopted for both ordinary kriging and sequential Gaussian simulation (SGS), ensuring methodological consistency and reproducibility.

For transparency, the fitted directional variogram models and parameters are summarized as follows (Figure 7). In the horizontal direction, a spherical model was adopted with nugget = 1500 ( $\text{m/s}^2$ ), sill = 12,000 ( $\text{m/s}^2$ ), and range  $a_h = 520$  m. In the vertical direction, an exponential model was used with nugget = 2000 ( $\text{m/s}^2$ ), sill = 10,500 ( $\text{m/s}^2$ ), and e-folding range parameter  $a_v = 20$  m (practical range  $\approx 3a_v$ ). The resulting anisotropy ratio ( $a_h/a_v$ ) is approximately 26, indicating substantially shorter correlation lengths in the vertical direction, which is consistent with layered alluvial deposits overlying bedrock and laterally segmented facies in the Seoul basin.

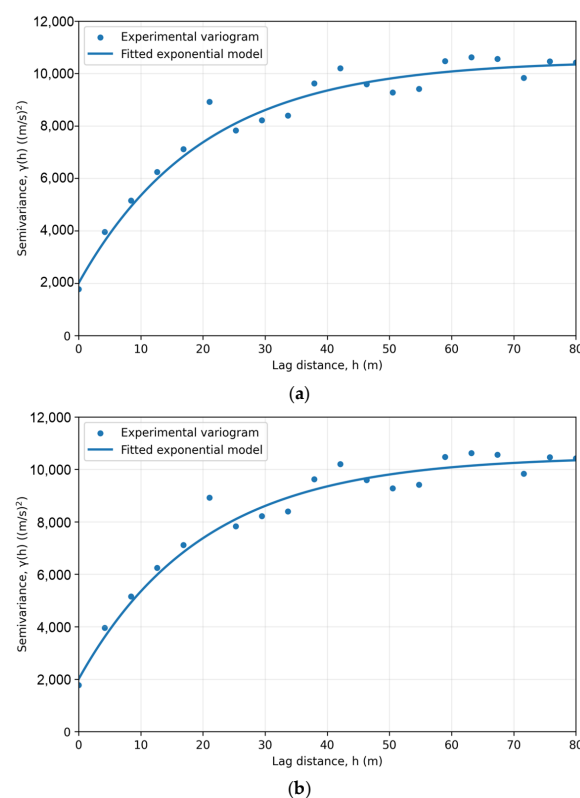


**Figure 7.** Directional experimental variograms and fitted models for shear-wave velocity ( $V_s$ ) in Seoul: (a) horizontal variogram, (b) vertical variogram.

Using the fitted variogram models, ordinary kriging was applied to estimate  $V_s$  on a regular three-dimensional grid [39,40]. The grid resolution was set to 10 m  $\times$  10 m in plan and 1 m in depth, yielding approximately 3.8 million voxels across the study area. At each voxel, kriging provides both an estimated  $V_s$  value and the associated kriging variance. Model performance was evaluated through cross-validation (leave-one-out or block-based validation, depending on data density), reporting standard metrics (e.g., mean error and RMSE) to quantify predictive accuracy and bias. The kriging-based model provides a high-resolution mean field that forms the deterministic baseline for deriving  $V_s$ -based site-response parameters.

While kriging estimates the local mean structure, uncertainty characterization is required to diagnose epistemic uncertainty arising from heterogeneous depositional environments and uneven data density. Therefore, sequential Gaussian simulation (SGS) was applied using the same variogram parameters and conditioning data [39,40]. One hundred conditional realizations of the  $V_s$  field were generated on the same 10 m  $\times$  10 m  $\times$  1 m grid, each honoring the measured  $V_s$  logs and reproducing the global mean and variance implied by the observed dataset and fitted variograms. From the ensemble, voxel-wise uncertainty metrics were computed as the standard deviation and coefficient of variation (COV) of  $V_s$ :

$$\sigma_{V_s}(x) = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (V_{s_j}(x) - \bar{V}_s(x))^2} \quad (7)$$

$$\text{COV}_{V_s}(x) = \frac{\sigma_{V_s}(x)}{\bar{V}_s(x)}, \quad (8)$$

where  $V_{s_j}(x)$  is the simulated  $V_s$  at location  $x$  in the  $j$ -th realization, mean  $(\bar{V}_s)(x)$  is the ensemble mean, and  $N$  is the number of realizations.

To support subsequent seismic vulnerability analysis, key  $V_s$ -based site-response parameters were computed consistently for each SGS realization and then summarized as ensemble statistics. Specifically, for each realization, depth to engineering bedrock ( $H$ ),  $V_{s30}$  depth-averaged soil shear-wave velocity ( $V_{s,soil}$ ), and fundamental site period ( $T_G$ ) were derived on the same grid. The corresponding COV maps were then computed using the ensemble mean and standard deviation of each parameter, providing spatially explicit measures of uncertainty for  $H$ ,  $V_{s30}$ ,  $V_{s,soil}$ , and  $T_G$ . For interpretability, the COV values are classified into four uncertainty levels: low uncertainty ( $\text{COV} < 0.15$ ), moderate uncertainty ( $0.15 \leq \text{COV} < 0.30$ ), high uncertainty ( $0.30 \leq \text{COV} < 0.45$ ), and very high uncertainty ( $\text{COV} \geq 0.45$ ). These ranges reflect typical behavior observed in dense and lithologically homogeneous bedrock domains (low COV), transitional zones between bedrock and soil layers (moderate COV), thick alluvial basins and river-adjacent heterogeneous facies (high COV), and reclaimed or data-sparse lowlands (very high COV).

On top of the continuous  $V_s$  field, an engineering-layer classification was constructed to assign each voxel to representative ground material classes (e.g., fill, alluvial soil, weathered soil, weathered rock, soft rock, hard rock). This classification integrated borehole stratigraphy, laboratory results, and interpreted geologic units within the 3D grid, allowing a physically interpretable linkage between mapped geologic domains and  $V_s$ -based stiffness patterns. The resulting model indicates that soft alluvium ( $V_s \approx 150$ – $250$  m/s) forms a continuous corridor through the Gangnam–Songpa–Yeouido axis, consistent with the main Han River alluvial plain, whereas shallow, high-velocity bedrock ( $V_s > 700$  m/s) dominates beneath the Bukhansan–Dobong mountainous belt. When vertically averaged over the upper 30 m, the  $V_{s30}$  patterns are consistent with NEHRP/HAZUS site classes: northern mountainous areas correspond primarily to Site Class B, the central urban plateau to Class C, and southern alluvial basins to Classes D–E [43,44]. These patterns agree with

independent proxies based on topographic slope and regional geology [42], supporting the validity of the constructed model.

Overall, the proposed 3D  $V_s$ -ground model provides a geostatistically consistent representation of subsurface stiffness that captures anisotropy and heterogeneity, enables  $V_s$  prediction at unsampled locations and depths, and supplies a common geospatial platform for HAZUS-compatible site-response parameterization. In this study, the SGS-derived uncertainty metrics are used to qualify the interpretation of site classification and vulnerability results and to identify locations where additional investigations would most effectively reduce epistemic uncertainty, whereas deterministic vulnerability maps are presented using the kriging-based mean fields for consistency with standard HAZUS-based regional planning practice.

### 2.5. Building Inventory Modeling for HAZUS Integration

For HAZUS-based seismic vulnerability assessment, a city-wide building inventory was constructed for Seoul using cadastral GIS and administrative building registries. The primary sources were the Seoul Open Data Plaza, the MOLIT public data portal, the “Seumter” building administration system (OpenAPI), and the national real-estate information service. These datasets provide both polygon footprints and centroid point geometries, each linked via unique identifiers (e.g., building registry primary keys) to attribute tables containing basic, structural, and area information. After cleaning and harmonization, the compiled dataset contains on the order of 650,000–700,000 buildings within the administrative boundary of Seoul, all projected in the Korea 2000/UTM-K coordinate system (EPSG:5179), which is consistent with the geotechnical database. Figure 8 illustrates the spatial distribution of building centroids, superimposed on the Han River and major fault traces, highlighting the very high building density along the central urban corridor and southern commercial districts.

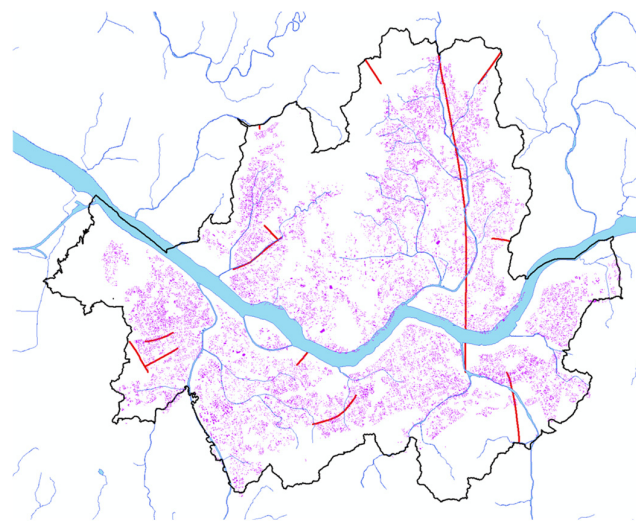


**Figure 8.** Spatial distribution of buildings in Seoul based on the cadastral GIS.

Key registry attributes include building name and address, total and per-floor areas, number of stories, primary use (residential, commercial, educational, office, industrial, etc.), structural type (reinforced concrete, steel, steel-reinforced concrete, masonry), year of completion, and parcel information. Additional tables provide exclusive and common floor areas, unit-level area distributions, and limited information on foundation type or special structural systems. These attributes were systematically mapped to the HAZUS building inventory schema, which requires occupancy class, general and specific building type, number of stories, year built, and replacement cost, among others [21,28,44]. Table 1

summarizes the correspondence between core Seoul building registry fields and the HAZUS inventory fields, and identifies items that require supplementation or interpretation. For example, the “primary use” code of the registry can be mapped almost directly to HAZUS occupancy classes (e.g., RES, COM, EDU, OFC), whereas mixed-use buildings require additional processing to distinguish primary versus secondary occupancy and, where possible, floor-by-floor use fractions.

**Table 1.** Key HAZUS inventory fields and correspondence with Seoul building registry attributes.

| Seoul Building Registry Attribute                               | Corresponding HAZUS Inventory Field(s)                                            | Additional Information/Enhancement Required                                                                                                                                                                           |
|-----------------------------------------------------------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Primary use code (building use/occupancy)                       | Occupancy Class                                                                   | Primary use code is almost directly mappable. For mixed-use buildings, primary vs. secondary uses and floor-by-floor occupancy information are needed.                                                                |
| Structural type (e.g., reinforced concrete, steel, masonry)     | General Building Type, Specific Building Type                                     | Detailed structural information such as roof type, wall material, and lateral-load resisting system may be missing and should be supplemented where possible.                                                         |
| Number of stories                                               | Number of Stories                                                                 | It must be verified whether basement levels are included or excluded; if necessary, separate fields for above-ground and basement stories should be defined.                                                          |
| Total floor area/building footprint area                        | Area/Footprint/Gross Floor Area                                                   | Differences between gross floor area used for floor-area ratio calculations and the actual exposed floor area for damage/loss estimation should be corrected.                                                         |
| Approval date/year of completion                                | Year Built, Design Level                                                          | A rule set is required to map year built to Korean seismic code eras in order to derive the seismic design level (Pre-/Low-/Moderate-/High-code, etc.).                                                               |
| Special structure flag, foundation type, seismic capacity notes | Foundation Type, Structural System Attributes, Seismic Design Level               | Even when foundation type or ground condition fields exist in the registry, they are often incomplete; additional geotechnical/structural information is needed to refine these attributes.                           |
| Exclusive/common floor area and floor-level unit information    | (Not an explicit core HAZUS field; supports exposure and occupancy distributions) | Floor-by-floor and unit-level area information is advantageous for detailed exposure, casualty, and economic loss estimation and can be used to derive weighting factors when HAZUS aggregates at the building level. |

Structural-type information in the registry (e.g., reinforced-concrete frame, steel frame, masonry, light-gauge steel) was mapped to the HAZUS general and specific building types (e.g., C1/C2 for reinforced-concrete moment or shear wall systems, S1–S3 for steel moment/braced frames, URM for unreinforced masonry). However, many records lack detailed attributes such as roof type, wall material, or explicit description of the lateral load-resisting system. Where possible, these gaps were reduced by using typical Korean design practices and auxiliary data (e.g., building height and footprint shape) to infer probable HAZUS types, while preserving explicit registry information as priority. For the number-of-stories field, the registry was examined to confirm whether basement levels were included; where ambiguity existed, separate fields for above-ground and basement stories

were created so that structural period and occupancy exposure could be modeled more consistently. Differences between gross floor area used for floor-area-ratio calculations and the actual exposed floor area relevant to damage and loss estimation were also reviewed and, where feasible, adjusted.

The year of completion (approval date) was used to assign each building to a Korean seismic design era and, by extension, to a HAZUS-consistent design-level category [43–45]. Four broad eras were defined: (i) Pre-code, for buildings constructed before the introduction of modern seismic provisions (pre-1988); (ii) Low-code, for buildings designed under initial seismic requirements (approximately 1988–2005); (iii) Moderate-code, reflecting the period of significantly strengthened provisions (approximately 2005–2009); and (iv) High-code, for buildings constructed after the most recent major code revisions (post-2016). A rule set was implemented to map “year built” to these eras and to encode the design level as an attribute that can be directly used to adjust fragility parameters in HAZUS. This mapping explicitly recognizes that a large portion of Seoul’s building stock predates modern seismic codes, implying substantial structural vulnerability in many districts.

Exclusive and common floor area tables, although not explicitly required by the core HAZUS inventory, provide valuable information for exposure and casualty modeling. Floor-by-floor area distributions were used to derive weighting factors for the vertical distribution of occupants and economic value within each building, which is particularly important in high-rise mixed-use complexes where residential, office, and commercial functions coexist. For special structures (e.g., large-span facilities, infrastructure-related buildings) and records that include foundation type or seismic-capacity notes, additional fields were created to store foundation type, structural system attributes, and qualitative seismic-capacity indicators. In many cases these fields are incomplete, but even partial information can improve the assignment of appropriate HAZUS building classes and damage functions when combined with geotechnical information on local site conditions.

Spatial analysis of the resulting inventory reveals clear regional patterns. High-rise reinforced-concrete and steel buildings, often with mixed commercial and office occupancy, are concentrated in Gangnam, Yeouido, and the central business districts of Jongno and Jung, forming dense clusters visible in Figure 8. In contrast, outer districts such as Eunpyeong, Dobong, Nowon, and parts of Guro and Geumcheon exhibit a higher proportion of low-rise masonry, light-gauge steel, and older non-ductile structures. When overlain with the 3D Vs–ground model and mapped site classes, these spatial contrasts indicate that seismically vulnerable building stocks often coincide with unfavorable ground conditions, especially along thick alluvial deposits and reclaimed land near the Han River.

In summary, the Seoul building inventory has been restructured into a HAZUS-compatible database through systematic field mapping and enhancement (Table 1). By combining structural type, height, use, year built, and detailed area information with site-specific geotechnical parameters (e.g.,  $V_{s30}$ , fundamental site period  $T_G$ , and bedrock depth), the inventory provides a robust input platform for probabilistic fragility assessment, loss estimation, and scenario-based earthquake risk mapping at the urban scale. Figure 6 and Table 1 together illustrate how the cadastral GIS and registry attributes were transformed into a standardized, internationally comparable framework suitable for the hybrid seismic vulnerability analyses presented in subsequent sections.

In this study, the building fragility functions are primarily adopted from the standard HAZUS framework, which provides a widely validated set of damage-state probability models for different structural systems, height classes, and design levels. Localization of the vulnerability assessment is therefore not achieved through direct recalibration of the fragility curves themselves, but through a Seoul-specific adaptation of both exposure characterization and seismic demand.

Specifically, Korean building inventory data are systematically mapped to the HAZUS structural taxonomy by explicitly accounting for structural type, number of stories, occupancy, and construction era relative to the evolution of Korean seismic design provisions. In parallel, the seismic demand input to the fragility functions is modified using site-specific response spectra derived from the locally calibrated 3D  $V_s$ -ground model, rather than relying on generic site-class amplification factors.

As a result, although the functional form and median parameters of the fragility curves follow the HAZUS defaults, the effective damage probabilities reflect localized ground-motion amplification, basin effects, and resonance characteristics that are specific to the Seoul metropolitan area. This approach yields a demand-adjusted, Seoul-adapted vulnerability model that captures the combined influence of local subsurface conditions and building stock characteristics, while maintaining consistency with internationally established fragility formulations.

The seismic vulnerability assessment in this study follows the HAZUS building taxonomy and associated fragility formulations, in which damage-state exceedance probabilities are defined as lognormal functions of spectral acceleration. In the present implementation, the fragility parameter sets (median and lognormal dispersion) are adopted from the standard HAZUS database and applied without re-calibration using Korea-specific post-earthquake damage observations. Localization is implemented through (i) exposure characterization and classification, whereby the Seoul building inventory is mapped to HAZUS structural classes and to Korean seismic design-era categories that are subsequently linked to HAZUS code levels, and (ii) seismic demand modification, whereby site-specific response spectra and amplification factors are derived from the city-scale 3D  $V_s$ -ground model and used to define the effective spectral acceleration input to the fragility functions. Accordingly, the framework should be interpreted as “Seoul-adapted” in terms of exposure and demand, while the vulnerability functions themselves remain HAZUS-based and constitute an important limitation of the current study.

## 2.6. HAZUS-Based Seoul-Adapted Seismic Vulnerability Modeling

The integrated geotechnical and building databases were coupled within a HAZUS-compatible framework to develop a Seoul-adapted seismic vulnerability and loss model. Each building is characterized by its structural system, height class, seismic design level, and occupancy type mapped to the HAZUS building taxonomy, while local site conditions are represented by  $V_{s30}$ , depth to engineering bedrock  $H$ , and the fundamental site period  $T_G$ . Although the overall structure follows the HAZUS Earthquake Model, the parameters and mapping rules were calibrated to reflect Korean seismic design provisions and the spatial characteristics of Seoul.

Design-era mapping to HAZUS design levels (consolidated). For fragility selection at inventory scale, each building is assigned to a HAZUS design level based on the year of construction using explicit thresholds aligned with major milestones of Korean seismic design provisions:

- Pre-1988 → Pre-code;
- 1988–2004 → Low-code;
- 2005–2016 → Moderate-code;
- 2017–present → High-code (or Very-high-code where additional performance evaluation or special-importance designation is documented).

This mapping is intended as a pragmatic harmonization for metropolitan-scale screening rather than a claim of strict code equivalence, and the associated assumptions are explicitly acknowledged as a limitation where building-specific detailing information is unavailable in the registry.

Site classification was carried out by overlaying the 3D  $V_s$ -ground model with the building inventory. For each building centroid,  $V_{s30}$  and the depth to engineering bedrock were extracted from the 10 m  $\times$  10 m  $\times$  1 m grid and converted to NEHRP-consistent site classes. Northern mountainous areas with shallow, stiff bedrock were predominantly classified as Site Class B, central Seoul as Site Class C, and southern alluvial basins as Site Classes D and E, illustrating strong intra-urban variability in site amplification.

The fundamental site period  $T_G$  was estimated using an empirically derived regression relationship obtained from the integrated geotechnical database:

$$T_G = 0.016 H + 0.12, \quad (9)$$

where  $H$  is the depth to engineering bedrock in meters ( $R^2 \approx 0.78$ ). This relation reflects the tendency for deeper sedimentary basins to exhibit longer fundamental periods and was used together with  $V_{s30}$  to refine period-dependent site-response parameters.

For each seismic scenario, rock-outcrop response spectra corresponding to 500-year, 1000-year, and 2400-year return periods were taken from Korean seismic zonation and design spectra. These spectra were modified using NEHRP-consistent amplification factors  $F_a$  and  $F_v$  as functions of  $V_{s30}$  and  $T_G$ , yielding site-specific response spectra  $S_a(T)$ , which were used as intensity measures in the fragility analysis.

Building attributes from the Seoul registry were mapped to the HAZUS structural taxonomy following explicit classification rules (Table 2). Structural types, height classes, and seismic design levels were inferred from construction year and code evolution, allowing each building to be assigned a HAZUS class and associated lognormal fragility functions. For a given spectral acceleration  $S_a$ , the probability of exceeding damage state  $k$  is expressed as

$$P[DS \geq k | S_a] = \Phi \left( \frac{\ln S_a - \ln \theta_k}{\beta_k} \right), \quad (10)$$

where  $\theta_k$  and  $\beta_k$  are the median capacity and dispersion for damage state  $k$ , respectively.

**Table 2.** Building classification rules for mapping the Seoul building registry to HAZUS building classes and design levels.

| Category                           | Subcategory/Class                   | Description/Mapping Rule                                                                                                                   |
|------------------------------------|-------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|
| Structural type                    | RC $\rightarrow$ C1, C2, C3         | Reinforced concrete buildings mapped to HAZUS C1 (moment frame) or C2 (shear wall); if infill walls govern lateral response, mapped to C3. |
|                                    | Steel $\rightarrow$ S1, S2, S3      | Steel buildings mapped to HAZUS S1 (moment frame), S2 (braced frame), or S3 (light-frame/panelized).                                       |
|                                    | Masonry $\rightarrow$ URM, RM1, RM2 | Masonry buildings classified as URM (unreinforced masonry) or RM1/RM2 (reinforced masonry).                                                |
|                                    | Timber $\rightarrow$ W1, W2         | Timber/wood-frame buildings mapped to HAZUS W1/W2.                                                                                         |
| Height class                       | Precast $\rightarrow$ PC1, PC2      | Precast concrete systems mapped to PC1/PC2.                                                                                                |
|                                    | Low-rise                            | 1–3 stories.                                                                                                                               |
|                                    | Mid-rise                            | 4–7 stories.                                                                                                                               |
|                                    | High-rise                           | $\geq 8$ stories.                                                                                                                          |
| Design level<br>(seismic code era) | Pre-code                            | Constructed before introduction of modern Korean seismic code; mapped to “Pre-code”.                                                       |
|                                    | Low-code                            | Constructed in early code era; mapped to “Low-code”.                                                                                       |
|                                    | Moderate-code                       | Constructed after strengthening of code provisions; mapped to “Moderate-code”.                                                             |
|                                    | High-code/<br>Very-high-code (VH)   | Constructed under the most recent, stringent seismic design standards.                                                                     |
| Mixed occupancy                    | Primary occupancy                   | Primary occupancy determined by dominant floor-area share; secondary occupancies treated as auxiliary if needed.                           |

Evaluating these fragility functions for each building yields damage-state probability vectors, from which expected damage indices, economic loss ratios, and casualty estimates are derived. These building-level results are subsequently aggregated to spatial analysis units (grids or administrative districts) to produce scenario-dependent maps of damage, loss, and casualty indicators, which serve as the basis for composite risk assessment described in the following section.

## 2.7. Composite Risk Index (CRI) for Integrated Seismic Risk Zoning

The Composite Risk Index (CRI) constitutes a key output of this study, as it integrates multiple dimensions of seismic impact into a single metric suitable for spatial comparison and risk zoning. The CRI is defined as a weighted linear combination of four component indices representing building damage, economic loss, casualties, and lifeline disruption. While the CRI framework allows the weighting factors to be adjusted to reflect local policy priorities, it is essential to explicitly document both the index computation procedure and the specific weights adopted in this study to ensure transparency and reproducibility of the resulting risk maps.

The CRI is evaluated at a predefined spatial analysis unit  $u$ , corresponding either to a regular grid cell or to an administrative district. For each unit  $u$ , component indices are first computed from building-level or asset-level results and subsequently combined through normalization and weighting. The CRI is expressed as

$$\text{CRI}(u) = \alpha I_{\text{damage}}^*(u) + \beta I_{\text{loss}}^*(u) + \gamma I_{\text{casualty}}^*(u) + \delta I_{\text{lifeline}}^*(u), \alpha + \beta + \gamma + \delta = 1, \quad (11)$$

where  $I_i^*(u)$  denotes the normalized index for component  $i$ , and  $\alpha, \beta, \gamma$ , and  $\delta$  are non-negative weighting factors.

Each component index is derived following a consistent aggregation–normalization procedure. For building-based components (damage, loss, and casualty), building-level quantities are aggregated to the spatial unit  $u$  using exposure-weighted averaging. Let  $b$  denote a building located within unit  $u$ ,  $x_{i,b}$  the building-level metric associated with component  $i$ , and  $w_b$  an exposure weight. In this study, gross floor area was adopted as the exposure weight because it is uniformly available in the Seoul building registry and provides a reasonable proxy for both replacement cost and occupancy. The aggregated, unnormalized component index is defined as

$$I_i(u) = \frac{\sum_{b \in \mathcal{B}(u)} w_b x_{i,b}}{\sum_{b \in \mathcal{B}(u)} w_b}, i \in \{\text{damage, loss, casualty}\}. \quad (12)$$

The building-level damage metric  $x_{\text{damage},b}$  was defined as the expected Damage State Index (DSI), computed from the HAZUS damage-state probabilities:

$$x_{\text{damage},b} = \text{DSI}_b = \sum_{k=0}^4 s_k P(\text{DS} = k)_b, s_k = \frac{k}{4}, \quad (13)$$

where  $k$  denotes the damage state (none, slight, moderate, extensive, complete),  $P(\text{DS} = k)_b$  is the probability of damage state  $k$  for building  $b$ , and  $s_k$  is a normalized severity score ranging from 0 to 1. Economic loss metrics  $x_{\text{loss},b}$  were defined as expected loss ratios combining structural, non-structural, and contents losses, while casualty metrics  $x_{\text{casualty},b}$  were defined as expected casualty rates under the selected occupancy scenario. Lifeline-related indices were computed using an analogous aggregation scheme based on asset segments, with exposure weights defined by segment length or serviced demand, depending on data availability.

Because the aggregated component indices  $I_i(u)$  differ in physical units and numerical ranges, each index was normalized prior to combination. Min–max normalization was applied across all spatial units within a given seismic scenario:

$$I_i^*(u) = \frac{I_i(u) - \min_{v \in \mathcal{U}} I_i(v)}{\max_{v \in \mathcal{U}} I_i(v) - \min_{v \in \mathcal{U}} I_i(v) + \varepsilon}, \quad (14)$$

where  $\mathcal{U}$  denotes the set of all analysis units and  $\varepsilon$  is a small constant introduced for numerical stability. This normalization yields dimensionless indices in the range  $[0, 1]$ , emphasizing the relative spatial contrast of each risk component across Seoul.

For the CRI maps presented in Figure 9, a baseline equal-weight scheme

$$(\alpha, \beta, \gamma, \delta) = (0.25, 0.25, 0.25, 0.25) \quad (15)$$

was adopted. This choice reflects a neutral assumption in which structural damage, economic loss, human impact, and lifeline disruption are treated as equally important dimensions of seismic risk. The equal-weight configuration was selected to avoid introducing implicit policy bias into the primary results and to provide a clear reference case for comparison.

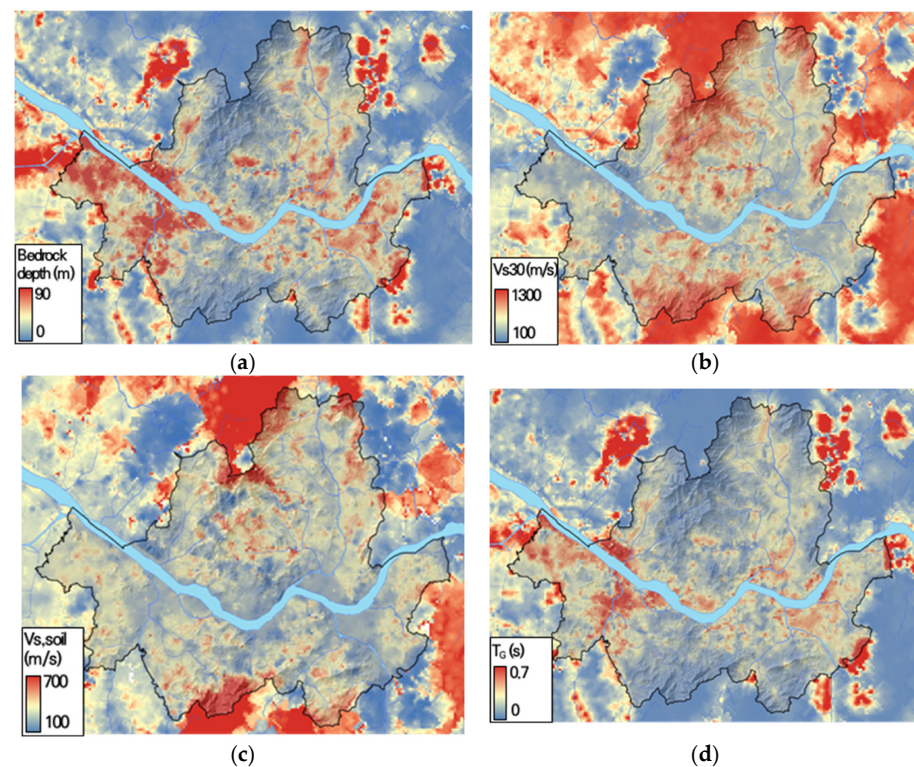


**Figure 9.** Seismic zonation of site response parameters: (a) bedrock depth; (b)  $V_{s30}$ ; (c)  $V_{s,soil}$ ; (d)  $T_c$ .

It is acknowledged that the spatial pattern of high-risk zones can be sensitive to the choice of weighting factors, particularly when different policy objectives prioritize economic loss, life safety, or infrastructure resilience. To explicitly address this sensitivity, additional CRI maps were generated using alternative weighting combinations representing different prioritization scenarios. The resulting changes in the spatial distribution and ranking of high-risk zones were quantified using overlap ratios and rank-correlation measures relative to the baseline case.

Table 2 summarizes the explicit classification rules used to map the Seoul building registry to the HAZUS building taxonomy and corresponding seismic design levels. This table plays a critical role in the overall framework, as it defines the structural and regulatory attributes that govern the selection of fragility functions and, consequently, all subsequent damage and loss estimates.

The classification is performed along four primary dimensions: structural system, height class, seismic design level, and occupancy. Structural types recorded in the national registry are first grouped according to their dominant lateral-load-resisting mechanisms and mapped to the closest HAZUS structural classes (e.g., reinforced-concrete moment frames to C1, shear-wall-dominated systems to C2). Where infill walls are expected to contribute significantly to lateral stiffness and strength, reinforced-concrete buildings are conservatively mapped to C3. Similar rule-based mappings are applied to steel, masonry, timber, and precast concrete systems to ensure consistency with the HAZUS taxonomy.

Height classes are defined using the number of above-ground stories, following the standard HAZUS thresholds for low-rise, mid-rise, and high-rise buildings. The seismic design level is inferred from the year of construction in relation to the historical evolution of Korean seismic design provisions, distinguishing pre-code, low-code, moderate-code, and high-code (or very-high-code) buildings. Mixed-occupancy buildings are classified based on the dominant floor-area share, while secondary occupancies are retained as auxiliary attributes where detailed loss or casualty modeling requires them.

By formalizing these mapping rules in Table 2, the framework ensures that the assignment of HAZUS building classes and design levels is transparent and reproducible. At the same time, it is acknowledged that this rule-based classification introduces epistemic uncertainty, particularly for older buildings or cases where detailed structural information is unavailable. This uncertainty propagates through the fragility-based damage calculations and ultimately influences the composite risk indices, underscoring the importance of explicitly documenting the classification assumptions. Approximately 95% of the reinforced-concrete buildings required inference-based classification of the dominant lateral-load-resisting system due to the absence of explicit structural system information in the registry.

To clarify how heterogeneous datasets are systematically transformed into quantitative damage and risk indicators, the main steps of the Seoul-adapted HAZUS-based seismic risk assessment are summarized in Table 3. Rather than representing independent analyses, these steps form a sequential workflow in which the outputs of each stage constitute the inputs for the next.

The process begins with the conversion of cadastral and building-record information into standardized HAZUS building classes and design levels (Step 1), using the classification rules defined in Table 2. This step establishes the structural and regulatory attributes required for fragility assignment. In Step 2, class-specific fragility parameters are assigned to each building, and conditional damage probabilities are evaluated as functions of site-specific spectral acceleration. Step 3 uses these fragility relationships to compute building-level damage-state probability vectors for the four standard HAZUS damage states.

Spatial aggregation is then performed in Step 4, where building-level damage probabilities are combined within predefined spatial analysis units, such as regular grid cells or administrative districts. This aggregation uses exposure-based weights (e.g., floor area or asset value) to ensure that larger or more heavily used buildings contribute proportionally to the aggregated indicators. Finally, in Step 5, aggregated damage probabilities are converted into derived indicators, such as the expected Damage State Index (DSI), economic loss ratios, and inputs for casualty estimation.

**Table 3.** Stepwise workflow of the HAZUS-based seismic risk assessment framework adapted for Seoul.

| Step | Process                                   | Main Operations                                                                                                                                                              | Outputs                                                                                     |
|------|-------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|
| 1    | Cadastral record → HAZUS class assignment | Use cadastral/building records to determine structural type, height class, and design level; map each building to a HAZUS building class (e.g., C1, S1, URM) and code level. | HAZUS building class and design level for each building.                                    |
| 2    | Fragility parameter assignment            | For each HAZUS class, assign fragility parameters; compute conditional probabilities for given spectral acceleration.                                                        | Building-level fragility functions and damage probabilities as a function of ( $S_a$ ).     |
| 3    | Damage-state probability computation      | Evaluate damage-state probabilities for each building: Slight, Moderate, Extensive, Complete using the assigned fragility curves.                                            | Building-level damage-state probability vectors.                                            |
| 4    | Spatial aggregation                       | Aggregate building-level probabilities to spatial units (e.g., 100 m or 500 m grid cells, dong/gu districts) using weights based on floor area, asset value, or population.  | Grid-based or administrative-unit damage probability maps.                                  |
| 5    | Risk and loss estimation                  | Derive composite indicators such as expected Damage State Index (DSI), economic loss ratios, and inputs to casualty and shelter-demand modules.                              | Seismic damage probability maps, DSI maps, and inputs to economic-loss and casualty models. |

Table 3 therefore provides the procedural backbone for the computation of the damage-, loss-, and casualty-related indices that later enter the Composite Risk Index (CRI). By explicitly organizing these steps, the table makes clear that the CRI is not an ad hoc metric, but the final outcome of a structured and traceable assessment chain.

While Table 3 focuses on building damage and loss estimation, Table 4 extends the framework to a multi-criteria evaluation of seismic risk, explicitly linking individual analysis components to the composite risk formulation. The steps summarized in Table 4 correspond directly to the four component indices used in the CRI definition.

Building damage analysis (Step 1) quantifies the probabilities of different structural damage states and forms the basis of the damage index used in the CRI. Lifeline functional-loss analysis (Step 2) evaluates the vulnerability of critical infrastructure systems using repair-rate models and fragility relationships, producing an infrastructure-related index that captures network-level functionality loss. Economic loss estimation (Step 3) integrates structural, non-structural, contents, and business-interruption losses, yielding an economic-loss index expressed in relative or normalized form. Casualty estimation (Step 4) combines damage-state probabilities with occupancy information and time-of-day scenarios to derive casualty indices representing potential human impact.

In Step 5, these heterogeneous indicators are normalized to a common scale and aggregated using a weighted linear combination, consistent with the CRI definition given in Equation (9). The final step (Step 6) maps the resulting indices onto spatial units and visualizes them as choropleth or hotspot maps, typically emphasizing upper quantiles (e.g., the top 10% highest-risk areas) for prioritization purposes.

**Table 4.** Workflow for multi-criteria evaluation of seismic damage, loss, and composite risk in Seoul.

| Step | Analysis Component                | Main Operation                                                                                                                                                                                                                                              |
|------|-----------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1    | Building damage analysis          | Apply seismic fragility curves to each building class to estimate the probability of Slight, Moderate, Extensive, and Complete damage states.                                                                                                               |
| 2    | Lifeline functional-loss analysis | Use lifeline repair rates (RR) and corresponding fragility relationships to evaluate the probability of functional loss for critical lifelines and infrastructure systems.                                                                                  |
| 3    | Economic loss estimation          | Combine structural, non-structural, contents, and business-interruption components to compute direct and indirect economic losses.                                                                                                                          |
| 4    | Casualty estimation               | Incorporate occupancy information and time-of-day scenarios to estimate casualties for each damage state, typically classified into Casualty Levels 1–4.                                                                                                    |
| 5    | Composite risk index calculation  | Normalize multiple indicators (0–1), then aggregate them using a weighted combination, e.g., Composite Risk Index = ( $\alpha$ )(damage index) + ( $\beta$ )(economic-loss index) + ( $\gamma$ )(casualty index) + ( $\delta$ )(infrastructure-loss index). |
| 6    | Result mapping and visualization  | Project indices and probabilities onto standard grids (e.g., 100 m or 500 m) or administrative units and create choropleth and hotspot maps, focusing on the top-risk quantiles (e.g., top 10%).                                                            |

By structuring the analysis in this manner, Table 4 clarifies the conceptual meaning of each term in the CRI formulation and demonstrates how the composite index synthesizes multiple dimensions of seismic impact. Importantly, it also highlights why the spatial pattern of high-risk zones can be sensitive to the choice of weighting factors: each CRI component originates from a distinct analytical pathway, reflects a different aspect of risk, and exhibits its own spatial variability.

### 3. Results

#### 3.1. Spatial Zonation of Seismic Site Effects

For the spatial characterization of site effects, the three-dimensional  $V_S$ -ground model described in Section 2 was aggregated onto a regular grid system that is compatible with the national standard grid proposed by the National Geographic Information Institute of Korea (NGII; notice No. 2022-2403). Seoul was partitioned into equal-area grid cells (primarily 100 m resolution for analysis, with 50 m and 250 m grids generated for specific visualization and aggregation purposes), and for each grid cell representative site-response parameters were computed by spatially averaging the 3D model values of the boreholes located within or in the vicinity of the cell. The parameters considered are the depth to engineering bedrock  $H$ , the time-averaged shear-wave velocity in the upper 30 m  $V_{S30}$ , the average shear-wave velocity of the soil column above the bedrock  $V_{S,soil}$ , and the fundamental site period  $T_G$ . Here  $H$  is defined as the depth at which  $V_s$  first reaches approximately 760 m/s,  $V_{S30}$  is evaluated as  $V_{S30} = 30 / \sum(d_i / V_{S,i})$  over the upper 30 m,  $V_{S,soil}$  is computed in an analogous manner from the ground surface down to depth  $H$ , and  $T_G$  is obtained as  $T_G = 4 \sum(d_i / V_{S,i})$ . These four parameters jointly describe the stiffness, thickness and dynamic response characteristics of the soil column and constitute the primary inputs for subsequent site-classification and amplification analyses.

The resulting spatial distributions of  $H$ ,  $V_{S30}$ ,  $V_{S,soil}$  and  $T_G$  are shown in Figure 9a–d. The bedrock depth map (Figure 9a) clearly delineates three major geotechnical domains. The northern mountainous region, including Bukhansan, Dobong and adjacent upland areas, exhibits shallow engineering bedrock, with typical depths on the order of 5–15 m. The central basin, covering districts such as Jongno, Jung and Yongsan, is characterized

by intermediate bedrock depths of approximately 20–35 m. In contrast, the southern alluvial plain, including Gangnam, Songpa, Yeouido and parts of Yeongdeungpo, reveals significantly greater thicknesses of unconsolidated deposits, with bedrock depths frequently exceeding 25–50 m and locally reaching even larger values along the Han River corridor. These patterns reflect the combined influence of long-term fluvial deposition, artificial reclamation and differential erosion between the granite-dominated uplands and the soft alluvial lowlands.

The spatial variation of  $V_{S30}$  (Figure 9b) mirrors the distribution of bedrock depth but also incorporates near-surface stiffness contrasts. In the northern uplands,  $V_{S30}$  attains values typically equal to or greater than about 700 m/s, indicating stiff to hard rock conditions representative of NEHRP Site Class B. The central basin exhibits intermediate  $V_{S30}$  values of roughly 400–500 m/s, consistent with weathered soil and weathered rock sequences and corresponding predominantly to Site Class C. The southern alluvial belt along the Han River is dominated by soft deposits with  $V_{S30}$  in the range of approximately 220–320 m/s, falling within Site Classes D–E. Localized zones of very low  $V_{S30}$  coincide with thick reclaimed or riverine alluvium, whereas relatively elevated  $V_{S30}$  values in parts of the eastern and western margins indicate areas where weathered rock rises closer to the surface despite being outside the main mountainous region.

The map of  $V_{S,soil}$  (Figure 9c) emphasizes the bulk stiffness of the entire soil column above the engineering bedrock. While  $V_{S30}$  is sensitive to the properties of the shallow 30 m,  $V_{S,soil}$  is influenced by the full thickness of the sedimentary package. As a result, differences between the two parameters are most evident in deep basin areas where the low-velocity soft layers are underlain by moderately stiff strata at greater depths. In many parts of the central basin,  $V_{S30}$  and  $V_{S,soil}$  are similar because the bedrock depth is comparable to or only moderately larger than 30 m. However, in the southern alluvial plain the contrast becomes more pronounced: even where  $V_{S30}$  remains low,  $V_{S,soil}$  may increase slightly with depth due to the presence of denser sands or weathered materials beneath the uppermost soft clays and silts. Conversely, in the northern uplands, where bedrock is shallow,  $V_{S30}$  and  $V_{S,soil}$  both reflect high stiffness values and are nearly indistinguishable. This comparison highlights that  $V_{S,soil}$  is particularly useful for identifying deep, low-stiffness basins whose dynamic response may not be fully captured by  $V_{S30}$  alone.

The fundamental site period  $T_G$  (Figure 9d) integrates the information contained in  $H$  and the depth-dependent  $V_s$  structure into a single dynamical descriptor. Short site periods prevail in the northern granitic mountains and in areas with shallow bedrock, where  $T_G$  is generally small, indicating that local ground motion will tend to amplify short-period components of the input excitation. In the central basin, intermediate values of  $T_G$  correspond to moderate basin thickness and stiffness, leading to resonance effects that are relevant for mid-rise buildings with natural periods in the corresponding range. The longest site periods occur in the deep southern alluvial corridors, particularly beneath Gangnam, Songpa, Yeouido and portions of the western floodplain, where thick, soft deposits and low  $V_{S,soil}$  values yield substantially larger  $T_G$ . These zones are expected to preferentially amplify long-period motions and thus pose elevated risk to mid- to high-rise structures and lifeline systems that are sensitive to displacement demands.

Taken together, the four maps in Figure 9 demonstrate that Seoul's seismic site effects are strongly controlled by the interplay between bedrock depth, near-surface stiffness and basin geometry. The northern mountainous region with shallow bedrock and high  $V_{S30}$  forms a stiff domain with relatively low amplification and short dominant periods, whereas the central weathered basin and the southern alluvial plain form progressively softer domains with increasing amplification potential and longer fundamental periods. When combined with the spatial distribution of building inventory and underground infrastructure,

these zonation results provide a physically consistent basis for the subsequent HAZUS-based seismic vulnerability and risk assessments presented in the following sections.

### 3.2. $V_S$ -Based Ground Model Visualization

The kriging- and SGS-based three-dimensional  $V_S$ -ground model provides a high-resolution representation of subsurface stiffness together with its spatial uncertainty across the Seoul basin. The kriging-derived mean  $V_S$  field delineates a continuous low-velocity corridor along the Gangnam–Songpa–Yeouido axis, where soft alluvial deposits typically exhibit  $V_S$  values of approximately 150–250 m/s and extend to depths exceeding 20–30 m. Within this corridor, the thickness of soft deposits varies sharply over short horizontal distances, reflecting the laterally heterogeneous depositional environments associated with the Han River and its tributary systems. In contrast, northern mountainous districts are characterized by shallow, high-velocity bedrock (generally  $V_S > 700$ –800 m/s) overlain by a thin soil mantle of roughly 5–15 m, consistent with granite-dominated terrains and limited sediment cover.

When integrated with the digital elevation model (DEM) and geological boundaries, the 3D  $V_S$ -ground model reveals systematic and physically interpretable linkages among geomorphology, lithology, and  $V_S$ -based site indices. Low-lying floodplains and reclaimed or filled zones along the Han River commonly coincide with low  $V_{S30}$  values and deep engineering bedrock, whereas steep northern slopes correspond to high  $V_{S30}$  and shallow bedrock. As a result, the derived maps of depth to engineering bedrock ( $H$ ),  $V_{S30}$ , depth-averaged soil shear-wave velocity ( $V_{S,soil}$ ), and fundamental site period ( $T_G$ ) exhibit coherent spatial patterns rather than random fluctuations (Figure 9a–d). Shallow bedrock and high  $V_{S30}$  values cluster in the northern and northeastern highlands, while deep bedrock, low  $V_{S30}$ , low  $V_{S,soil}$ , and long  $T_G$  values concentrate in the southern and southwestern alluvial plains. These parameter fields jointly characterize not only the amplification potential but also the predominant period of site response, providing a physically consistent basis for NEHRP site classification and for adjusting input response spectra in the subsequent HAZUS-based vulnerability analysis.

Beyond deterministic visualization, the ensemble of 100 SGS realizations enables explicit characterization of epistemic uncertainty associated with heterogeneous subsurface conditions and uneven data density. For each voxel, uncertainty is quantified using the coefficient of variation (COV), defined as the ratio of the ensemble standard deviation to the ensemble mean of the parameter of interest. Figure 10 presents the spatial distribution of COV for (a) bedrock depth  $H$ , (b)  $V_{S30}$ , (c)  $V_{S,soil}$ , and (d)  $T_G$ , all derived consistently from the SGS ensemble on the same 10 m  $\times$  10 m grid. For interpretability, uncertainty levels are categorized into four classes: low uncertainty ( $\text{COV} < 0.15$ ), moderate uncertainty ( $0.15 \leq \text{COV} < 0.30$ ), high uncertainty ( $0.30 \leq \text{COV} < 0.45$ ), and very high uncertainty ( $\text{COV} \geq 0.45$ ). These ranges correspond to typical geotechnical contexts, with low COV values indicating dense borehole coverage and relatively homogeneous bedrock or shallow weathered rock, and high to very high COV values indicating thick alluvium, reclaimed land, and river-adjacent zones characterized by strong lateral facies variability and limited local data support.

The COV maps exhibit a robust and interpretable spatial structure. Elevated uncertainty (high to very high COV) is concentrated along the Han River floodplain and reclaimed lowlands, highlighting areas where local predictions of  $V_S$ -based parameters are particularly sensitive to subsurface heterogeneity. Conversely, northern mountainous regions underlain by shallow competent bedrock show consistently low COV, indicating that  $V_S$ -based parameters are well constrained and spatially stable. Importantly, this reveals that locations with similar deterministic site indices (e.g., comparable  $V_{S30}$  or amplifica-

tion potential) may differ substantially in confidence level depending on local uncertainty, underscoring the added value of the SGS-based analysis.

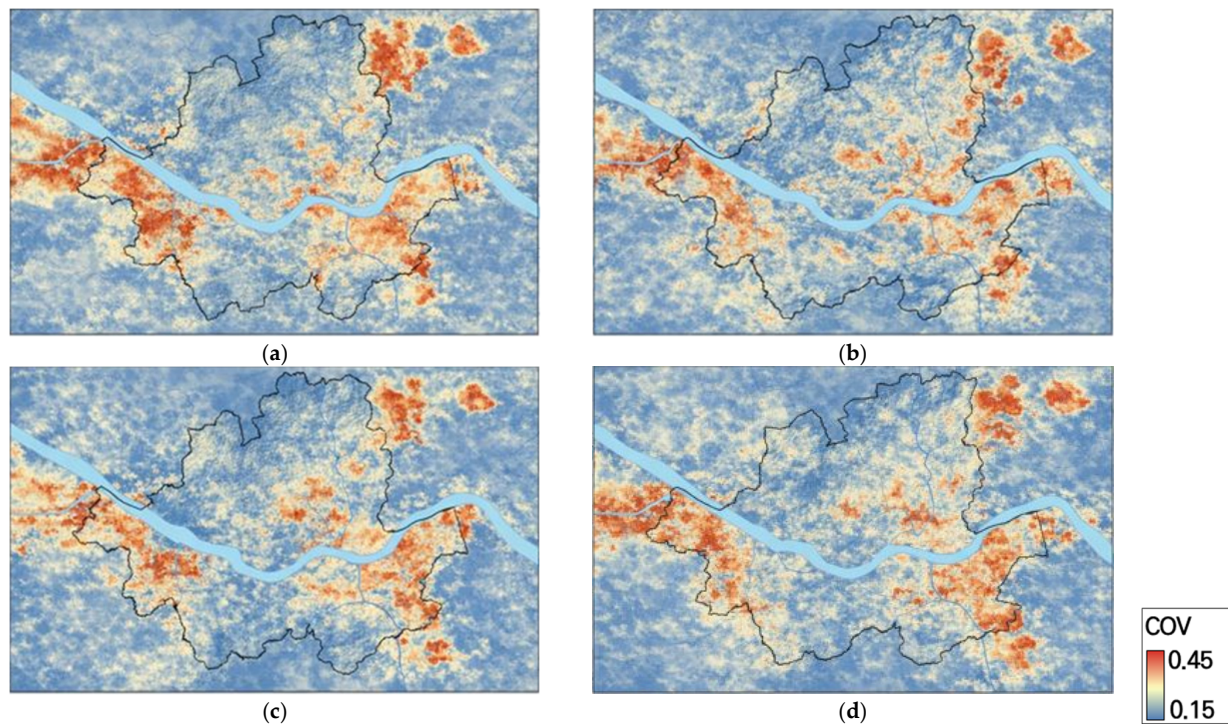


**Figure 10.** Spatial uncertainty (COV) of Vs-based site-response parameters inferred from 100 SGS realizations: (a) bedrock depth; (b)  $V_{S30}$ ; (c)  $V_{S,soil}$ ; (d)  $T_G$ .

In this study, final HAZUS-based vulnerability and loss estimates are computed using the kriging-derived mean fields to maintain compatibility with standard deterministic regional loss-estimation workflows. However, the SGS-derived uncertainty fields play a complementary role by qualifying interpretation, supporting sensitivity screening, and identifying zones where additional geotechnical investigations would most effectively reduce epistemic uncertainty in site classification, amplification parameters, and vulnerability assessment. Thus, the combined visualization of mean fields and COV maps provides a more complete and transparent representation of both seismic site conditions and their associated confidence levels for urban-scale seismic risk analysis.

### 3.3. Spatial Distribution of Seismic Damage Probabilities

The influence of key geotechnical parameters on structural fragility is mediated through their effect on site-specific ground-motion amplification and response spectra. In the proposed framework,  $V_{S30}$  and depth to engineering bedrock ( $H$ ) are not used merely as site classification indicators, but explicitly control the modification of input seismic demand prior to fragility evaluation. Specifically, lower  $V_{S30}$  values and greater bedrock depths are associated with higher amplification factors and longer predominant site periods, which shift the surface response spectra toward larger spectral accelerations at periods relevant to mid- and high-rise buildings. As a result, the effective spectral acceleration  $S_a(T_1)$  applied to the fragility curves increases systematically in soft-soil basins compared with shallow-rock sites under the same input motion.

Because HAZUS fragility curves are expressed as conditional probabilities of exceeding discrete damage states given spectral acceleration, this site-dependent modification of  $S_a$  directly translates into higher probabilities of extensive and complete damage for buildings located on deep alluvial deposits. Conversely, sites characterized by high  $V_{S30}$  and shallow

bedrock experience reduced amplification and lower effective demand, leading to a leftward shift in the fragility function and reduced damage probabilities. This mechanism explains the pronounced spatial contrasts observed in the vulnerability maps, where districts such as Gangnam, Songpa, and Yeouido exhibit elevated damage probabilities despite similar regional shaking levels, while northern mountainous districts show consistently lower vulnerability. The explicit linkage between geotechnical parameters, site response, and fragility-based damage estimation reinforces the physical consistency of the proposed hybrid framework.

The building inventory analysis, combined with the localized site-response parameters, highlights pronounced spatial variations in seismic vulnerability across Seoul. In several districts, particularly in older residential neighborhoods of the outer boroughs, buildings constructed before the introduction of modern Korean seismic codes constitute more than one third of the stock, indicating a substantial proportion of structures that were not explicitly designed for seismic loading. Districts such as Eunpyeong, Dobong, and Nowon exhibit a high prevalence of low-rise masonry and non-ductile reinforced-concrete buildings, which are associated with relatively high fragility even under moderate shaking. In contrast, central and southern business districts, including Jongno, Jung, Gangnam, and Yeouido, are dominated by mid-rise and high-rise reinforced-concrete and steel buildings that generally conform to at least low-code or moderate-code design standards, although pockets of older, non-ductile construction remain embedded within these high-density areas.

The damage-state probabilities presented in Figure 11 represent the conditional probabilities that a building exceeds or is classified into each discrete damage state defined in the HAZUS framework, namely Slight (minor), Moderate, Extensive, and Complete damage. These probabilities are derived directly from the corresponding fragility curves, which express the likelihood of reaching or exceeding a given damage state as a function of spectral acceleration at the fundamental period of the structure.

For clarity, the probability values shown in Figure 11 are continuous quantities ranging from 0 to 1. A higher probability of “minor damage” therefore indicates a greater likelihood that buildings within a given grid cell or district experience at least slight structural or non-structural damage under the considered seismic scenario, rather than representing a categorical threshold or binary outcome. Similarly, probabilities associated with moderate, extensive, and complete damage states quantify the expected spatial variation in the severity of structural damage, conditional on the site-specific seismic demand. Accordingly, the spatial patterns observed in Figure 11 should be interpreted as maps of expected damage-state likelihoods rather than deterministic damage classifications. This probabilistic representation allows direct comparison among districts and highlights areas where the same regional ground motion results in markedly different damage expectations due to local site amplification and building characteristics.

The HAZUS-based damage assessment translates these structural and geotechnical contrasts into explicit spatial patterns of damage probability for a representative design-level earthquake scenario. Figure 11 illustrates the spatial distribution of building damage probabilities by damage state on a 100 m grid: slight, moderate, extensive, and complete damage (Figure 11a–d, respectively). For the slight damage state, non-zero probabilities are widespread over most developed areas, but the highest concentrations occur in the southern alluvial corridor and in older low-rise districts, reflecting the combined effects of soft soil amplification and vulnerable structural typologies. The moderate damage probabilities are more spatially localized, forming clusters in parts of Yeouido, Gangnam, Songpa, and selected inner-city districts where dense mid-rise construction overlies thick, low-velocity deposits and where design levels are predominantly pre-code or low-code.

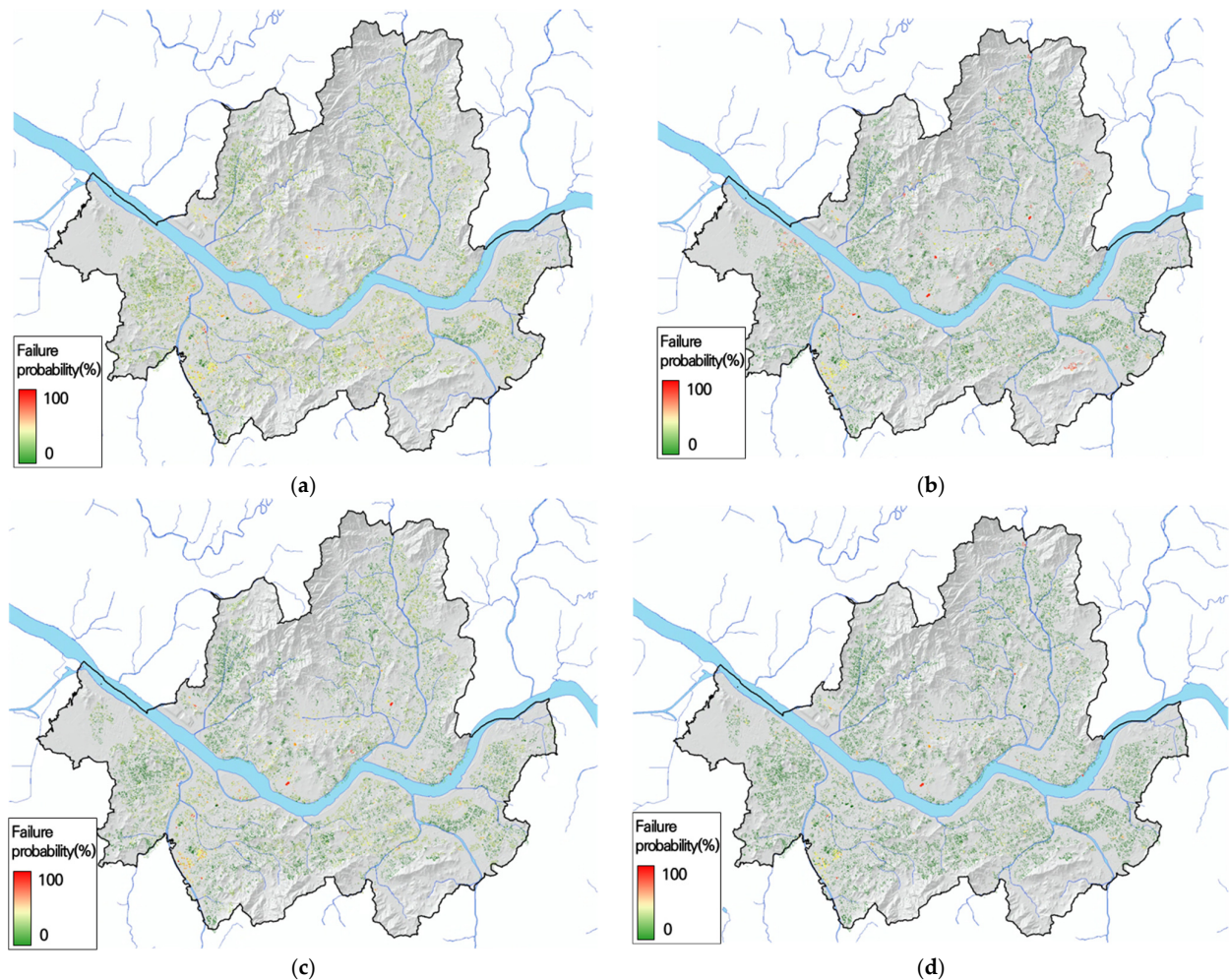


**Figure 11.** Spatial distribution of damage probabilities by damage state: (a) Slight; (b) Moderate; (c) Extensive; (d) Complete.

Extensive damage probabilities are generally low over the metropolitan area, but distinct hotspots emerge at the overlap of unfavorable ground conditions and structurally vulnerable building stocks. These hotspots include segments of the Han River floodplain with deep basins and long site periods, as well as neighborhoods characterized by aging masonry and non-ductile reinforced-concrete buildings. The complete damage state is associated with the lowest probabilities and is confined to a small number of grid cells, typically where high shaking intensity coincides with pre-code, brittle construction on soft alluvium. The progressive reduction in spatial extent from slight to complete damage states illustrates the expected filtering effect of structural capacity and code level on the transition from non-structural cracking and minor damage to partial or total collapse.

A comparison of the damage probability maps with the site-response zonation in Figure 11 demonstrates that regions with long fundamental site periods and low  $V_{S30}$  values tend to experience elevated probabilities of moderate and extensive damage, particularly where mid-rise or high-rise buildings with periods close to the basin resonance dominate the inventory. Conversely, northern mountainous areas classified as NEHRP Site Class B exhibit consistently low damage probabilities across all damage states, despite including some older building stock, because the underlying rock is stiff and the expected amplification is limited. This spatial interplay between ground motion amplification, structural typology, and design level confirms that a purely structure-based assessment would underestimate risk in soft-basin areas, whereas a purely site-based assessment would fail to capture

the mitigating influence of modern code-compliant construction in parts of the central business districts.

The gridded damage probability results also provide a quantitative basis for subsequent aggregation into administrative units and for deriving composite indicators such as the Damage State Index, economic loss ratios, and casualty estimates. In particular, grid cells that exhibit simultaneously high probabilities of moderate or greater damage and high building exposure (in terms of floor area or replacement cost) can be designated as priority zones for detailed microzonation, retrofitting programs, and emergency response planning.

### 3.4. Seismic Vulnerability Mapping for Seoul

The integration of the 3D  $V_s$ -ground model, the harmonized building inventory, and the localized HAZUS implementation yields spatially explicit maps of seismic vulnerability for the entire Seoul metropolitan area. For the representative 500-year return-period scenario considered in this study, the model outputs include the grid-based economic loss ratio (ELR), the casualty fatal ratio (CFR), and the composite seismic risk index (CRI), all computed at the 100 m grid-cell level and mapped at 100 m resolution (Figure 12). ELR represents the expected direct and indirect economic loss normalized by the total exposed replacement value of buildings and contents within each grid cell. CFR is defined as the ratio of expected fatalities to the exposed population, accounting for occupancy type and time-of-day assumptions. The CRI integrates these outputs by aggregating the normalized damage index, ELR, CFR, and lifeline functional-loss ratio using the baseline equal-weight scheme  $(\alpha, \beta, \gamma, \delta) = (0.25, 0.25, 0.25, 0.25)$  defined in Section 2.7, thereby providing a single spatial measure of multi-dimensional seismic risk.

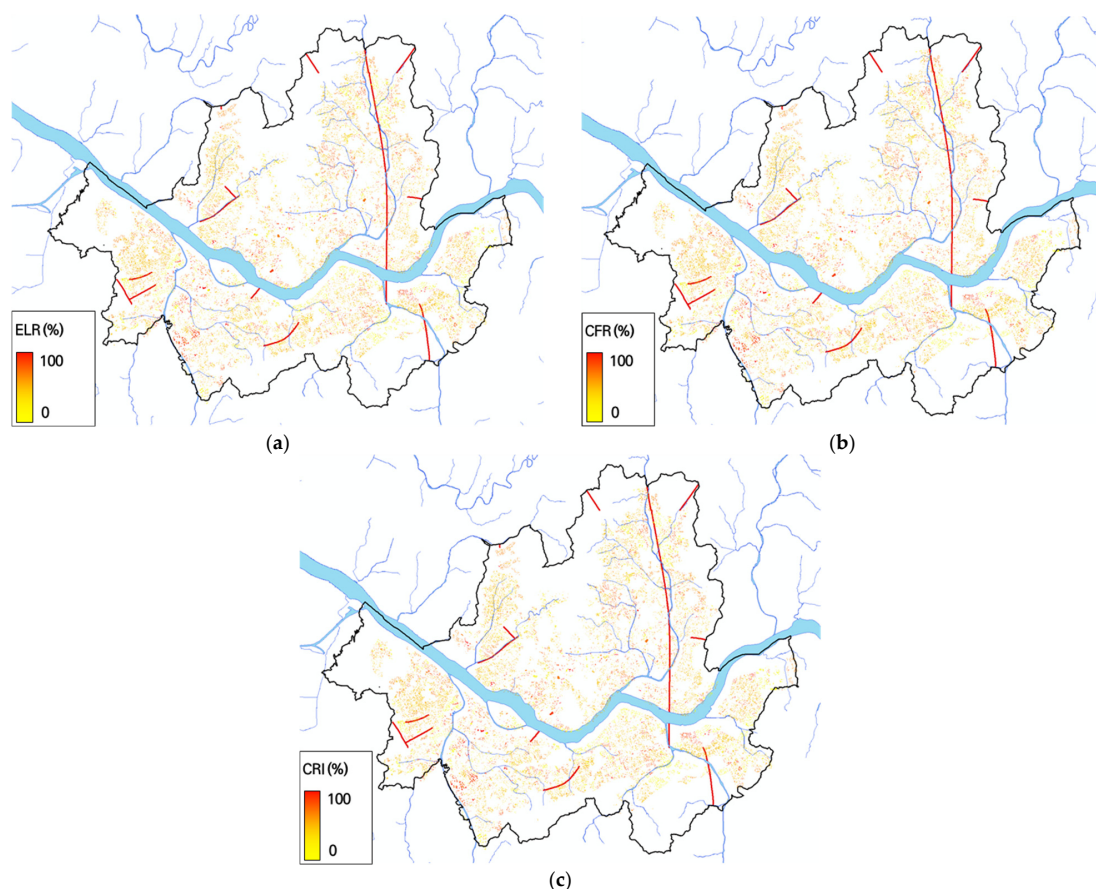


**Figure 12.** Spatial distribution of (a) economic loss ratios (ELR), (b) casualty fatal ratio (CFR), and (c) composite seismic risk index (CRI) in Seoul.

Figure 12a illustrates the spatial distribution of ELR across Seoul. The highest economic loss ratios are concentrated in the central business districts and along the southern alluvial corridor, particularly in Gangnam, Songpa, Yeouido, and parts of Yongsan and Jung. In these areas, thick soft deposits identified in the Vs-based ground model coincide with clusters of high-rise and mid-rise commercial and office buildings with large exposure values, leading to substantial expected repair, replacement, contents, and business-interruption costs even for moderate levels of structural damage. Elevated ELR values are also observed along several major transportation and commercial corridors, reflecting the combined effect of dense building stocks, high asset values, and unfavorable site amplification. In contrast, northern districts underlain by shallow bedrock and characterized by lower building heights and exposure generally exhibit low to moderate ELR levels, except in localized pockets where older non-ductile buildings or local sedimentary basins are present. This spatial pattern indicates that ELR is primarily governed by the intensity of economic exposure, modulated by local site amplification.

The spatial pattern of CFR shown in Figure 12b is partially aligned with, but not identical to, the ELR distribution. High CFR values emerge in densely populated residential zones where a large proportion of the building stock consists of pre-code or low-code low-rise masonry and non-ductile concrete buildings. These areas include older residential neighborhoods in Eunpyeong, Dobong, Nowon, and parts of Seongdong and Gwanak, where the combination of structural fragility and high nighttime occupancy leads to relatively high expected fatality ratios under the modeled scenario. Central business districts exhibit moderate CFR values despite their high exposure, because a larger fraction of the building stock satisfies moderate- or high-code requirements, which mitigates collapse probabilities even when economic losses remain substantial. This contrast highlights that life-safety risk and economic-loss risk are not spatially coincident and therefore need to be evaluated using distinct indicators prior to integration.

Figure 12c presents the spatial distribution of the CRI, which combines damage, loss, casualty, and infrastructure-failure components into a single normalized index. Because each CRI component originates from a distinct analytical pathway and exhibits its own spatial variability, the resulting CRI pattern reflects the compounded influence of subsurface conditions, structural vulnerability, human exposure, and infrastructure concentration. High-CRI zones appear where unfavorable ground conditions, vulnerable building stocks, and dense population and infrastructure coexist. These zones include the Gangnam–Songpa–Yeouido corridor along the Han River, where strong site amplification in thick alluvial deposits overlaps with high-rise commercial and residential developments, as well as parts of Jung and Yongsan that host critical governmental, transportation, and utility facilities. Additional CRI hotspots are observed in several outer districts where old low-rise masonry housing is concentrated on soft ground or in proximity to major lifeline corridors, indicating that relatively modest economic exposure can still translate into high composite risk when structural fragility and site effects dominate. Conversely, mountainous northern districts underlain by shallow, stiff bedrock and characterized by higher proportions of newer reinforced-concrete buildings generally display low CRI values, although isolated cells with elevated risk remain where local sedimentary basins or clusters of non-ductile structures are present.

Overall, the spatial distributions of ELR, CFR, and CRI demonstrate that seismic risk in Seoul is governed by the interplay among subsurface conditions, structural characteristics, and human exposure, rather than by any single factor. Even under the same seismic hazard scenario, areas experiencing similar levels of ground shaking can exhibit markedly different risk profiles depending on site class, building inventory, and occupancy patterns. The consistency between the mapped patterns and the known characteristics

of soft-ground urban damage observed in recent Korean inland earthquakes, such as the 2017 Pohang event, supports the validity of the integrated assessment framework. The resulting vulnerability and risk maps therefore provide a quantitative basis for prioritizing seismic retrofitting, allocating emergency-response resources, and formulating land-use and redevelopment policies that explicitly account for both ground-motion amplification and the spatial distribution of exposure in Seoul.

### 3.5. Quantitative Comparison with Simplified Site-Classification Approaches

To quantitatively assess the added value of the proposed framework, a comparative analysis was conducted against a simplified site-classification-based approach representative of previous urban-scale seismic vulnerability studies. In the simplified approach, site conditions were characterized using spatially uniform or zonation-based  $V_{S30}$  values, without explicit consideration of engineering bedrock depth or three-dimensional variability in subsurface stiffness. All other components of the analysis, including the building inventory, fragility functions, seismic scenario, and aggregation procedures, were kept identical to isolate the effect of site characterization.

Figure 13 compares the grid-level distributions of the expected Damage State Index (DSI), economic loss ratio (ELR), and Composite Risk Index (CRI) obtained from the two approaches for the 500-year return-period scenario. While the citywide mean values of DSI and ELR differ only moderately between the two cases, the proposed 3D  $V_s$ -ground-model-based approach produces substantially greater spatial variability. In particular, localized zones of elevated damage and loss emerge along deep alluvial basins and basin margins, where strong site amplification associated with thick soft deposits is captured only when depth-dependent and three-dimensional ground properties are considered.

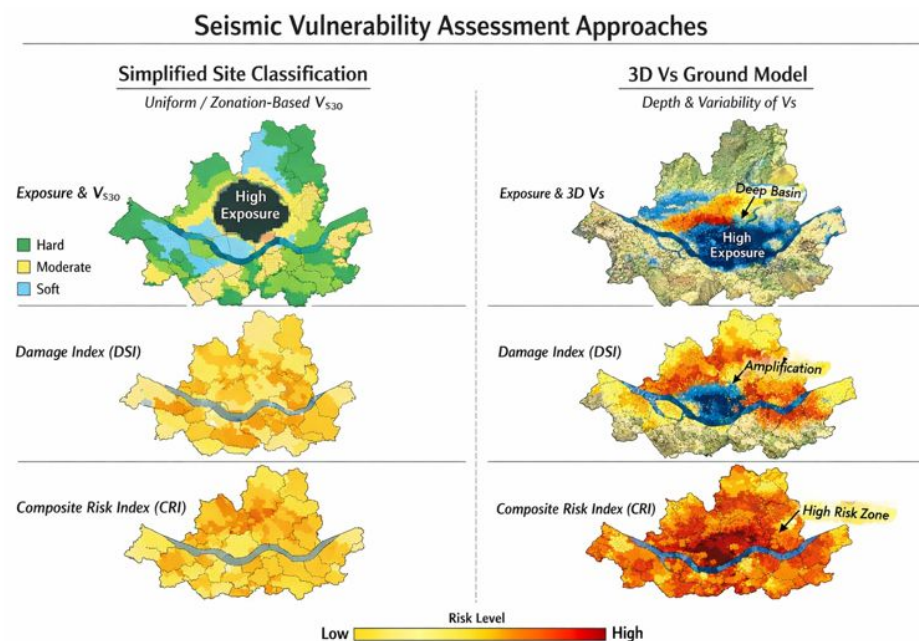


**Figure 13.** Conceptual comparison between simplified site-classification-based and 3D  $V_s$ -ground-model-based seismic vulnerability assessment frameworks.

This difference is further quantified by examining the upper tail of the risk distribution. For the top 10% highest-risk grid cells ranked by CRI, the proposed framework yields a higher mean CRI value and a wider interquartile range compared to the simplified approach, indicating stronger discrimination among high-risk zones. Moreover, the spatial overlap between the top 10% high-risk cells identified by the two approaches is limited,

with the simplified method capturing primarily large, contiguous exposure-driven areas while missing smaller but critical hotspots associated with local geotechnical conditions.

These results demonstrate that simplified site classifications tend to smooth spatial contrasts and underestimate risk concentration in areas where unfavorable subsurface conditions vary sharply over short distances. By contrast, the proposed framework explicitly resolves three-dimensional variability in soil stiffness and basin geometry, leading to a more differentiated and physically interpretable representation of seismic vulnerability. The quantitative differences observed in DSI, ELR, and CRI distributions provide direct evidence of the added value of the proposed approach, particularly for urban risk zoning and prioritization applications where identifying localized high-risk areas is essential.

Figure 13 is included to conceptually illustrate the source of the observed quantitative differences between the two approaches. It is not intended to replace the quantitative comparison, which is presented in terms of grid-level distributions and upper-tail statistics of DSI, ELR, and CRI. Figure 13 provides a schematic illustration of the fundamental differences between a simplified site-classification-based approach and the proposed 3D  $V_s$ -ground-model-based framework. The maps are conceptual representations intended to illustrate relative spatial patterns, not direct numerical outputs of the seismic risk calculations. In the simplified approach, site effects are represented using uniform or zonation-based  $V_{s30}$  values, leading to comparatively smooth spatial patterns of damage and risk that are largely driven by exposure distribution. In contrast, the proposed framework explicitly incorporates three-dimensional variability in subsurface stiffness and engineering bedrock depth, enabling localized amplification effects associated with deep alluvial basins and basin-margin transitions to be resolved. The color scale indicates relative (normalized) levels of damage and risk, and is used solely to highlight conceptual differences in spatial variability between the two approaches. Quantitative results and statistical comparisons are presented separately in the Section 3.

#### 4. Discussion

The hybrid seismic vulnerability framework developed in this study highlights the critical importance of explicitly coupling large-scale geotechnical information, three-dimensional  $V_s$ -ground models, and localized fragility functions when assessing seismic risk at the urban scale. The Seoul application shows that the spatial variability of subsurface conditions is pronounced even over short distances, with bedrock depth,  $V_{s30}$ ,  $V_{s,soil}$ , and fundamental period  $T_G$  exhibiting strong contrasts between the northern mountainous region, the central weathered basin, and the southern alluvial plain (Figure 9). In the northern granite-dominated zones,  $V_{s30}$  typically exceeds about 700 m/s and engineering bedrock is encountered within 5–15 m, whereas in the Gangnam–Songpa–Yeouido corridor  $V_{s30}$  falls to roughly 220–320 m/s and bedrock depths commonly reach 25–50 m. These differences translate directly into variations in site amplification and seismic demand that cannot be captured by district-averaged parameters or uniform site-class assignments. Earlier studies in Korea and elsewhere have often employed simplified zonations or adopted a single NEHRP class per administrative district [6–8,15,18], but the present analysis demonstrates that such approaches tend to mask low- $V_s$  corridors and localized soft-soil pockets that are strongly controlled by geomorphological and sedimentary histories [39–42].

The locally calibrated, multi-parameter  $N$ - $V_s$  regression model represents a second key contribution of this work. Conventional empirical relationships such as those by Ohta and Goto, Imai and Tonouchi, or Hasancebi and Ulusay [20–22] showed mean absolute errors on the order of 100–250 m/s when directly applied to Seoul’s heterogeneous alluvium, leading to systematic over- or underestimation of  $V_s$  in specific stratigraphic units. By incorporating soil classification, geological age, and depth into the regression form, the

proposed model reduces prediction uncertainty by more than twenty percent relative to these traditional equations and reproduces measured downhole and VSPs with mean relative differences generally below ten percent. This confirms that  $N$ - $V_s$  correlations must reflect local depositional environments, aging effects, and stress histories if they are to provide reliable inputs to city-scale  $V_{S30}$  and  $T_G$  mapping [25–31]. In areas without direct seismic tests, the calibrated regression ensures that  $V_s$  and  $V_{S30}$  can be estimated with sufficient accuracy to support HAZUS-type site classification without introducing systematic bias in the assignment of site classes and amplification factors. Although these diagnostics support the stability of the calibrated regression, further refinement could be achieved by incorporating additional explanatory covariates (e.g., fines content, plasticity, and groundwater conditions) and by expanding the paired seismic–SPT dataset in underrepresented zones to reduce epistemic uncertainty.

The geostatistical construction of the 3D  $V_s$ –ground model further advances urban seismic hazard characterization by providing a physically consistent and uncertainty-aware description of the subsurface. Ordinary kriging, guided by direction-dependent variograms, yielded low mean prediction errors and an RMSE of about 85 m/s, improving upon inverse distance weighting by approximately twenty-five percent. Sequential Gaussian simulation then generated an ensemble of realizations that preserved the global mean and variance of  $V_s$  while revealing spatial patterns of uncertainty associated with limited data density, reclaimed land, and riverine alluvium. The resulting maps of bedrock depth,  $V_{S30}$ ,  $V_{S,soil}$ , and  $T_G$  (Figure 9) clarify the geotechnical controls on site response, and they can be directly combined with DEM-derived topography and geological boundaries to support voxel-based wave-propagation analyses and one- or three-dimensional site-response simulations. The explicit quantification of uncertainty also offers a rational basis for identifying locations where additional investigations or dense microzonation surveys would most effectively reduce epistemic uncertainty [39–41].

Use of SGS uncertainty products (scope statement). In this study, Sequential Gaussian Simulation (SGS) realizations are used to quantify and visualize epistemic uncertainty in the 3D  $V_s$ –ground model (e.g., voxel-wise dispersion across realizations) as diagnostic products that support interpretation and future data-acquisition prioritization. However, the seismic vulnerability, loss, and CRI maps reported in this manuscript are presented as deterministic best-estimate (expected-value) outputs for each scenario, and SGS uncertainty is not propagated through the full damage/loss workflow to derive map-level confidence intervals. Full end-to-end uncertainty propagation that jointly samples hazard, site response, fragility parameters, and SGS-based  $V_s$  variability is therefore stated as a limitation and a topic for future research.

Although the present study employs SGS to quantify subsurface uncertainty, the final seismic vulnerability and risk maps are presented in deterministic form, following standard HAZUS-based planning practice. Full propagation of geotechnical uncertainty through nonlinear site-response analysis and fragility functions to obtain confidence intervals or probabilistic loss distributions would require significantly greater computational effort and additional assumptions regarding ground-motion variability and structural response. Such extensions are beyond the scope of this paper and are identified as an important direction for future research.

When the  $V_s$ -based ground model is integrated with the detailed building inventory, the compound nature of seismic vulnerability in a high-density metropolis becomes evident. Districts such as Gangnam, Songpa, and Yeouido are simultaneously characterized by deep, low-velocity alluvial deposits and by clusters of mid-rise reinforced-concrete buildings and high-rise steel towers that concentrate economic value and human exposure. Even where these structures satisfy low- or moderate-code seismic provisions, their dynamic interaction

with long-period soft basins tends to amplify demands near the fundamental site period, leading to elevated probabilities of extensive and complete damage under scenario ground motions [9,18,32]. In contrast, many northern districts benefit from shallow, stiff bedrock, higher  $V_{S30}$ , and lower building densities, and the proportion of pre-code or non-ductile structures is smaller, so that expected damage levels remain comparatively low for similar regional shaking. This contrast confirms that seismic risk is governed by the interaction of ground conditions, building stock, and exposure rather than by any single factor, and underscores the need to evaluate these components jointly rather than in isolation.

The HAZUS-based workflow adapted in this study also enables a more nuanced interpretation of risk by distinguishing among physical damage, economic consequences, and human impacts. Damage-state probability maps derived from the fragility curves (Figure 11) reveal where structural performance is likely to deteriorate to moderate, extensive, or complete damage states, whereas the economic loss ratios (ELR), casualty fatal ratios (CFR), and composite risk index (CRI) maps (Figure 12) show that economic and life-safety risks are not necessarily co-located. Central business districts and major commercial corridors exhibit high ELR due to large exposure values and significant non-structural and business-interruption components, while older low-rise residential neighborhoods on soft or intermediate ground can display higher CFR despite more modest asset values. CRI combines these dimensions with lifeline vulnerability to identify hot-spot zones where unfavorable ground, fragile structures, dense population, and critical infrastructure overlap. This multi-criteria perspective is consistent with contemporary resilience initiatives such as RASOR, TELES, and J-SHIS-based systems [10–14,43], and it demonstrates the utility of composite indices for prioritizing retrofitting, emergency planning, and infrastructure investment at the city scale.

At the same time, the study has limitations that should be acknowledged. The geotechnical database, although large, remains spatially uneven, with data sparsity in certain industrial or restricted areas and potential biases in historical investigations. The  $N$ - $V_s$  regression and kriging models assume second-order stationarity and approximate Gaussian behavior, which may not fully capture strongly non-linear or anisotropic features in complex geological settings. The building inventory, while extensively cleaned and harmonized, still contains uncertainties in structural details such as reinforcement detailing, infill wall configuration, and foundation type, all of which influence real structural performance but are difficult to represent explicitly at the metropolitan scale. Moreover, the fragility functions used in this study are largely derived from international HAZUS datasets, with only limited local calibration; as a result, they may not fully reflect Korean construction practices, material properties, or code enforcement levels. Finally, the present application focuses on ground shaking and structural damage, with lifeline, liquefaction, and permanent ground deformation effects represented only in a simplified or preliminary manner.

Nevertheless, the authors acknowledge that the fragility functions employed in this study are largely derived from international HAZUS datasets and have not been fully recalibrated using post-earthquake damage observations or large-scale experimental results specific to Korean construction practices. At present, such empirical datasets remain limited for Seoul, particularly for moderate-to-strong ground motions that would be required for statistically robust calibration. Accordingly, the vulnerability model proposed herein should be interpreted as Seoul-adapted in terms of seismic demand, exposure representation, and site-response coupling, rather than as a fully localized fragility formulation. Future work will focus on incorporating emerging post-earthquake damage databases, experimental test results, and forensic investigations from recent Korean earthquakes to refine fragility parameters and reduce potential systematic bias.

A key limitation of the present work is that the building fragility functions are not calibrated against Korea-specific post-earthquake damage datasets for Seoul. While HAZUS provides a comprehensive and widely used fragility library, direct transfer of US-based fragility parameter sets to Korean construction practice may introduce systematic bias, particularly for building classes where detailing, materials, and code provisions differ. In this study, we mitigate this limitation by (i) explicitly stratifying the building inventory by Korean seismic design eras and linking these categories to the closest HAZUS code levels to reflect differences in expected seismic performance, and (ii) localizing seismic demand through the 3D  $V_s$ -ground model and site-specific response spectra, which substantially affects the spectral acceleration input to fragility evaluation. Nevertheless, the resulting vulnerability estimates should be interpreted as “Seoul-adapted HAZUS-based” outcomes rather than fully validated local fragilities. Future work should prioritize development and calibration of Korean fragility functions using systematic post-earthquake damage databases, targeted experimental evidence, and/or physics-based numerical fragility analyses aligned with Korean structural typologies and seismic design standards.

Although the present study focuses on ground shaking amplification and structural vulnerability, other geotechnical failure mechanisms such as liquefaction, lateral spreading, and earthquake-induced settlement are also recognized as important contributors to seismic risk in certain geological settings. These phenomena were not explicitly analyzed in the current Results in order to maintain a clear and focused scope on site response and fragility-based damage estimation. Nevertheless, the proposed framework and the underlying three-dimensional geotechnical ground model provide a natural foundation for extending the analysis to liquefaction and ground deformation in future work. Parameters such as  $V_s$ , stratigraphy, groundwater depth, and depositional environment derived from the same database can be directly coupled with established liquefaction triggering and consequence models, enabling a fully integrated, multi-hazard seismic risk assessment for metropolitan areas.

Despite these limitations, the framework is broadly consistent with recent international efforts that integrate multi-resolution ground models, dense monitoring networks, and building inventories within GIS-based platforms, and it is well suited for extension toward a more comprehensive smart-city seismic resilience strategy. The voxel-based  $V_s$  model provides a natural foundation for performance-based liquefaction and settlement analyses, for coupling with PGV- and PGD-driven lifeline fragility models, and for embedding in digital-twin environments that assimilate real-time strong-motion data and asset information [3–5,29,45]. In this broader context, the methodology presented here can be regarded as a core decision-support component that transforms heterogeneous geotechnical and structural data into actionable information for scenario planning, emergency response, and long-term risk-informed land-use and infrastructure management.

## 5. Conclusions

This study developed and demonstrated an integrated framework that combines three-dimensional geotechnical ground modeling with a HAZUS-based urban seismic vulnerability assessment, using Seoul, Korea, as a case study. By systematically linking a large geotechnical database, a regionally calibrated  $N$ - $V_s$  regression model, a high-resolution 3D  $V_s$ -ground model, and a harmonized building inventory mapped to the HAZUS taxonomy and Korean seismic design eras, the work provides a spatially detailed and physically consistent characterization of seismic risk in a complex metropolitan environment.

A unified geotechnical database consisting of 63,124 boreholes and associated seismic surveys was first constructed through multi-stage quality control that combined statistical screening, kriging-based cross-validation, spatial autocorrelation analysis using Moran’s  $I$ ,

and cluster-based validation. Approximately 7.5% of the original records were removed as outliers, which improved spatial coherence and reduced error metrics such as relative RMSE, while stabilizing the  $V_{S30}$  distribution at levels consistent with plausible NEHRP site classes for the basin [37–42]. On this basis, a multi-parameter  $N$ – $V_s$  regression model was calibrated using 1200 collocated  $N$ – $V_s$  pairs and explicit soil-type and geological-age attributes. The locally derived equations reduced prediction uncertainties by more than twenty percent relative to conventional empirical correlations [20–22] and reproduced downhole and VSPs with mean relative errors generally below ten percent, thereby providing robust  $V_s$  estimates for  $V_{S30}$  classification at locations without direct seismic testing.

Using the QC-filtered  $V_s$  dataset and the calibrated  $N$ – $V_s$  model, a 3D  $V_s$ –ground model was constructed on a  $10\text{ m} \times 10\text{ m} \times 1\text{ m}$  regular grid through ordinary kriging and sequential Gaussian simulation. The kriging predictions exhibited mean errors on the order of 0.8% and an RMSE of approximately 85 m/s, outperforming inverse distance weighting by about twenty-five percent, while the stochastic simulations generated an ensemble of realizations that captured both subsurface heterogeneity and spatial uncertainty. Derived maps of bedrock depth,  $V_{S30}$ ,  $V_{S,soil}$ , and  $T_G$  (Figure 9) delineated three major geotechnical zones within Seoul: shallow, high-velocity rock in the northern mountains corresponding primarily to NEHRP Site Class B; intermediate stiffness weathered soil and rock in the central basin corresponding to Site Class C; and thick low-velocity alluvial deposits in the southern plain corresponding to Site Classes D–E. These zonations revealed strong lateral gradients in expected site amplification and fundamental period that would be obscured by coarser district-scale classifications.

In parallel, the building inventory of approximately 650,000–700,000 structures was standardized by mapping cadastral and registry attributes to the HAZUS structural taxonomy and to Korean seismic design-level categories (pre-code, low-code, moderate-code, and high-/very-high-code). Structural type, height class, occupancy, and year-built information were used to assign each building to a HAZUS class and design level, enabling the systematic application of fragility curves and associated damage-state probabilities [9]. Site-specific response spectra were obtained by adjusting spectral accelerations for local  $V_{S30}$ , bedrock depth, and  $T_G$ , and these spectra were used as inputs to the fragility functions. The resulting damage-state probability maps (Figure 11) and derived economic loss ratios, casualty fatal ratios, and composite risk indices (Figure 12) revealed clear spatial clustering of seismic risk. Southern districts such as Gangnam, Songpa, and Yeouido, where deep soft deposits coincide with dense clusters of mid- to high-rise buildings and high asset values, exhibited substantially higher probabilities of extensive and complete damage and higher ELR and CRI than northern mountainous districts where shallow bedrock and higher  $V_{S30}$  values dominate. At the same time, several outer residential areas with older low-rise masonry and non-ductile concrete buildings showed elevated CFR despite more moderate economic exposure, emphasizing the need to distinguish economic and life-safety dimensions of risk.

Taken together, these results demonstrate that the hybrid methodology proposed in this study offers a robust foundation for urban-scale seismic risk management. By explicitly integrating geotechnical, structural, and socio-economic information within a common GIS-based framework, the approach enables data-driven identification of high-risk corridors and districts, supports the prioritization of retrofitting and land-use interventions, and provides quantitative input for emergency response planning and recovery budgeting. Although the present analysis focuses on ground shaking and structural damage, the structure of the framework facilitates its extension to multi-hazard applications, including liquefaction, settlement, and lifeline disruption, by coupling the voxel-based  $V_s$  model with

PGV- and PGD-sensitive fragility relationships and with performance-based liquefaction and subsidence models.

The methodology is readily transferable to other Korean cities and to international contexts, provided that sufficient geotechnical and building data can be assembled and appropriately calibrated. The experience from Seoul underscores the necessity of moving beyond simplified, regionally uniform models toward integrated, city-specific ground and vulnerability models that can resolve the true spatial complexity of urban systems. In this sense, the proposed framework constitutes a practical stepping stone toward digital-twin-enabled seismic resilience, in which continuously updated geotechnical and structural information is linked to real-time hazard monitoring and scenario-based decision-support tools for sustainable urban risk reduction.

**Author Contributions:** Conceptualization, H.-S.K.; methodology, H.-S.K.; software, H.-S.K.; validation, H.-S.K. and J.-H.L.; formal analysis, H.-S.K.; investigation, J.-H.L.; resources, H.-S.K.; data curation, H.-S.K. and J.-H.L.; writing—original draft preparation, H.-S.K.; Funding acquisition: J.-H.L.; writing—review and editing, J.-H.L.; visualization, H.-S.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data supporting the findings of this perspective paper are available from the corresponding author upon reasonable request.

**Acknowledgments:** This work was supported by the Korea Institute of Energy Technology Evaluation and Planning(KETEP) and the Ministry of Trade, Industry & Energy(MOTIE) of the Republic of Korea (NO. RS-2025-02318006). And the author expresses sincere gratitude to the Seoul Institute and Dongguk University for providing access to geotechnical and building inventory datasets and for their continuous support of this research.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Galal, K.; Naimi, M. Effect of soil conditions on the response of reinforced concrete tall structures to near-fault earthquakes. *Struct. Des. Tall Spec. Build.* **2008**, *17*, 541–562. [[CrossRef](#)]
- Sun, C.-G.; Chun, S.-H.; Ha, T.-G.; Chung, C.-K.; Kim, D.-S. Development and application of a GIS-based tool for earthquake-induced hazard prediction. *Comput. Geotech.* **2008**, *35*, 436–449. [[CrossRef](#)]
- Mortezaei, A.; Mohajer Tabrizi, M. Damage assessment of RC buildings subjected to the different strong motion duration. *J. Phys. Conf. Ser.* **2015**, *628*, 012061. [[CrossRef](#)]
- Kim, H.S.; Chung, T.K.; Kim, M. Real-Time Spatiotemporal Seismic Fragility Assessment of Structures Based on Site-Specific Seismic Response and Sensor-Integrated Modeling. *Sensors* **2025**, *25*, 5171. [[CrossRef](#)] [[PubMed](#)]
- Leyton, F.; Ruiz, S.; Sepúlveda, S.A.; Contreras, J.P.; Rebolledo, S.; Astroza, M. Microtremors' HVSR and its correlation with surface geology and damage observed after the 2010 Maule earthquake (Mw 8.8) at Talca and Curicó, Central Chile. *Eng. Geol.* **2013**, *161*, 26–33. [[CrossRef](#)]
- Drabek, T.E.; Hoetmer, G.J. *Emergency Management: Principles and Practice for Local Government*; International City Management Association: Washington, DC, USA, 1991.
- Xu, L.; Liu, G. The study of a method of regional environmental risk assessment. *J. Environ. Manag.* **2009**, *90*, 3290–3296. [[CrossRef](#)]
- Felicioni, L.; Lupíšek, A.; Hájek, P. Major European stressors and potential of available tools for assessment of urban and buildings resilience. *Sustainability* **2020**, *12*, 7554. [[CrossRef](#)]
- Samuel, M.A.; Xiong, E.; Haris, M.; Lekeufack, B.C.; Xie, Y.; Han, Y. Assessing seismic vulnerability methods for RC-frame buildings pre- and post-earthquake. *Sustainability* **2024**, *16*, 10392. [[CrossRef](#)]
- Applied Technology Council. *Earthquake Damage Evaluation Data for California (ATC-13)*; Applied Technology Council: Redwood City, CA, USA, 1985.

11. Singhal, A.; Kiremidjian, A. Bayesian updating of fragilities with application to RC frames. *J. Struct. Eng.* **1998**, *124*, 922–929. [[CrossRef](#)]
12. Shinozuka, M.; Feng, M.Q.; Lee, J.; Naganuma, T. Statistical analysis of fragility curves. *J. Eng. Mech.* **2000**, *126*, 1224–1231. [[CrossRef](#)]
13. Cornell, A.C.; Jalayer, F.; Hamburger, R.; Foutch, D. Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines. *J. Struct. Eng.* **2002**, *128*, 526–533. [[CrossRef](#)]
14. Gardoni, P.; Der Kiureghian, A.; Mosalam, K.M. Probabilistic capacity models and fragility estimates for reinforced concrete columns based on experimental observations. *J. Eng. Mech.* **2002**, *128*, 1024–1038. [[CrossRef](#)]
15. Au, S.K.; Beck, J.L. Subset simulation and its application to seismic risk based on dynamic analysis. *J. Eng. Mech.* **2003**, *129*, 901–917. [[CrossRef](#)]
16. Schotanus, M.I.J.; Franchin, P.; Lupoi, A.; Pinto, P.E. Seismic fragility analysis of 3D structures. *Struct. Saf.* **2004**, *26*, 421–441. [[CrossRef](#)]
17. Lupoi, G.; Franchin, P.; Lupoi, A.; Pinto, P.E. Seismic fragility analysis of structural systems. In Proceedings of the 13th World Conference on Earthquake Engineering, Vancouver, BC, Canada, 1–6 August 2004. Paper No. 4008.
18. Kiureghian, D.A. Structural reliability methods for seismic safety assessment: A review. *Eng. Struct.* **1996**, *18*, 412–424. [[CrossRef](#)]
19. Pinto, P.E.; Giannini, R.; Franchin, P. *Seismic Reliability Analysis Methods*; IUSS Press: Pavia, Italy, 2004.
20. Luna, R.; Frost, J.D. Spatial liquefaction analysis system. *J. Comput. Civ. Eng.* **1998**, *12*, 48–56. [[CrossRef](#)]
21. National Institute of Building Sciences (NIBS). *Earthquake Loss Estimation Technology (HAZUS)*; Federal Emergency Management Agency: Washington, DC, USA, 1997.
22. ASTM D2487-85; Standard Practice for Classification of Soils for Engineering Purposes (Unified Soil Classification System). ASTM International: West Conshohocken, PA, USA, 1985.
23. Chung, C.K.; Kim, H.-S.; Sun, C.-G. Real-time assessment framework of spatial liquefaction hazard in port areas considering site-specific seismic response. *Comput. Geotech.* **2014**, *61*, 241–253. [[CrossRef](#)]
24. Sun, C.-G.; Kim, H.-S.; Cho, H.-I.; Jeong, B.-S. Real-time acceleration monitoring-based rapid earthquake response for ports in Korea. *Mar. Georesour. Geotechnol.* **2018**, *36*, 564–578. [[CrossRef](#)]
25. Han, J.W.; Kim, M.; Kim, H.S.; Chung, T.K.; Chung, C.K. Geostatistical-learning-based site-optimum 3D integration of borehole logs and geophysical data in urban area in South Korea. *Soils Found.* **2025**, *65*, 101684. [[CrossRef](#)]
26. EduPro Civil Systems, Inc. *ProShake User's Manual*; EduPro Civil Systems, Inc.: Redmond, WA, USA, 1997.
27. Box, G.E.P.; Lucas, H.L. Design of experiments in non-linear situations. *Biometrika* **1959**, *46*, 77–90. [[CrossRef](#)]
28. Federal Emergency Management Agency (FEMA). *HAZUS-MH Technical Manual (MR1)*; Federal Emergency Management Agency: Washington, DC, USA, 2003.
29. Korea Meteorological Administration (KMA). Earthquake: Notice, 2017 Pohang Earthquake. Available online: <https://www.weather.go.kr/w/eqk-vol/archive/major-eqk/pohang.do> (accessed on 10 February 2025).
30. Campbell, K.W.; Bozorgnia, Y. Updated near-source ground-motion (attenuation) relations for the horizontal and vertical components of peak ground acceleration and acceleration response spectra. *Bull. Seismol. Soc. Am.* **2003**, *93*, 314–331. [[CrossRef](#)]
31. Kim, H.S.; Chung, C.K. Integrated system for site-specific earthquake hazard assessment with geotechnical spatial grid information based on GIS. *Nat. Hazards* **2016**, *82*, 981–1007. [[CrossRef](#)]
32. Sun, C.G.; Kim, H.S. GIS-based regional assessment of seismic site effects considering the spatial uncertainty of site-specific geotechnical characteristics in coastal and inland urban areas. *Geomat. Nat. Hazards Risk* **2017**, *8*, 1592–1621. [[CrossRef](#)]
33. Toprak, S.; Holzer, T.L. Liquefaction potential index: Field assessment. *J. Geotech. Geoenviron. Eng.* **2003**, *129*, 315–322. [[CrossRef](#)]
34. Ohta, Y.; Goto, N. Empirical shear wave velocity equations in Japan. *Earthq. Eng. Struct. Dyn.* **1978**, *6*, 167–187. [[CrossRef](#)]
35. Imai, T.; Tonouchi, K. Correlation between N-value and shear wave velocity for soils in Japan. *Soils Found.* **1982**, *22*, 61–65.
36. Hasancebi, N.; Ulusay, R. Empirical correlations between shear wave velocity and penetration resistance for ground shaking assessments. *Bull. Eng. Geol. Environ.* **2007**, *66*, 203–213. [[CrossRef](#)]
37. Kramer, S.L. *Geotechnical Earthquake Engineering*; Prentice Hall: Upper Saddle River, NJ, USA, 1996.
38. Moran, P.A.P. Notes on continuous stochastic phenomena. *Biometrika* **1950**, *37*, 17–23. [[CrossRef](#)]
39. Journel, A.G.; Huijbregts, C.J. *Mining Geostatistics*; Academic Press: London, UK; New York, NY, USA, 1978.
40. Deutsch, C.V.; Journel, A.G. *GSLIB: Geostatistical Software Library and User's Guide*, 2nd ed.; Oxford University Press: New York, NY, USA, 1998.
41. Seyhan, E.; Stewart, J.P. Semi-empirical nonlinear site amplification from NGA-West2 data and simulations. *Earthq. Spectra* **2014**, *30*, 1241–1256. [[CrossRef](#)]
42. Wald, D.J.; Allen, T.I. Topographic slope as a proxy for seismic site conditions and amplification. *Bull. Seismol. Soc. Am.* **2007**, *97*, 1379–1395. [[CrossRef](#)]
43. Building Seismic Safety Council. *NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures (FEMA 450)*; Building Seismic Safety Council, National Institute of Building Sciences: Washington, DC, USA, 2004.

44. Federal Emergency Management Agency (FEMA). *Multi-Hazard Loss Estimation Methodology, Earthquake Model: HAZUS-MH 2.1 Technical Manual*; Federal Emergency Management Agency: Washington, DC, USA, 2013.
45. *KDS 17 10 00*; Seismic Design Standard for Buildings. Korean Ministry of Land, Infrastructure and Transport (MOLIT): Sejong, Republic of Korea, 2018.
46. Naik, S.P.; Gwon, O.; Park, K.; Kim, Y.-S. Land Damage Mapping and Liquefaction Potential Analysis of Soils from the Epicentral Region of 2017 Pohang Mw 5.4 Earthquake, South Korea. *Sustainability* **2020**, *12*, 1234. [[CrossRef](#)]
47. *ISO 22475-1:2021*; Geotechnical Investigation and Testing—Sampling Methods and Groundwater Measurements—Part 1: Technical Principles for Execution. International Organization for Standardization (ISO): Geneva, Switzerland, 2021.
48. *ASTM D420-98(2003)*; Standard Guide for Site Characterization for Engineering Design and Construction Purposes. ASTM International: West Conshohocken, PA, USA, 2003.
49. *BS EN ISO 14688-1:2018*; Geotechnical Investigation and Testing—Identification and Classification of Soil—Part 1: Identification and Description. British Standards Institution (BSI): London, UK, 2018.
50. *BS EN ISO 14688-2:2018*; Geotechnical Investigation and Testing—Identification and Classification of Soil—Part 2: Principles for a Classification. British Standards Institution (BSI): London, UK, 2018.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.