technologies

MDPI

Article

Machine Learning Delta Correction for Empirical and Hybrid Radiowave Propagation Models Toward Deterministic Predictions at 3.6 GHz

Tamás István Unger * and **Miklós Kuczmann**

Doctoral School of Multidisciplinary Engineering Sciences (MMTDI), Széchenyi István University, Egyetem tér 1, H-9026 Győr, Hungary; kuczmann@ga.sze.hu

* Correspondence: unger.tamas@nmhh.hu

Abstract

Deterministic radio wave propagation models provide high accuracy in complex outdoor environments but remain computationally impractical for large-scale network planning and spectrum management. In contrast, empirical and hybrid models offer low complexity at the expense of reduced accuracy, systematic bias, and limited terrain sensitivity. This paper proposes a unified delta learning framework that enhances fast baseline propagation models by learning a data-driven correction toward a deterministic Parabolic Equation Modeling (PEM) reference. A key novelty lies in a compact, physics-informed feature representation that replaces the full terrain profile with an 18-dimensional vector combining local geometric descriptors, global terrain characteristics, and baseline responses, enabling accurate correction with low-dimensional input. The study also provides the first systematic investigation of delta-based correction across multiple widely used propagation models. The framework is evaluated for free-space propagation, ITU-R P.1546, ITU-R P.1812, and ITU-R P.452 using ridge regression, kernel ridge regression, gradient boosting regression trees, and a neural network model. Model performance is assessed in terms of error reduction, bias mitigation, robustness across learning algorithms, and profile-level generalization to previously unseen propagation paths within the considered terrain categories. Results show substantial error reduction, with up to twofold improvement for simpler baseline models and consistent gains for hybrid models, while preserving computational efficiency.

Keywords: wave propagation modeling; radiowave propagation; delta learning; machine learning; radio engineering; machine learning-based correction; computational efficiency

Academic Editor: Manoj Gupta

Received: 19 May 2026
Revised: 10 June 2026
Accepted: 12 June 2026
Published: 15 June 2026

1. Introduction

1.1. Role of Radio Wave Propagation Modeling in Modern Wireless Systems

Wireless communication infrastructures have become a fundamental pillar of today's information society [1]. Beyond the well-known and widely used public mobile communication systems, radio-frequency technologies increasingly support industrial automation, transportation and logistics systems [2], the operation of critical infrastructures [3,4], and digital services [5,6] in everyday life in general [7,8]. Owing to their widespread adoption and extensive use, it is unsurprising that the requirements imposed on telecommunication networks are continuously increasing: reliable coverage [9], high data rates [10], low latency [11], and ever-improving spectrum efficiency with minimal interference [12] have all become key performance criteria. It is evident that these requirements can be best met

by achieving the most accurate possible understanding of radiowave propagation and by being able to predict it in advance, often already during the network and infrastructure planning phase, prior to actual deployment and implementation.

Radio wave propagation models therefore play a central role in the development, planning, and operation of wireless communication systems. These models enable the prior estimation of path loss, coverage characteristics, and spatial field strength levels before the actual deployment of a network [13]. This capability is particularly important in coverage planning, where the reliable determination of service area boundaries is a fundamental requirement. Equally significant is their role in intra-system and cross-border interference studies, where the assessment of mutual interference between different coexisting systems is necessary [14]. Beyond these applications, propagation models are indispensable tools for spectrum management and network optimization tasks [15], which require the rapid and consistent evaluation of a large number of configurations, parameter combinations, and environmental conditions. The application of more advanced and accurate models results in substantial savings in both time and cost for network planners and operators [16], while simultaneously supporting telecommunications regulatory authorities worldwide in carrying out international frequency coordination activities.

1.2. Unified Abstraction of Propagation Models

To facilitate the understanding of the general operation of radio wave propagation models and their relative roles, it is useful to interpret the different approaches within a unified abstraction space. From this perspective, the various models provide descriptions of the same physical phenomenon with differing levels of detail and complexity, while a substantial portion of the underlying input information remains common.

Formally, a given radio wave propagation model $\mathcal{M}_k$ can be regarded as a mapping [17,18] that transforms the space of geometric and environmental parameters into the space of propagation-related characteristics:

$$\mathcal{M}_k : \mathcal{X} \times \mathcal{Y}_k \to \mathcal{Z}, \tag{1}$$

where $\mathcal{X}$ denotes the set of geometric and environmental input parameters (such as terrain profile, transmitter–receiver geometry, frequency, etc.), while $\mathcal{Y}_k$ represents the space of model-specific parameters and configuration settings. The output space $\mathcal{Z}$ typically describes the spatial distribution of path loss or electric field strength, which is often interpreted not as a scalar quantity but as a distance- or height-dependent profile. This is illustrated in Figure 1, which schematically depicts, as an example, the electric field strength and the path loss as functions of the distance from the investigated source (e.g., a base station).

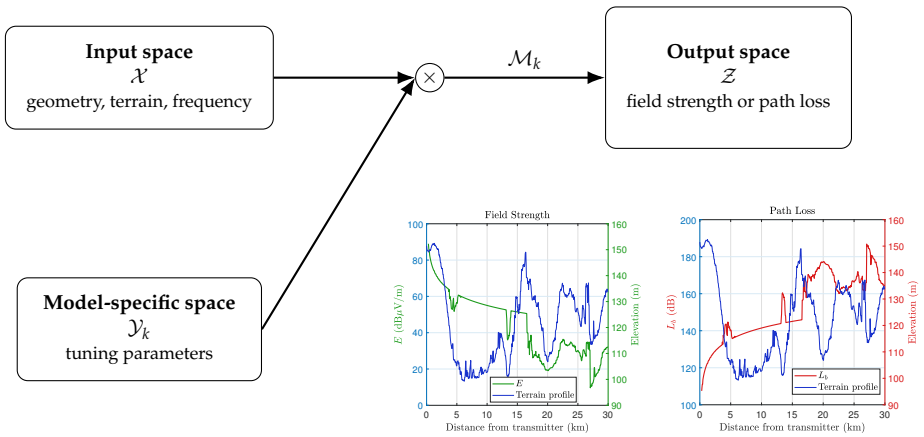


Figure 1. Abstract representation of a radio wave propagation model as a mapping.

It should be noted that the path loss and the electric field strength estimated by the model are naturally not independent of each other; they can be converted into one another using the following relationship [19]:

$$E = \mathrm{ERP} - L + 10\log_{10} f + 79.35, \quad (2)$$

where E denotes the electric field strength in dBμV/m, ERP is the effective radiated power of the transmitter in dBm, L is the path loss in dB, and f is the frequency in MHz.

1.3. Classification of Propagation Modeling Approaches

The unified abstraction framework introduced in the form of a mapping makes it possible to interpret different radio wave propagation models not as isolated estimation techniques, but as mappings of the same input space with varying levels of resolution. Consequently, the differences between models can primarily be understood in terms of how extensively and under what assumptions the available technical, geometric, and environmental information is exploited to construct the output profile.

In practice, radio wave propagation models can be classified into three fundamental categories [20]: empirical, deterministic, and hybrid approaches. The main characteristics of the different model categories are summarized in Table 1.

Table 1. Comparison of propagation model categories.

Category	Modeling Basis	Input Information	Accuracy/Complexity	Terrain Sensitivity
Empirical	statistical relations	averaged environment, simple geometry	limited accuracy, low complexity	limited
Hybrid	physical and empirical modeling	partial physics, geometry	moderate accuracy and complexity	moderate
Deterministic	physical laws, numerical methods	detailed terrain and environment	high accuracy, high complexity	strong

Empirical models, which are based on observational and measurement data, are characterized by high computational efficiency and ease of application [21]; however, their accuracy and terrain sensitivity are limited. In contrast, deterministic models describe the physical mechanisms of wave propagation in detail, relying on physical equations—most notably the Maxwell equations—which leads to significantly higher accuracy but often at the cost of computational complexity that challenges practical applicability [22]. Hybrid models occupy an intermediate position between these two extremes: certain aspects of wave propagation (e.g., diffraction effects) are modeled using physical principles, while other components are described using empirical relations derived from measurement data [23].

The most important distinction between propagation models therefore lies not merely in their accuracy, but in the trade-off between accuracy, computational complexity, and terrain sensitivity [24]. This trade-off is particularly relevant in large-scale planning and spectrum management tasks, where the accuracy of deterministic models and the efficiency of empirical and hybrid models would ideally be available simultaneously. For this reason, the evaluation and comparison of propagation models can only be considered well-founded if the combined impact of these three aspects is taken into account [25].

1.4. Overview of Propagation Models Considered in This Study

Radio-frequency waves propagating in an outdoor environment are fundamentally governed by Maxwell's equations [26], which describe the interactions between the electric and magnetic field components in time and space [27,28]. In a source-free, linear, isotropic medium, these equations can be reduced to the wave equation for the electric field. Assum-

ing a time-harmonic field and slowly varying material parameters, the wave equation can be further reduced to the Helmholtz equation, which forms the basis of most deterministic outdoor propagation models [29]:

$$\nabla^2 \mathbf{E}(\mathbf{r}) + k^2 \mathbf{E}(\mathbf{r}) = 0, \tag{3}$$

where $\mathbf{E}(\mathbf{r})$ is the vector of the complex amplitude of the electric field strength, $\mathbf{r}$ denotes the spatial position vector, and ∇^2 is the Laplace operator, which includes the second-order spatial derivatives [30]. The parameter k is the wavenumber, which describes the electromagnetic properties of the medium (magnetic permeability μ and electric permittivity ε) and also incorporates the frequency dependence: $k = 2\pi f\sqrt{\mu\varepsilon}$ [31].

It is easy to see that in a realistic terrestrial propagation environment, an analytical solution of the Helmholtz equation is practically unattainable. This is due to the presence of uneven terrain [32], environmental inhomogeneities, and large propagation distances, not to mention continuously changing weather conditions, which also cause ongoing variations in the electrical properties of environmental elements and the conductivity of the atmosphere. As a result of these limitations, a wide range of propagation models has been developed in practice, which approximate the underlying physical processes at different levels of abstraction, with varying degrees of accuracy and computational complexity [33]. In this subsection, we briefly outline the wave propagation models used in this paper, including the simplest terrain-independent free-space propagation model, the empirical ITU-R P.1546 model [34], the hybrid ITU-R P.1812 [35] and ITU-R P.452 [36] models, as well as deterministic Parabolic Equation Modeling (PEM) [37–39]. The qualitative differences observed between the outputs of the various propagation models are exemplified in Figure 2 for an identical propagation scenario.

The simplest analytical model can be established based on the physical laws of free-space wave propagation. It is important to note that although this is a deterministic model, it cannot be fitted into any of the categories described in Section 1.3, since it is a simplified analytical model that takes into account neither environmental nor terrain conditions, but only the radiation frequency and the distance of the receiving point from the transmitter. It can also be used as a reference: this model assumes ideal wave propagation conditions (in practice, ensuring the clearance of the first-order Fresnel zone), and thus provides the maximum achievable field strength and the minimum attainable path loss.

Starting from the Friis transmission equation [40], the power received by the receiver takes the following form:

$$P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi d}\right)^2, \tag{4}$$

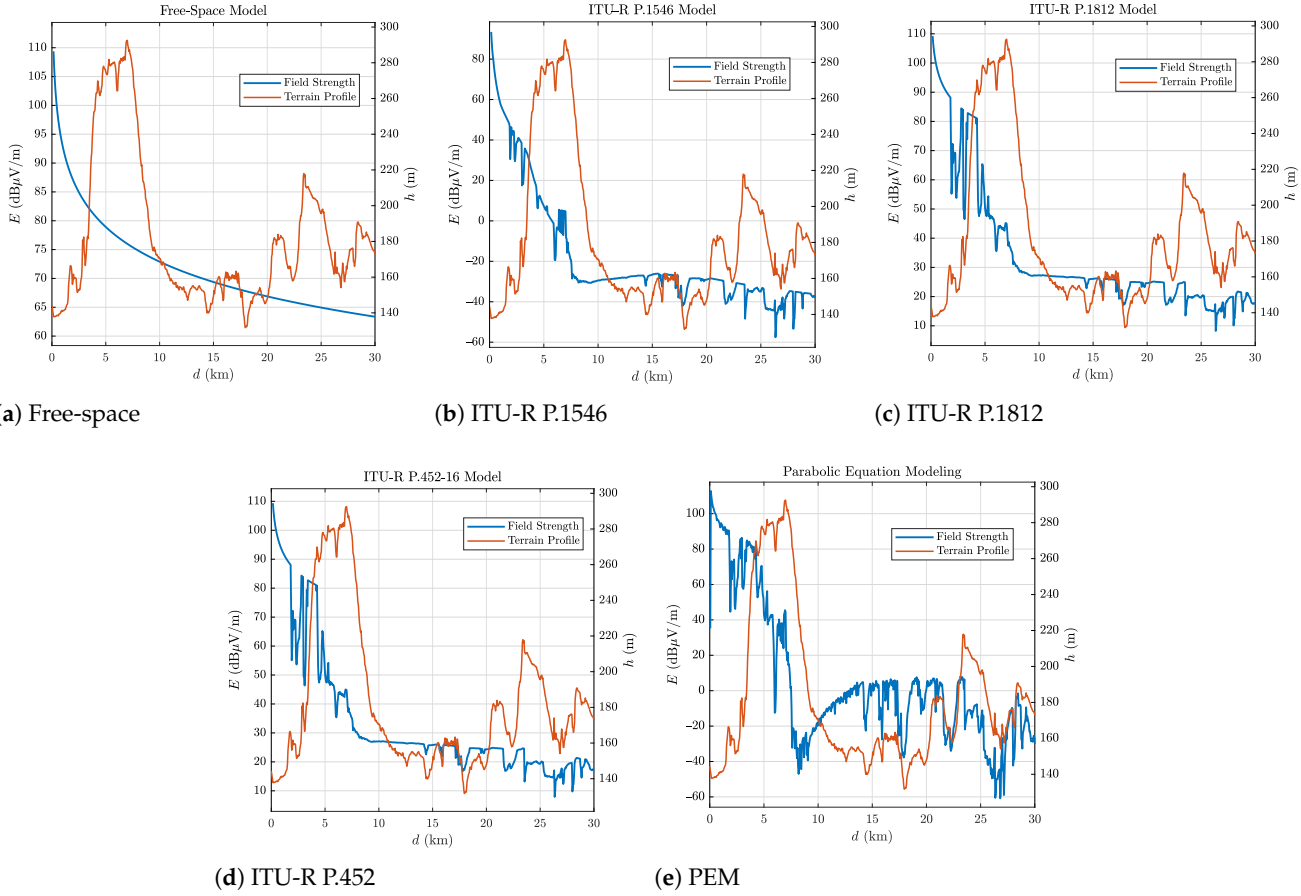
where P_t denotes the transmitted power, G_t and G_r are the gains of the transmitting and receiving antennas, respectively, λ is the wavelength, and d is the distance between the transmitter and the receiver. From this, the free-space path loss can be interpreted as the ratio (quotient) of the transmitted and received powers:

$$L_{FS} = \left(\frac{4\pi d}{\lambda}\right)^2. \tag{5}$$

Switching to a logarithmic scale (expressed in dB), and specifying the distance in kilometers and the frequency in megahertz, the free-space path loss simplifies to the following form:

$$L_{FS} = 32.44 + 20\log_{10}(d) + 20\log_{10}(f), \tag{6}$$

where the constant 32.44 dB originates from the geometric factor 4π, the speed of light, and the unit conversions [41]. As the formula shows, terrain conditions and, for example, the receiver height are irrelevant for this model.



(**a**) Free-space (**b**) ITU-R P.1546 (**c**) ITU-R P.1812

(**d**) ITU-R P.452 (**e**) PEM

Figure 2. Example outputs of the considered propagation models for the same scenario.

The ITU-R P.1546 Recommendation is a purely empirical outdoor propagation model originally developed for large-scale terrestrial radio services and is applicable in the frequency range of 30 MHz to 4000 MHz. The model is based on the statistical analysis of extensive measurement campaigns and predicts the electric field strength primarily as a function of distance, frequency, transmitter and receiver antenna heights, as well as time and location variability percentages. The influence of terrain is not represented explicitly in a geometric sense, but rather incorporated through statistical correction factors, such as effective antenna heights and terrain clearance angles. This principle is illustrated in Figure 3, which shows the family of measurement-based propagation curves used by the model at different frequencies and antenna heights, forming the basis for the interpolation and extrapolation procedures applied during prediction.

As a consequence, ITU-R P.1546 exhibits low computational complexity and can be robustly applied over large areas and long propagation paths. However, its sensitivity to detailed terrain features is inherently limited. The model is therefore best suited for describing average propagation conditions and is not capable of reliably capturing the impact of localized terrain obstacles or sharp diffraction effects. In the present study, the ITU-R P.1546 model serves as a representative example of empirical propagation approaches, illustrating the trade-off between low computational cost and limited terrain sensitivity, and providing a natural baseline for comparison with the hybrid and deterministic models discussed subsequently.

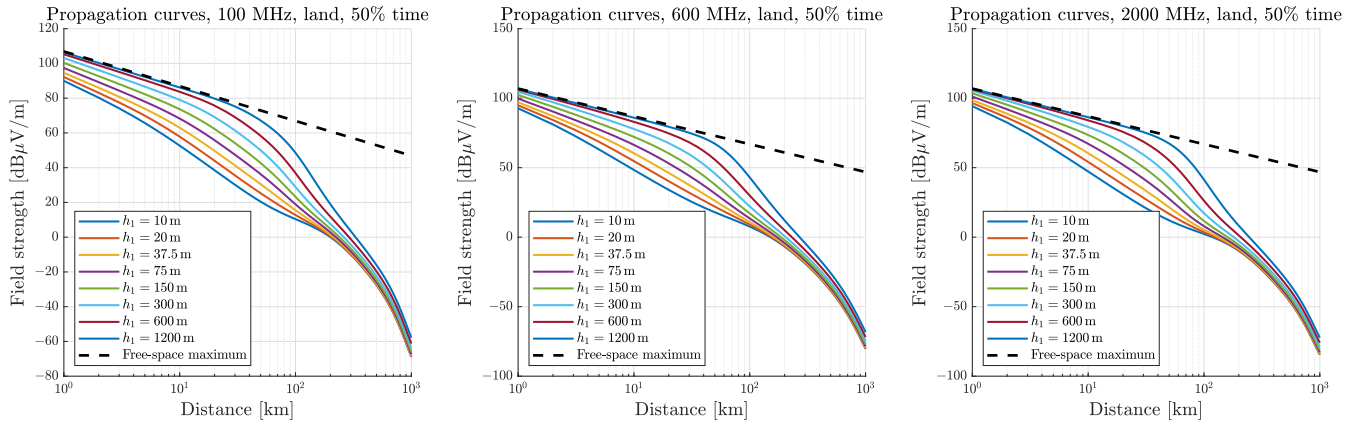


Figure 3. ITU-R P.1546 propagation curves based on measurements for 3 frequencies.

Parabolic Equation Modeling (PEM) provides a deterministic numerical approximation of electromagnetic wave propagation derived from the Helmholtz equation under the assumption of predominantly forward propagation. By introducing the reduced complex field $\psi(x,z)$ and neglecting backward-propagating wave components, the Helmholtz equation can be transformed into the parabolic equation

$$\frac{\partial \psi(x,z)}{\partial x} = -\frac{j}{2k_0}\left(\frac{\partial^2 \psi(x,z)}{\partial z^2} + k_0^2\left(n^2 - 1\right)\psi(x,z)\right), \tag{7}$$

where x denotes the horizontal propagation direction, z is the height coordinate above ground level, $k_0 = 2\pi/\lambda$ is the free-space wavenumber corresponding to wavelength λ, and n denotes the refractive index profile of the propagation medium [42].

This formulation enables terrain and atmospheric effects to be incorporated explicitly through boundary conditions and refractivity profiles. As a result, PEM is capable of accurately modeling key propagation phenomena such as diffraction, shadowing, and refraction over irregular terrain. This increased physical fidelity, however, comes at the cost of significantly higher computational complexity compared to empirical and hybrid propagation models. In the present study, PEM is therefore employed as a deterministic reference model against which the behavior and accuracy of the empirical and hybrid approaches are assessed. In this context, the term reference model denotes a fixed deterministic numerical benchmark rather than an experimentally verified ground truth. PEM is used because it resolves terrain-dependent propagation effects with higher physical detail than the considered empirical and hybrid models, but its outputs remain conditional on the adopted numerical discretization, boundary conditions, refractivity profile, antenna configuration, polarization, and implementation settings.

Between these two extremes lie the hybrid propagation models ITU-R P.452 and ITU-R P.1812, which combine empirical formulations with selected deterministic components.

The ITU-R P.452 model is primarily intended for point-to-point interference and coordination studies. It decomposes the total path loss into multiple components, including free-space loss, diffraction over terrain obstacles, tropospheric scatter, and anomalous propagation effects. Diffraction losses are computed using simplified physical models based on digital terrain profiles, while atmospheric effects are parameterized using climatological refractivity data. This structure enables ITU-R P.452 to capture key physical effects with moderate computational effort, making it suitable for long-range and interference-limited scenarios.

The ITU-R P.1812 model extends the hybrid concept toward point-to-area and high-resolution terrestrial coverage prediction. It combines empirical formulations with deterministic diffraction modeling and incorporates detailed digital elevation and clutter

data. Compared to ITU-R P.1546, it exhibits significantly higher terrain sensitivity, while remaining computationally tractable for large-scale simulations. In this study, ITU-R P.1812 represents an intermediate level of physical fidelity between empirical models and full-wave deterministic methods.

2. State of the Art, Research Gap and Contributions

In recent years, machine learning approaches have been playing an increasingly important role in predicting the propagation of radio-frequency waves [43,44]. One line of research focuses on fully data-driven prediction, where the model directly learns the relationship between the environment and geometry from measured or simulated outputs [45]. These methods often achieve good fits on the training dataset, but they are highly data-intensive, offer limited physical interpretability, and their generalization can be sensitive to domain shifts in the environment, such as a new terrain type or a significantly different topographical pattern [46,47]. Consequently, the role of input representation and feature engineering has become a prominent topic in the literature: beyond purely coordinate- or grid-based representations, physically motivated descriptors that summarize the environment and terrain are appearing with increasing frequency [48].

The other direction is the hybrid approach, which does not replace classical propagation models but rather complements or corrects them [49,50]. The basic idea is that fast, widely used empirical and hybrid models can be retained as a baseline, while the learning component is used exclusively to compensate for the missing physics, or, more precisely, for systematic deviations [51,52]. Residual or so-called delta learning approaches are particularly relevant: instead of predicting the full output, the learning algorithm aims to estimate as accurately as possible only the difference between the baseline and the reference model(s). The advantage is that the built-in structure of the baseline model (such as frequency and distance dependence, the effects of terrain conditions, etc.) is already implicitly present, so the learning algorithm has a smaller task to solve. This improves learnability and also reduces the risk of overfitting [53].

Deterministic wave-propagation models can be regarded as a reference in the literature, as they are capable of a more detailed treatment of terrain and propagation phenomena [54]. Their practical application, however, runs into computational limitations in large-scale network planning and spectrum-management tasks, such as those encountered by telecommunications regulatory authorities and service providers. This naturally leads to the basic idea that deterministic results can be effectively used as training or reference signals, making them suitable for shifting fast baseline models toward deterministic behavior [55].

Although an increasing number of studies have investigated the application of machine learning techniques to radiowave propagation modeling [56,57], most existing works focus on broad frequency ranges, measurement-driven datasets, or fully data-driven prediction schemes [58]. At present, there is a lack of studies specifically targeting the 3.6 GHz band, despite its widespread use in contemporary terrestrial mobile and industrial communication systems [59]. Moreover, the majority of published approaches either replace classical models entirely or address the correction of a single baseline model, without systematically establishing a relationship between a deterministic reference and multiple empirical and hybrid baseline models under realistic terrain conditions. In addition, many studies rely on a single learning paradigm, which limits insight into the robustness and generalizability of the reported improvements.

The present study addresses these gaps by proposing a unified delta-learning framework that explicitly targets the residual between a deterministic Parabolic Equation Modeling (PEM) reference and several widely used ITU-based empirical and hybrid propagation models in the 3.6 GHz band. The analysis is conducted in the 3.4–3.8 GHz frequency

range, which has been harmonized within the European Union [60] for terrestrial electronic communications services, with frequency arrangements and technical conditions defined in ECC Decision (11)06 [61]. According to the Radio Regulations [62], this band is allocated on a co-primary basis to mobile services, excluding aeronautical mobile applications, and is designated for mobile use in the European Table of Frequency Allocations and Applications (ECA Table) [63]. Beyond its regulatory relevance, this band represents a transitional propagation regime, offering a favorable balance between coverage and capacity while already exhibiting increased sensitivity to terrain features compared to traditional sub-GHz bands.

Within this framework, multiple machine-learning techniques are applied and evaluated in parallel, enabling a comparative assessment of their ability to learn systematic model deviations and to generalize across heterogeneous environments. The proposed approach is further validated across distinct topographical categories, including urban hilly, flat, and mountainous terrains, thereby providing a comprehensive evaluation of model performance, learning stability, and correction efficiency as functions of both terrain complexity and learning strategy.

3. Propagation Scenarios and Data Description

The objective of the study is to analyze realistic wave-propagation scenarios that place different levels of demand on the terrain sensitivity of empirical, hybrid, and deterministic propagation models. To this end, the simulations are based on real topographic data and cover markedly different geomorphological environments across the territory of Hungary. The spatial distribution and geographical context of the digital terrain data used in the investigations are illustrated in Figure 4. The figure presents the digital elevation model of Hungary and its immediate surroundings, which serves as the topographic basis for the subsequent simulations and terrain-profile-based analyses.

Figure 4. Digital elevation model of Hungary and its immediate surroundings based on the Copernicus GLO-30 dataset.

For terrain modeling, we used data from the Copernicus GLO-30 Digital Elevation Model (DEM), which were accessible via the OpenTopography portal [64]. Copernicus GLO-30 is a global digital elevation model of uniform quality which, with its spatial resolution of approximately 30 m, is suitable for terrain-profile-based radiowave propagation studies. The dataset is available in the WGS84 coordinate reference system and provides uniform-quality elevation data with extensive spatial coverage, making it suitable for generating terrain profiles (longitudinal sections) for radiowave propagation studies.

From the digital elevation model, straight-line terrain profiles were generated along the links between the transmitter and the receiver using a custom-developed Python-based processing script. The script uses sub-areas extracted from the Copernicus GLO-30 digital elevation model as input and then randomly generates coordinate pairs within the study area such that the transmitter–receiver distance is fixed at 30 km in all cases.

Between the resulting coordinate pairs, terrain profiles are extracted with a sampling interval of 50 m, ensuring an adequately resolved representation of terrain details for subsequent wave-propagation calculations. The script allows parameterized control over the number of generated profiles, thereby ensuring that a sufficient and statistically meaningful number of samples is available for machine-learning-based training and testing procedures.

The extracted terrain profiles underwent a uniform preprocessing step and were used as identical inputs for deterministic Parabolic Equation Modeling (PEM) simulations, for the empirical and hybrid baseline models, and for the delta-learning-based correction framework, ensuring direct and consistent comparability of the results produced by the different models.

The model evaluation was conducted in three terrain-wise clearly distinct environments, which represent a broad spectrum of propagation conditions in Hungary, as illustrated in Figure 5.

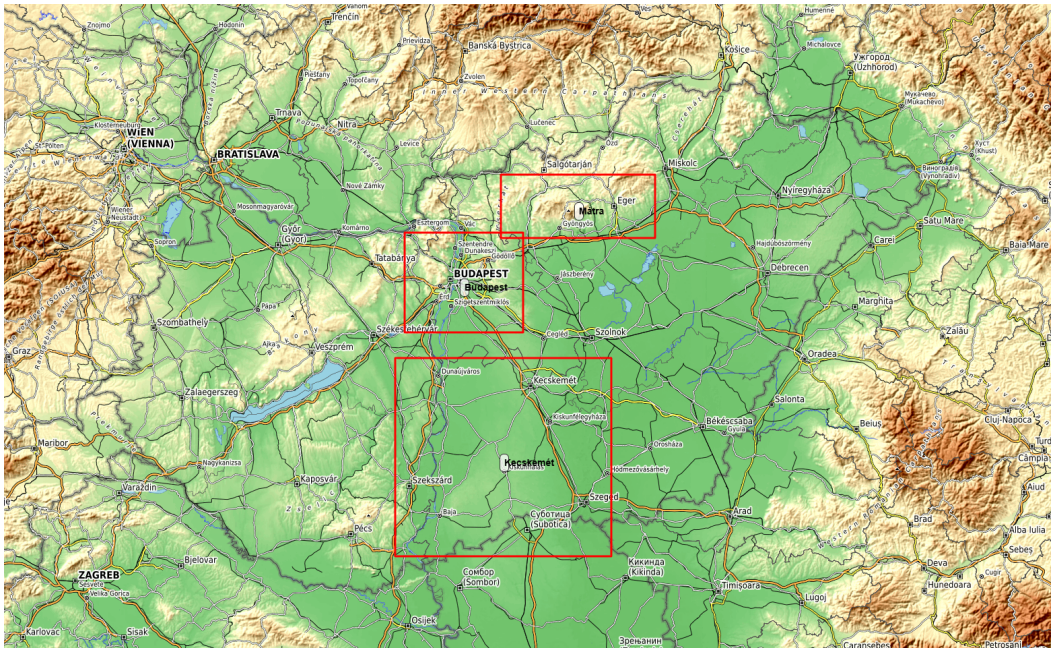


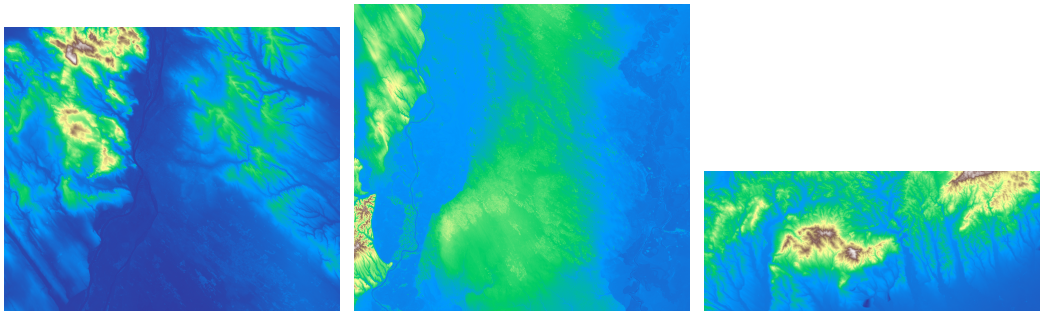
Figure 5. Geographical extent of the selected propagation scenarios highlighted with red boxes.

These environments were deliberately selected to span markedly different levels of terrain complexity, enabling a systematic investigation of how empirical, hybrid, and deterministic propagation models respond to topographical variations:

- Urban–hilly environment (Budapest region): This scenario represents an urbanized area characterized by moderately undulating terrain. The combination of fine-scale elevation variations and complex geometrical conditions makes this environment suitable for assessing local terrain sensitivity and the ability of the models to capture diffraction and shadowing effects in built-up surroundings.
- Flat environment (Kecskemét region): The flatland scenario is characterized by nearly homogeneous elevation conditions with minimal topographical variability. It serves as a reference environment in which empirical propagation models are expected to perform more favorably, and where learning-based corrections primarily mitigate systematic model deviations rather than pronounced terrain-induced effects.

- Mountainous environment (Mátra region): The mountainous scenario exhibits significant elevation differences, steep slopes, and highly complex terrain geometry. These conditions pose a substantial challenge for empirical and hybrid models and provide a suitable test case for evaluating the extent to which delta-based correction toward a deterministic reference improves prediction accuracy under strong topographical complexity.

To further characterize the topographical properties of the investigated environments, Figure 6 presents elevation heatmaps derived from the Copernicus GLO-30 DEM for each selected scenario.



(**a**) Budapest region (**b**) Kecskemét region (**c**) Mátra region

Figure 6. Elevation heatmaps of the scenarios derived from the Copernicus GLO-30 DEM.

These heatmaps provide a spatial overview of the elevation distribution within the selected areas and highlight the increasing level of terrain variability from the flatland environment through the urban–hilly region to the mountainous scenario. The visual comparison of the elevation patterns complements the scenario definitions by illustrating the differing geomorphological characteristics that influence wave propagation behavior and, consequently, the performance of the evaluated propagation models.

To complement the qualitative visual analysis provided by the elevation heatmaps, and to ensure reproducibility, a quantitative summary of the investigated terrain scenarios is given in Table 2. The table reports the geographical boundaries of the selected areas together with key elevation-based statistics, providing a compact numerical characterization of the terrain complexity across the flat, urban–hilly, and mountainous environments. These metrics serve as a consistent basis for interpreting the subsequent propagation modeling results and for relating model performance to underlying topographical properties.

Table 2. Geographical boundaries and key statistics of the investigated terrain scenarios.

Region	$\mathbf{X_{min}}$	$\mathbf{Y_{min}}$	$\mathbf{X_{max}}$	$\mathbf{Y_{max}}$
Budapest	18.6685	47.2029	19.6353	47.7542
Kecskemét	18.5889	45.9513	20.3522	47.0628
Mátra	19.4540	47.7228	20.7079	48.0717
Metrics	**Budapest**	**Kecskemét**	**Mátra**	
Min elevation [m]	84.46	52.84	25.54	
Max elevation [m]	761.30	319.01	1034.34	
Elevation range [m]	676.84	266.17	1008.80	
Mean elevation [m]	173.05	106.68	264.36	
Standard deviation [m]	81.17	23.26	147.2	
Mean slope [°]	4.06	1.88	7.27	

Finally, Figure 7 illustrates the general wave-propagation scenario and the data flow applied throughout the study. For each generated terrain profile, identical geometric

and environmental inputs are provided to the empirical and hybrid baseline propagation models, as well as to the deterministic Parabolic Equation Modeling (PEM) solver. The resulting propagation outputs are evaluated along the same transmitter–receiver paths, ensuring strict consistency across all modeling approaches. The investigated transmitting antenna (base station or source) is assumed in all cases to be 30 m above ground level with an effective radiated power of 46 dBm, which are considered an average value in the examined 3.6 GHz frequency band. The receiving point is fixed at a height of 3 m above ground level in all cases, and at this height the resulting field strength level (i.e., the path loss) is examined at distances of up to a maximum of 30 km from the base station (source).

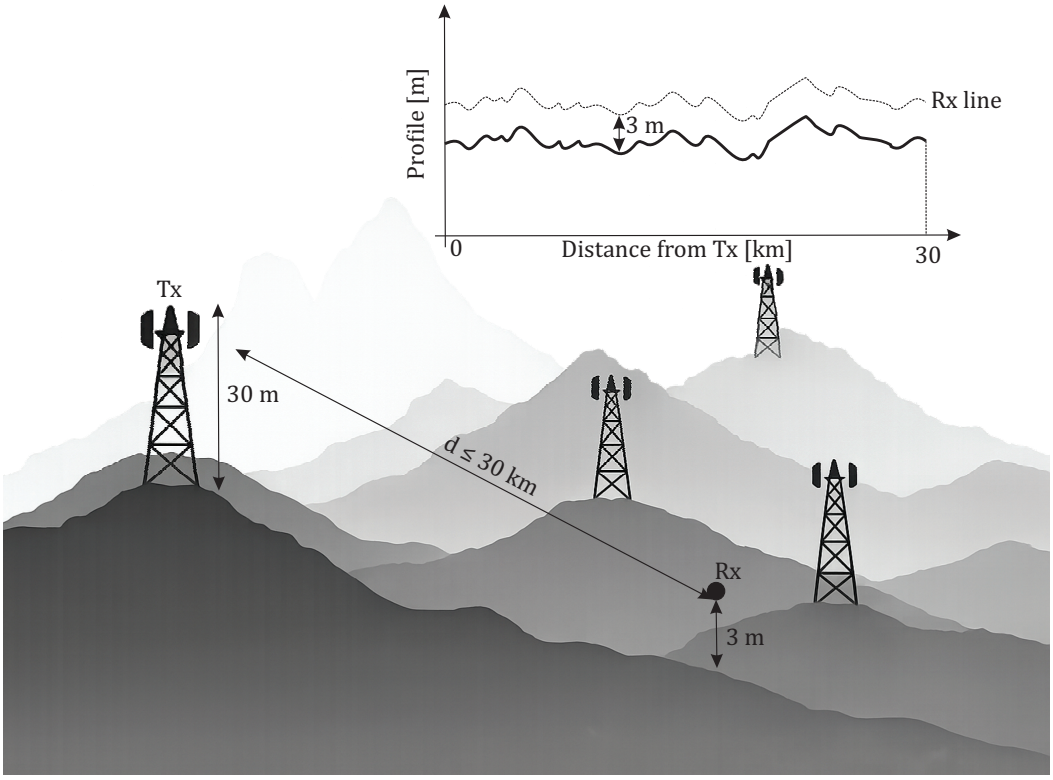


Figure 7. Conceptual illustration of the wave-propagation scenario. Identical terrain profiles and geometric inputs are used for all propagation models.

These one-dimensional curves will serve as input parameters during the training of the models, and naturally, results of the same type will also be used as references during the execution and validation of the simulations.

Our investigations were carried out by following the previously described methodology. For the ITU-R P.1546, ITU-R P.452, and ITU-R P.1812 propagation models, the time percentage was uniformly set to 10%, while the location percentage was fixed at 50%. In the case of the ITU-R P.452 model, the prediction type was set to Average Year. The surface and atmospheric parameters were defined as follows: the temperature was set to 20 °C, the atmospheric pressure to 1013.25 hPa, and the water vapour density to 8 g/m^2. Furthermore, the atmospheric refractivity (N_0) was set to 330 N-units, and the refractivity index lapse rate was set to 45 N-units/km. The same set of parameters was applied in the ITU-R P.1812 model and during the line-based calculations in order to ensure comparability between the different propagation models. For the simulations, iterative MATLAB implementations of the propagation models were applied. The PEM simulations were computed in an iterative manner using PETOOL v2.0 [65], which had been reverse-engineered during our previous research work [42].

It should be emphasized that the validation performed in this study is simulation-based. No independent field measurement campaign was used for model validation. The deterministic PEM results are employed as a high-fidelity numerical reference, and the learning task is formulated as the correction of empirical and hybrid baseline models toward this deterministic reference. Accordingly, the reported performance improvements quantify the ability of the proposed framework to reproduce PEM-like predictions for previously unseen terrain profiles within the considered terrain categories, rather than its direct agreement with real-world measurement data.

It should also be clarified that PEM is not treated in this work as an absolute ground truth in the experimental sense, but as a deterministic numerical reference generated under a fixed and reproducible simulation configuration. The purpose of using PEM is to provide a physically more detailed reference solution than the empirical and hybrid baseline models, especially with respect to terrain-dependent diffraction, shadowing, and refractive effects. Nevertheless, PEM predictions are themselves dependent on numerical and physical modeling choices, including the range–height discretization, absorbing boundary treatment, terrain-profile representation, refractivity assumptions, antenna parameters, polarization, and the specific implementation of the propagation solver. Therefore, the reported errors and improvements should be interpreted as deviations relative to the adopted PEM configuration rather than as absolute propagation errors with respect to the real environment.

4. Delta Learning Framework

Consider N radio links of length 30 km, each of which can be characterized by a terrain profile. Along the profile, the terrain elevation and the wave-propagation quantities are given at a discrete set of L sample points.

Let

$$i \in \{1, \ldots, N\} \text{ and } j \in \{1, \ldots, L\}$$

denote the profile and the sample-point indices, respectively. At the distance point d_j, the terrain elevation $h_i(d_j) \in \mathbb{R}$, the output field strength of the baseline wave-propagation model $b_i(d_j) \in \mathbb{R}$, and the field strength of the reference wave-propagation model $y_i(d_j) \in \mathbb{R}$ are available. In matrix form:

$$\mathbf{H} = [h_i(d_j)], \mathbf{B} = [b_i(d_j)], \mathbf{Y} = [y_i(d_j)], \qquad \mathbf{H}, \mathbf{B}, \mathbf{Y} \in \mathbb{R}^{N \times L}.$$

In our study, we work with a fixed raster of 50 m, from which it immediately follows that $L = 601$.

According to the basic assumption of delta learning, the reference can be expressed as the sum of the baseline and an additive correction:

$$y_i(d_j) = b_i(d_j) + \delta(d_j), \tag{8}$$

where the residual field strength level is

$$\delta_i(d_j) = y_i(d_j) - b_i(d_j), \text{ in matrix form: } \boldsymbol{\Delta} = \mathbf{Y} - \mathbf{B}.$$

It is clear that, in what follows, the learning task concerns exclusively the field $\boldsymbol{\Delta}$ [66].

In constructing the model, we started from the assumption that the systematic error of the baseline model at a given point depends primarily on the local terrain geometry. To formalize this, we introduce a centered, finite-extent sliding window with which we scan

all N terrain profiles of length L. To formalize this, we introduce a centered, finite-extent sliding window as follows:

$$\mathcal{W}_j = \{k \in \{1, \ldots, L\} \mid |k - j| \leq w\}, \qquad w = 5. \tag{9}$$

It can be seen that the maximum size of the sliding window is

$$|\mathcal{W}_j| = 2w + 1 = 11,$$

while at the edges of the terrain profiles, $|\mathcal{W}_j| < 11$ without extrapolation.

To illustrate this, Figure 8 shows a local window around a given distance point, highlighting the center point and the ± 5-point neighborhood. Within this framework, all input features are defined on a per-profile, per-distance-point basis and are constructed exclusively from quantities that are available both during training and inference.

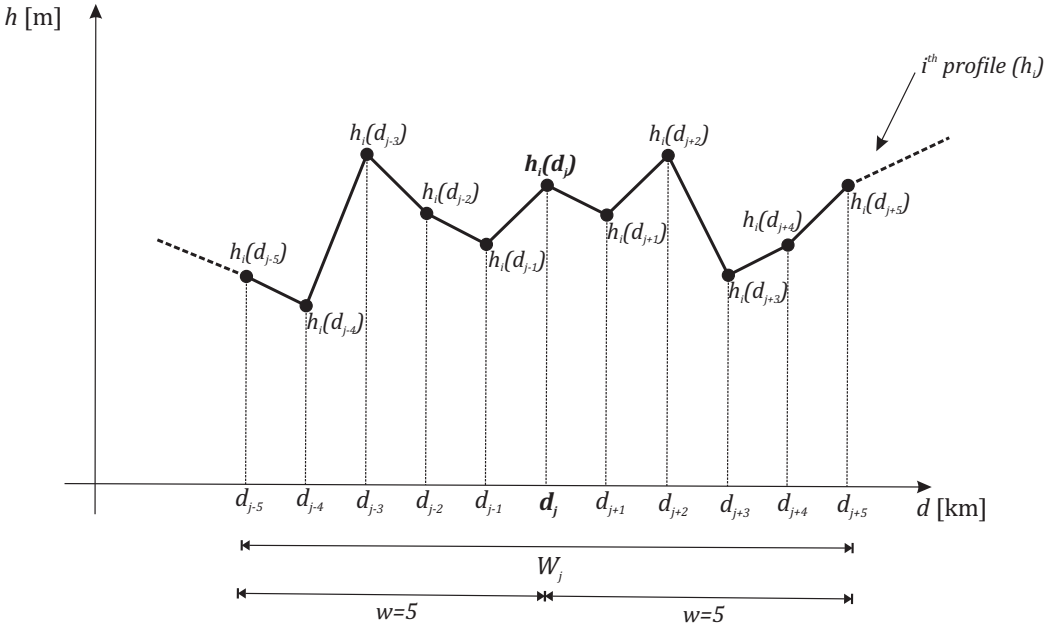


Figure 8. Illustration of $\mathcal{W}_j$ sliding window on the terrain profile.

Formally, for the i-th profile and the distance point d_j, a feature vector $\boldsymbol{\varphi}_{i,j} \in \mathbb{R}^8$ is assigned, which encodes the local terrain geometry, the baseline model output, and the relative position along the propagation path. The individual components of the feature vector are defined as follows.

First, the baseline propagation model output at the given location, $b_i(d_j)$, is included explicitly as a feature. This term represents the first-order physical or empirical estimate of the propagation loss and provides a natural reference level for the learned correction. Second, the absolute terrain height at the evaluation point, $h_i(d_j)$, is used to capture large-scale elevation effects, such as ridge tops or valley floors, which may not be fully accounted for by the baseline model. To characterize the local terrain geometry, discrete spatial derivatives of the terrain profile are introduced. The first derivative (local slope) is approximated by a central difference scheme,

$$s_i(d_j) \approx \frac{h_i(d_{j+1}) - h_i(d_{j-1})}{2\Delta d}, \tag{10}$$

where Δd denotes the spatial sampling interval of the terrain profile. This quantity describes whether the terrain is locally rising or falling and provides information related to diffraction and shadowing effects. The second derivative (local curvature) is approximated as

$$c_i(d_j) \approx \frac{h_i(d_{j+1}) - 2h_i(d_j) + h_i(d_{j-1})}{(\Delta d)^2}. \tag{11}$$

The curvature allows the model to distinguish between locally convex and concave terrain features, even when absolute heights are similar.

In addition to pointwise geometric descriptors, several local statistical measures are computed over the sliding window $\mathcal{W}_j$. The local mean terrain height is defined as

$$\overline{h}_i(d_j) = \frac{1}{|\mathcal{W}_j|} \sum_{k \in \mathcal{W}_j} h_i(d_k), \tag{12}$$

which represents the average elevation level in the immediate neighborhood of the evaluation point and reduces sensitivity to small-scale fluctuations.

The local terrain variability is quantified by the standard deviation

$$\sigma_h(d_j) = \sqrt{\frac{1}{|\mathcal{W}_j|} \sum_{k \in \mathcal{W}_j} \left(h_i(d_k) - \overline{h}_i(d_j) \right)^2}, \tag{13}$$

which serves as a measure of terrain roughness and captures how strongly the elevation varies within the local window.

To further describe the geometric complexity of the terrain, the mean absolute slope within the window is introduced:

$$\overline{|s|}_i(d_j) = \frac{1}{|\mathcal{W}_j|} \sum_{k \in \mathcal{W}_j} |s_i(d_k)|. \tag{14}$$

Unlike the pointwise slope, this quantity provides a robust indicator of the overall steepness of the surrounding terrain, independent of slope direction.

Finally, the relative position of the evaluation point along the propagation path is included through a normalized distance coordinate,

$$\frac{d_j}{d_{\max}}, \tag{15}$$

where $d_{\max}$ denotes the total path length. This feature enables the model to account for systematic position-dependent effects along long paths.

Collecting all components, the complete local feature vector used for delta learning at profile i and distance point d_j is given by

$$\boldsymbol{\varphi}_{i,j} = \begin{bmatrix} b_i(d_j) \\ h_i(d_j) \\ s_i(d_j) \\ c_i(d_j) \\ \overline{h}_i(d_j) \\ \sigma_h(d_j) \\ \overline{|s|}_i(d_j) \\ d_j / d_{\max} \end{bmatrix} \in \mathbb{R}^8. \tag{16}$$

This feature representation combines pointwise and window-based descriptors to capture both local terrain geometry and broader contextual information, while remaining fully compatible with the delta learning formulation and applicable to unseen terrain profiles [67].

While the feature representation defined in Equation (16) captures the local terrain geometry and pointwise characteristics of the propagation path, the systematic deviation of baseline propagation models may also depend on the global structure of the terrain profile as a whole. In particular, large-scale elevation variations, overall terrain roughness, and the statistical characteristics of the baseline response may influence the magnitude and spatial structure of the residual field.

To incorporate such profile-level contextual information, the local feature representation is extended with a set of global descriptors that characterize the entire terrain profile. These quantities depend only on the profile index i and are independent of the distance sample d_j. Consequently, they provide a global context shared by all points of a given propagation path.

Formally, for each terrain profile i, a global descriptor vector $\mathbf{g}_i \in \mathbb{R}^{10}$ is defined, and the extended feature vector takes the form

$$\boldsymbol{\phi}_{i,j} = \begin{bmatrix} \boldsymbol{\varphi}_{i,j} \\ \mathbf{g}_i \end{bmatrix} \in \mathbb{R}^{18}, \tag{17}$$

where $\boldsymbol{\varphi}_{i,j} \in \mathbb{R}^8$ denotes the local feature vector defined in (16). The global descriptors are constructed from the terrain profile $h_i(d_j)$ and the baseline model output $b_i(d_j)$ as follows.

The first group of global features describes the statistical properties of the terrain elevation along the entire propagation path. The mean terrain elevation is defined as

$$g_i^{(1)} = \frac{1}{L} \sum_{j=1}^{L} h_i(d_j), \tag{18}$$

while the standard deviation of the elevation is given by

$$g_i^{(2)} = \sqrt{\frac{1}{L} \sum_{j=1}^{L} \left(h_i(d_j) - g_i^{(1)} \right)^2}. \tag{19}$$

The overall elevation range (relief) of the terrain profile is defined as

$$g_i^{(3)} = \max_{1 \le j \le L} h_i(d_j) - \min_{1 \le j \le L} h_i(d_j). \tag{20}$$

To characterize the overall geometric complexity of the terrain, global descriptors derived from the local slope $s_i(d_j)$ are introduced. The mean absolute slope is defined as

$$g_i^{(4)} = \frac{1}{L} \sum_{j=1}^{L} |s_i(d_j)|, \tag{21}$$

while the upper range of terrain steepness is characterized by the 90th percentile of the absolute slope distribution,

$$g_i^{(5)} = \text{percentile}_{90}\left(|s_i(d_j)| \right). \tag{22}$$

Furthermore, the cumulative elevation gain (total climb) and cumulative descent along the propagation path are defined as

$$g_i^{(6)} = \sum_{j=1}^{L-1} \max\big(h_i(d_{j+1}) - h_i(d_j), 0\big), \tag{23}$$

$$g_i^{(7)} = \sum_{j=1}^{L-1} \max\big(h_i(d_j) - h_i(d_{j+1}), 0\big). \tag{24}$$

Finally, global statistical descriptors of the baseline propagation model output are included to capture systematic characteristics of the baseline response along the entire path. The mean baseline response is given by

$$g_i^{(8)} = \frac{1}{L} \sum_{j=1}^{L} b_i(d_j), \tag{25}$$

its standard deviation by

$$g_i^{(9)} = \sqrt{\frac{1}{L} \sum_{j=1}^{L} \Big(b_i(d_j) - g_i^{(8)}\Big)^2}, \tag{26}$$

and the overall dynamic range by

$$g_i^{(10)} = \max_j b_i(d_j) - \min_j b_i(d_j). \tag{27}$$

Since the global descriptors depend only on the terrain profile index i, they are identical for all distance samples of a given profile. In the training data, these quantities are therefore associated with each profile–distance pair by replication, ensuring that the regression model receives both local geometric information and global terrain context.

The resulting feature representation combines pointwise descriptors of local terrain geometry with large-scale profile characteristics, providing a richer representation of the propagation environment while remaining fully compatible with the delta-learning formulation.

It is important to clarify the intended role of the proposed 18-dimensional feature representation. The feature vector is not designed to serve as a complete surrogate representation of the deterministic electromagnetic field, nor is it expected to explicitly encode all wave phenomena represented by the PEM solver, such as detailed diffraction, reflection, interference, or small-scale multipath effects. Instead, its role is to provide a compact, physics-informed description of the terrain and baseline-model response that is sufficient for learning the dominant systematic component of the baseline-PEM discrepancy.

This interpretation follows directly from the delta-learning formulation adopted in this work. The learning algorithm does not predict the full PEM reference field from the terrain profile alone. Rather, the empirical or hybrid propagation model first provides a baseline prediction that already incorporates distance dependence, antenna-height effects, frequency dependence, and, depending on the model, simplified terrain-dependent propagation mechanisms. The machine-learning model is then trained only to estimate the residual correction between this baseline prediction and the PEM reference. Consequently, the feature representation is required to characterize the systematic error structure of the baseline model, not the complete physical propagation process.

Although some of the descriptors are statistical in nature, the representation is not limited to mean elevation and standard deviation. It combines pointwise local terrain height, local slope, local curvature, sliding-window statistics, mean absolute local slope,

normalized propagation distance, global elevation relief, slope percentiles, cumulative elevation gain and descent, and statistical descriptors of the baseline response along the full profile. These quantities jointly encode local geometry, profile-level terrain context, and the behavior of the baseline propagation model.

The use of such a compact representation is a deliberate trade-off between physical completeness, interpretability, data efficiency, and computational cost. A higher-dimensional input, such as the full terrain profile or a two-dimensional environmental grid, could in principle retain more information about complex propagation mechanisms. However, it would also increase the dimensionality of the learning problem, require substantially more training data, reduce interpretability, and increase the risk of overfitting, especially under profile-level validation. Since the present work targets fast delta correction for repeated large-scale inference, the proposed low-dimensional feature representation was selected to capture the dominant residual patterns while maintaining robust trainability and low inference cost.

Similarly, convolutional, recurrent, or attention-based neural-network architectures operating directly on raw terrain profiles could in principle be used to learn spatial dependencies along the propagation path. A systematic benchmark of such raw-profile deep-learning models is, however, outside the scope of the present study, which focuses on an interpretable, low-dimensional, physics-informed delta-correction framework applicable across multiple empirical and hybrid baseline models.

Accordingly, the remaining prediction error after correction may partly originate from propagation phenomena that are not fully represented by the proposed descriptors. This limitation should be considered when interpreting the results, particularly in highly complex terrain where diffraction, shadowing, and interference effects may vary rapidly along the propagation path.

Following this, the delta learning task can be formulated in a mathematically unambiguous manner. Given that each terrain profile contains multiple distance samples that are in close spatial relationship with one another, the learning problem can naturally be interpreted at the level of individual profile–distance-point pairs. For the sake of simplifying the notation, we introduce the composite index $n = (i, j)$, where i denotes the index of the terrain profile, while j denotes the index of the distance point within the given profile. Accordingly, the residual (delta) value—interpreted as the difference between the reference and the baseline model—as well as its associated feature vector can be expressed as follows:

$$\delta_n = \delta_i(d_j), \qquad \boldsymbol{\phi}_n = \boldsymbol{\phi}_{i,j}. \tag{28}$$

With this notation, the complete training set can be defined in the following compact form:

$$\mathcal{D} = \{(\boldsymbol{\phi}_n, \delta_n)\}_{n=1}^{NL}, \tag{29}$$

where N denotes the number of (training) terrain profiles used, and L denotes the number of distance samples within a single profile, which is fixed to 601 in our investigations, resulting from the analysis of 30 km sections using a 50 m raster. This formulation emphasizes that the learning task consists of a large number of samples that are structured, where the individual samples are not independent of one another.

Let

$$g : \mathbb{R}^{18} \to \mathbb{R}, \tag{30}$$

be an arbitrary regression model defined as a mapping that projects the previously defined eight-dimensional feature vector onto a scalar correction value. The goal of delta learning is to determine a function $\hat{g}$ that approximates the residual field as accurately

as possible in the sense of supervised learning. This leads to the following empirical risk minimization problem [68,69]:

$$\hat{g} = \arg\min_{g} \sum_{n=1}^{NL} \ell(\delta_n, g(\phi_n)) + \Omega(g), \tag{31}$$

where $\ell(\cdot)$ denotes a pointwise loss function, while $\Omega(g)$ is a regularization term that controls the complexity of the model.

It is important to note that this formulation is mathematically equivalent to minimizing the reconstruction error of the reference field. Since $\delta_n = y_n - b_n$, by substituting this into the previous expression and assuming that $\ell(\cdot)$ is shift-invariant,

$$\hat{g} = \arg\min_{g} \sum_{n=1}^{NL} \ell(y_n, b_n + g(\phi_n)) + \Omega(g), \tag{32}$$

can be written, which clearly expresses that the learned model does not replace the baseline estimate but is coupled to it as an additive correction.

After training the regression function $\hat{g}$, the estimation of the reference field for new terrain profiles can be carried out directly. At distance point d_j of the i-th profile, the estimated reference value is given by

$$\hat{y}_i(d_j) = b_i(d_j) + \hat{g}(\phi_{i,j}), \tag{33}$$

which, in matrix form, takes the following shape:

$$\hat{\mathbf{Y}} = \mathbf{B} + \hat{\mathbf{\Delta}}, \tag{34}$$

where $\hat{\mathbf{Y}}$ denotes the residual field produced by the learned model. The additive structure ensures that the physical or empirical foundations embedded in the baseline model are preserved, while systematic deviations are corrected in a data-driven manner. The model we propose thus provides, in this novel form, the possibility of more accurate estimation.

At this point, it is worth highlighting two additional aspects. The first concerns the statistical interpretation of delta learning. The framework presented here implicitly assumes the following decomposition:

$$y_i(d_j) = b_i(d_j) + \delta_{sys}\left(\boldsymbol{\phi}_{i,j}\right) + \varepsilon_i(d_j), \tag{35}$$

where $\delta_{sys}(\cdot)$ represents the feature-dependent, deterministic systematic bias of the baseline model, while $\varepsilon_i(d_j)$ denotes a non-modelable random component that encompasses measurement noise, errors arising from numerical approximations, and other physical effects neglected due to additional considerations.

In this context, the objective of $\hat{g}$ is to approximate the δ_{sys} component, whereas the term $\varepsilon_i(d_j)$ defines the lower bound on the achievable accuracy.

The second aspect concerns validation considerations: samples taken along terrain profiles are strongly correlated; therefore, the residual values associated with consecutive distance points cannot be regarded as statistically independent:

$$\delta_i(d_j) \not\perp \delta_i(d_{j+1}). \tag{36}$$

This correlation implies that a training–test split based on random selection of individual distance points would lead to biased, overly optimistic performance estimates. To avoid this, training and validation must be performed at the profile level, meaning that

entire terrain profiles are assigned exclusively to either the training or the test set. This approach provides a more realistic assessment of the model's generalization capability for new, previously unseen propagation paths.

By combining the feature representation, the optimization objective, and the reconstruction step, the proposed delta learning approach can be summarized in the following compact form:

$$\widehat{\mathbf{Y}} = \mathbf{B} + \widehat{g}(\Phi(\mathbf{H}, \mathbf{B})), \tag{37}$$

where the mapping $\Phi(\cdot)$ constructs the feature vectors from the terrain profiles and the baseline model outputs. In this way, the learned model preserves the physical foundations of the baseline approach, while reducing its systematic errors in a data-driven manner [70].

It is important to emphasize that the unified framework in this study refers to a common correction methodology and data-processing pipeline rather than to a single regression model shared across all baseline propagation models and terrain conditions. The same delta-learning formulation, feature construction procedure, PEM reference generation, profile-level validation strategy, and evaluation metrics are applied consistently throughout the study. However, the residual structure between a baseline propagation model and the PEM reference is strongly dependent on both the baseline model and the morphology of the terrain profile. In practical propagation-simulation environments, the topographic data used for terrain profile extraction are typically accompanied by morphological information describing the character of the terrain and land environment.

Such MORPHO data may distinguish between several detailed categories, such as urban, suburban, agricultural, forested, flat, hilly, or mountainous areas. Although these categories may be more detailed than the three broad terrain classes considered in this work, they can be mapped by a preliminary classification step into the disjoint terrain groups used in the present study: urban–hilly, flat, and mountainous environments.

Let m_i denote the morphological descriptor associated with the i-th terrain profile, and let

$$\kappa : \mathcal{M} \to \mathcal{C} \tag{38}$$

be a terrain-class assignment function that maps the available morphological information from the original MORPHO class set $\mathcal{M}$ to the broader terrain-class set

$$\mathcal{C} \;=\; \{\text{urban-hilly}, \text{flat}, \text{mountainous}\}. \tag{39}$$

For a new terrain profile, the corresponding terrain class is therefore obtained as

$$c_i \;=\; \kappa(m_i), \qquad c_i \in \mathcal{C}. \tag{40}$$

Within this class-conditioned framework, separate correction functions are trained for the relevant baseline model and terrain class. If q denotes the baseline propagation model, the corresponding correction function can be written as

$$\hat{g}_{q,c} : \mathbb{R}^{18} \to \mathbb{R}, \qquad q \in \mathcal{Q}, \quad c \in \mathcal{C}, \tag{41}$$

where $\mathcal{Q}$ denotes the set of investigated baseline propagation models. During inference, the terrain class of the new profile is first determined from the available MORPHO/TOPO information, and the correction model trained for the corresponding class is selected. The corrected prediction is then obtained as

$$\hat{y}_i^{(q)}(d_j) \;=\; b_i^{(q)}(d_j) + \hat{g}_{q,c_i}\left(\boldsymbol{\phi}_{i,j}^{(q)}\right), \qquad c_i \;=\; \kappa(m_i). \tag{42}$$

This formulation provides a unified operational workflow while avoiding the direct pooling of heterogeneous residual distributions from physically different terrain conditions. A single regressor trained jointly across all terrain classes and baseline models would need to learn residual patterns associated with flat, urban–hilly, and mountainous profiles simultaneously. Since these environments exhibit substantially different diffraction, shadowing, and terrain-sensitivity characteristics, such pooling may increase the risk of reduced accuracy or negative transfer. The proposed morphology-aware selection step therefore allows the framework to use the correction model that is most appropriate for the terrain profile under investigation, while preserving a common delta-learning formulation across all cases.

Accordingly, the framework should be interpreted as a class-conditioned correction approach rather than as a single universal model intended for direct transfer between arbitrary terrain morphologies.

For reproducibility, the model-selection settings of the non-linear regressors are summarized here. For the Gradient Boosting Regressor, a limited grid search was performed using profile-level 5-fold inner cross-validation on the training pool. The candidate values were 200 and 400 learning cycles, learning rates of 0.05 and 0.10, and maximum tree splits of 10 and 30. For the Multilayer Perceptron, a limited architecture search was performed with the number of neurons in the first hidden layer selected from 20, 40 and the number of neurons in the second hidden layer selected from 0, 20, 40, where 0 denotes a single-hidden-layer network. The remaining MLP settings were fixed: tansig hidden-layer activations, a linear output layer, scaled conjugate-gradient training (trainscg), MSE loss, and a maximum of 300 epochs. In both cases, the selected configuration was the one minimizing the mean validation MAE of the reconstructed field-strength prediction after delta correction.

It should be emphasized that the generalization assessed in this work is profile-level generalization within the considered terrain categories. Complete terrain profiles are separated between training and testing, so that all distance samples belonging to a given propagation path are assigned exclusively to one subset. This prevents information leakage caused by the strong spatial correlation of neighboring samples along the same profile and provides a realistic assessment for previously unseen propagation paths.

However, this validation setting does not imply unrestricted transfer across fundamentally different terrain morphologies or geographical domains. A correction model trained only on flat terrain is not expected to fully capture the residual structure observed in mountainous terrain, where diffraction, shadowing, and elevation-induced effects are substantially stronger. For this reason, the proposed framework adopts a morphology-aware structure in which separate correction functions are trained for the relevant terrain classes and selected during inference using available MORPHO/TOPO information.

Consequently, the present results demonstrate profile-level generalization within the investigated flat, urban–hilly, and mountainous terrain categories, but they do not establish cross-terrain or cross-country transferability. The latter would require additional geographically independent validation datasets and is therefore left for future work.

5. Computational Cost Considerations

The proposed delta-learning framework relies on deterministic Parabolic Equation Modeling results as reference data. Consequently, the computational cost of generating the PEM-based training labels is not eliminated by the proposed method and must be clearly distinguished from the cost of applying the trained correction model.

As discussed in our previous work [19], PEM is computationally substantially more demanding than empirical or hybrid propagation models. The main reason is that PEM requires the discretization of the full two-dimensional computational domain in the range-

height plane. Let N_x denote the number of discretization points in the horizontal propagation direction and N_z the number of discretization points in the vertical direction. The numerical solution must be obtained on this complete grid in order to enforce the required boundary conditions and to close the computational domain, even if the final quantity of interest is needed only along a receiver-height line or at a finite number of receiver locations.

The computational burden is strongly affected by the wavelength. As the frequency increases, the wavelength decreases, and a finer spatial discretization is required to preserve numerical accuracy. In addition, the maximum propagation distance determines the horizontal size of the computational domain, while the terrain-elevation range and the chosen upper boundary determine its vertical size. Therefore, hilly or mountainous profiles typically require a larger computational domain than flat terrain profiles, which increases both memory consumption and computation time. Depending on the numerical solution technique, the approximate computational complexity of PEM can be expressed as $\mathcal{O}(N \log N)$ for Split-Step Fourier Method (SSFM)-based implementations, where N denotes the number of grid samples involved in the transform operation, or as $\mathcal{O}(N_x N_z)$ for finite-difference formulations such as Crank–Nicolson-type schemes. In contrast, empirical and hybrid models usually require only curve interpolation, closed-form evaluations, or one-dimensional terrain-profile processing, and therefore have substantially lower marginal computational cost.

For this reason, the computational efficiency of the proposed framework should be interpreted in an amortized sense. Let C_{PEM} denote the cost of generating one PEM reference profile, C_{base} the cost of evaluating the empirical or hybrid baseline model, C_{feat} the cost of feature extraction, C_{ML} the cost of evaluating the trained correction model, and C_{train} the cost of training the regression model. For N_{train} training profiles and N_{pred} new prediction profiles, the total cost of the proposed workflow can be written schematically as

$$C_{\mathrm{total}} = N_{\mathrm{train}}(C_{\mathrm{PEM}} + C_{\mathrm{base}} + C_{\mathrm{feat}}) + C_{\mathrm{train}} + N_{\mathrm{pred}}(C_{\mathrm{base}} + C_{\mathrm{feat}} + C_{\mathrm{ML}}). \tag{43}$$

By contrast, direct deterministic prediction for all new profiles would require approximately

$$C_{\mathrm{direct}} = N_{\mathrm{pred}} C_{\mathrm{PEM}}. \tag{44}$$

Thus, the proposed approach does not remove the initial PEM cost associated with the generation of reference labels. Instead, it transfers this cost to an offline training stage. After the correction model has been trained, new propagation profiles can be evaluated using only the fast baseline model, the feature-extraction step, and the learned correction. The marginal inference cost is therefore

$$C_{\mathrm{infer}} = C_{\mathrm{base}} + C_{\mathrm{feat}} + C_{\mathrm{ML}}, \tag{45}$$

which is substantially lower than C_{PEM} for repeated large-scale evaluations.

The break-even condition can be expressed by comparing the one-time offline cost with the per-profile savings obtained during inference. Neglecting the comparatively small baseline and feature-extraction costs during PEM label generation, the proposed framework becomes computationally advantageous when

$$N_{\mathrm{pred}}[C_{\mathrm{PEM}} - (C_{\mathrm{base}} + C_{\mathrm{feat}} + C_{\mathrm{ML}})] > N_{\mathrm{train}} C_{\mathrm{PEM}} + C_{\mathrm{train}}. \tag{46}$$

This condition emphasizes that the method is not primarily intended to reduce the cost of a small number of isolated deterministic simulations. Its advantage arises when the trained delta-correction model is reused for many new terrain profiles, as in large-scale network planning, coverage studies, or spectrum-management workflows. In this sense, the

framework provides low marginal inference cost after an offline PEM-based reference-data generation stage, rather than eliminating the computational expense of PEM itself.

A qualitative summary of the computational role of each modeling component is provided in Table 3.

To make the computational-efficiency discussion more explicit, Table 4 provides an algorithmic comparison of the baseline, deterministic, and ML-corrected propagation workflows. The purpose of this comparison is not to report implementation-specific wall-clock runtimes, but to characterize the dominant computational operations that determine the marginal cost of applying each model to a new terrain profile. This distinction is important because absolute runtimes depend strongly on the hardware platform, available memory, software environment, programming language, numerical libraries, compiler settings, parallelization strategy, and the degree of optimization of the individual implementations. In particular, a direct timing comparison between a PEM solver and empirical or hybrid propagation-model codes may be misleading if the implementations are not optimized to the same extent or if they use different numerical backends.

Table 3. Qualitative comparison of the computational roles and costs of the considered propagation-modeling approaches within the proposed delta-learning workflow.

Approach	Main Computational Operation	Role in the Proposed Workflow	Relative Marginal Cost
Empirical baseline models	Curve interpolation, extrapolation, and closed-form correction terms	Fast baseline prediction and input to the delta-correction model	Low
Hybrid baseline models	One-dimensional terrain-profile processing combined with empirical and simplified physical submodels	Terrain-sensitive baseline prediction and input to the delta-correction model	Low to moderate
PEM reference model	Numerical solution on a two-dimensional range–height computational grid with boundary conditions	Offline generation of deterministic reference labels for supervised learning	High
Trained delta-correction model	Feature extraction followed by regression-model evaluation	Online correction of baseline predictions for new terrain profiles	Low

For this reason, the comparison in Table 4 is formulated at the workflow level. It separates the dominant operations required by each propagation approach from the additional cost introduced by the delta-learning correction. Empirical models are dominated by closed-form expressions, curve interpolation, extrapolation, and empirical correction terms. Hybrid ITU-R models additionally require one-dimensional terrain-profile processing and diffraction-related calculations. In contrast, PEM requires the numerical solution of a two-dimensional range–height propagation problem with boundary conditions, which leads to substantially higher computational and memory requirements. This difference in algorithmic structure is the main reason why deterministic models are computationally demanding for repeated large-scale network-planning and spectrum-management tasks.

The notation used in Table 4 is defined as follows. The variable n denotes the number of one-dimensional terrain-profile samples used by the empirical and hybrid propagation models. The symbols N and M denote the range and height grid dimensions of the two-dimensional PEM computational domain, respectively. These quantities are conceptually different: n refers to the sampled terrain profile along the propagation path, whereas N and M refer to the numerical field-computation grid used by PEM. The proposed delta-learning framework uses a compact feature vector of dimension $d = 18$, which is independent

of the full PEM range–height grid. For the gradient boosting regressor, T denotes the number of regression trees and L_{tree} denotes the characteristic number of leaves or terminal regions in a tree. For the MLP model, h denotes the number of arithmetic operations associated with the dense feed-forward layers. Finally, C_{base} denotes the computational cost of evaluating the corresponding baseline propagation model, including the required terrain-profile processing.

Table 4. Algorithmic computational complexity of the baseline, deterministic, and ML-corrected propagation workflows.

Model/Workflow	Dominant Operation	Approximate Inference Complexity
Free-space	Closed-form distance- and frequency-dependent formula	$\mathcal{O}(1)$
ITU-R P.1546	Curve interpolation and empirical correction terms	$\mathcal{O}(1)$–$\mathcal{O}(\log n)$
ITU-R P.452	Terrain-profile analysis, diffraction calculation, and statistical corrections	$\mathcal{O}(n)$
ITU-R P.1812	Terrain-profile analysis and hybrid diffraction-related corrections	$\mathcal{O}(n)$
PEM reference	Numerical range–height field solution with boundary conditions	$\mathcal{O}(N \log N)$ or $\mathcal{O}(N_x N_z)$
Baseline + ridge correction	Baseline evaluation, feature extraction, and linear correction	$\mathcal{O}(C_{\text{base}} + d)$
Baseline + kernel ridge correction	Baseline evaluation, feature extraction, and kernel prediction	$\mathcal{O}(C_{\text{base}} + n_{\text{tr}} d)$
Baseline + gradient boosting correction	Baseline evaluation, feature extraction, and regression-tree traversal	$\mathcal{O}(C_{\text{base}} + T \log L_{\text{tree}})$
Baseline + MLP correction	Baseline evaluation, feature extraction, and dense forward pass	$\mathcal{O}(C_{\text{base}} + h)$

As shown in Table 4, the ML-corrected workflows inherit the computational order of the underlying baseline propagation model, with only a low-dimensional regression-inference overhead added after feature extraction. This overhead is independent of the two-dimensional PEM grid and does not require solving a deterministic range–height field problem during inference. For linear models such as ridge regression, the additional online cost is essentially proportional to the feature-vector dimension. For tree-based and neural-network corrections, the additional cost is determined by the trained model size, but it remains an inference-only operation on compact feature vectors rather than a numerical electromagnetic field solution.

The computational advantage of the proposed framework should therefore be interpreted in an amortized sense. PEM is still required during the offline reference-generation stage, where it provides deterministic labels for supervised learning. However, once the correction model has been trained, repeated prediction for new terrain profiles no longer requires PEM simulation. Instead, the online workflow consists of baseline-model evaluation, construction of the 18-dimensional feature vector, and regression-model inference. Thus, the proposed method does not eliminate the computational cost of PEM itself, but transfers it to the offline training stage and preserves a substantially lower marginal inference complexity for repeated large-scale prediction tasks.

6. Delta Correction with Ridge Regression

This section presents the theoretical foundations of the delta correction approach based on ridge regression (L2 regularization). We introduce the basic principles of the method and describe the formulation of the proposed delta correction algorithm. Furthermore, we outline the evaluation methodology used to assess the performance of the model.

6.1. Theoretical Overview and Algorithm Description

The core assumption of the ridge regression model is that the delta field is not random noise, but is largely determined by the local geometric properties of the terrain and the response of the baseline model, and that it is linearly related to these factors. Accordingly, to each profile–point pair we assign the previously defined feature vector $\boldsymbol{\phi}_{i,j} \in \mathbb{R}^{18}$. The model formulation for each profile–point pair is as follows [71,72]:

$$\delta_i(d_j) = \mathbf{w}^\top \boldsymbol{\phi}_{i,j} + b + \varepsilon_i(d_j), \tag{47}$$

where $\mathbf{w} \in \mathbb{R}^{18}$ is the regression weight vector, b is a global bias term, and $\varepsilon_i(d_j)$ denotes the residual that cannot be modeled. This implies that the systematic error of the baseline model can be approximated in feature space by a globally linear operator of low complexity [73,74].

Reusing the composite indexing $n = (i, j)$, where $n = 1, \ldots, NL$, the feature (design) matrix and the target vector take the following form:

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\phi}_{1,1}^\top \\ \boldsymbol{\phi}_{1,2}^\top \\ \vdots \\ \boldsymbol{\phi}_{N,L}^\top \end{bmatrix} \in \mathbb{R}^{NL \times 18}, \qquad \boldsymbol{\delta} = \begin{bmatrix} \delta_1(d_1) \\ \delta_1(d_2) \\ \vdots \\ \delta_N(d_L) \end{bmatrix} \in \mathbb{R}^{NL}. \tag{48}$$

The regression model in compact form is:

$$\boldsymbol{\delta} = \boldsymbol{\Phi}\mathbf{w} + b\mathbf{1} + \boldsymbol{\varepsilon}, \tag{49}$$

where $\mathbf{1} \in \mathbb{R}^{NL}$ is a column vector of ones.

The estimation of the regression parameters could initially be formulated as a classical least-squares fitting problem, in which the sum of squared estimation errors is minimized. However, this approach is not sufficient for the present task. The terrain-based features exhibit strong mutual correlations, for example, between elevation, rolling mean and standard deviation, as well as between slope and curvature. As a consequence, the matrix $\boldsymbol{\Phi}^\top \boldsymbol{\Phi}$ may become poorly conditioned, leading to unstable regression coefficients with excessively large magnitudes and, ultimately, to poor generalization performance on unseen terrain profiles. An additional difficulty arises from the dense sampling along each terrain profile, which implies that the training data cannot be regarded as a collection of independent samples; without regularization, this would result in overfitting.

For these reasons, the learning problem is formulated as a regularized least-squares optimization task [75]:

$$(\hat{\mathbf{w}}, \hat{b}) = \arg\min_{\mathbf{w}, b} \left[\sum_{n=1}^{NL} \left(\delta_n - \mathbf{w}^\top \boldsymbol{\phi}_n - b \right)^2 + \lambda \|\mathbf{w}\|_2^2 \right]. \tag{50}$$

The symbols $\hat{\mathbf{w}}$ and $\hat{b}$ denote the estimates of the regression weights and the bias term obtained by solving the L2-regularized least-squares optimization problem on the training data. The second term, $\|\mathbf{w}\|_2^2 = \sum_k w_k^2$, is the squared Euclidean norm of the regression weight vector. For this reason, the method is commonly referred to as L2 regularization.

From a physical and statistical perspective, regularization constrains the model from excessively amplifying individual features; instead of reproducing sharp, profile-specific fluctuations, the baseline–PEM discrepancy is described as a smooth, moderate correction. The regularization parameter $\lambda > 0$ controls the trade-off between data fidelity and model simplicity.

The constant (bias) term is not regularized. This choice is motivated by the fact that the bias represents a global offset between the baseline model and the reference solution; penalizing this term would be unjustified and would leave a systematic residual error in the model. The selective application of regularization is implemented through an augmented matrix formulation. We introduce the augmented regression matrix

$$\boldsymbol{X}_a = \begin{bmatrix} \boldsymbol{\Phi} & \mathbf{1} \end{bmatrix}, \qquad \boldsymbol{\theta} = \begin{bmatrix} \mathbf{w} \\ b \end{bmatrix}, \tag{51}$$

together with the regularization matrix

$$\boldsymbol{R} \;=\; \mathrm{diag}(1, 1, \ldots, 1, 0), \tag{52}$$

where the zero entry corresponds to the bias term, indicating that it is excluded from the L2 penalty.

The resulting optimization problem is quadratic and admits a closed-form solution given by the solution of the following normal equation [76]:

$$\hat{\boldsymbol{\theta}} = \left(\boldsymbol{X}_a^\top \boldsymbol{X}_a + \lambda \boldsymbol{R} \right)^{-1} \boldsymbol{X}_a^\top \boldsymbol{\delta}. \tag{53}$$

This formulation directly illustrates the effect of regularization: the term $\lambda \boldsymbol{R}$ stabilizes the normal equation along the diagonal directions, improves numerical conditioning, and guarantees a unique and stable solution. The resulting parameters define a deterministic correction operator that improves the baseline model output in an additive manner while preserving its underlying physical structure.

The regularization parameter λ is a hyperparameter that controls the trade-off between data fidelity and model smoothness. Its value cannot be derived analytically from the regression problem itself and must therefore be selected using an external criterion. In this work, λ is determined by profile-level K-fold cross-validation [77]. In the proposed cross-validation scheme, the basic unit of data splitting is a complete terrain profile rather than an individual distance sample. The set of N profiles is partitioned into K disjoint folds, each fold consisting of a subset of profile indices. During cross-validation, all distance samples belonging to a given profile are assigned exclusively either to the training set or to the validation set. Consequently, one fold typically contains multiple profiles, and the procedure evaluates generalization performance at the profile level rather than at the point level. For each candidate value $\lambda \in \Lambda$, the model is trained on $K - 1$ folds and evaluated on the remaining fold, and the procedure is repeated for all folds [78]. The average validation error is then computed as

$$\mathrm{MAE}_{\mathrm{CV}}(\lambda) = \frac{1}{K} \sum_{k=1}^{K} \mathrm{MAE}^{(k)}(\lambda). \tag{54}$$

The optimal regularization parameter is selected according to

$$\lambda^* = \arg\min_{\lambda \in \Lambda} \mathrm{MAE}_{\mathrm{CV}}(\lambda), \tag{55}$$

i.e., the value that minimizes the cross-validated mean absolute error across terrain profiles. This selection strategy ensures that the resulting model achieves the best generalization performance on previously unseen terrain profiles.

Prior to regression, feature standardization is essential, as the Ridge penalty is not scale invariant. A critical aspect of the proposed implementation is, however, how this standardization is performed. Instead of computing feature statistics over the entire dataset, the mean and standard deviation are determined independently for each cross-validation fold, using exclusively the samples belonging to the training profiles of the given fold.

Formally, in the k-th cross-validation fold, let $\mathcal{T}^{(k)} \subset \{1, \ldots, NL\}$ denote the index set of training samples. The mean and standard deviation of the m-th feature are computed as

$$\mu_m^{(k)} = \frac{1}{|\mathcal{T}^{(k)}|} \sum_{n \in \mathcal{T}^{(k)}} \phi_{n,m}, \qquad \sigma_m^{(k)} = \sqrt{\frac{1}{|\mathcal{T}^{(k)}| - 1} \sum_{n \in \mathcal{T}^{(k)}} \left(\phi_{n,m} - \mu_m^{(k)} \right)^2}. \tag{56}$$

These statistics are then used to standardize *both* the training and validation samples according to

$$\phi_{n,m}^{\text{std}} = \frac{\phi_{n,m} - \mu_m^{(k)}}{\sigma_m^{(k)}}. \tag{57}$$

This step is crucial. If standardization were performed using statistics computed over the entire dataset, information from the validation profiles would implicitly leak into the training process, leading to overly optimistic validation results. The proposed implementation correctly avoids such data leakage by performing feature standardization independently within each fold.

Once the optimal regularization parameter λ^* has been selected, the final model is trained using all available profiles. In this case, the feature statistics are computed over the full dataset (μ_m^{all} and σ_m^{all}), and stored as part of the trained model. During inference on new terrain profiles, these global statistics are applied to ensure consistency between training and deployment.

The output of the trained model is not the delta field itself, but rather the reconstructed reference field, obtained by additively correcting the baseline model:

$$\hat{y}_i(d_j) = b_i(d_j) + \hat{\delta}_i(d_j), \tag{58}$$

where the estimated delta term is given by

$$\hat{\delta}_i(d_j) = \hat{\mathbf{w}}^\top \left(\frac{\boldsymbol{\phi}_{i,j} - \boldsymbol{\mu}}{\boldsymbol{\sigma}} \right) + \hat{b}. \tag{59}$$

This formulation ensures that the physical structure of the baseline propagation model is preserved, while its systematic bias is corrected in a data-driven manner.

6.2. Model Performance Evaluation and Metrics

The performance of the proposed delta learning–based correction schemes was evaluated in a manner consistent with the physical characteristics of radiowave propagation and the profile-level data separation applied throughout the study. Model accuracy was assessed exclusively at the terrain-profile level, and all results are reported in terms of electric field strength in dBμV/m.

Deterministic Parabolic Equation Modeling (PEM) outputs served as reference field-strength profiles, while empirical and hybrid propagation models, as well as their delta-corrected variants, provided the corresponding estimates.

Let i denote a test terrain profile with $L = 601$ distance samples. The pointwise prediction error is defined as

$$e_{i,j} = E_i^{\text{ref}}(d_j) - E_i^{\text{est}}(d_j). \tag{60}$$

All error metrics were first computed for each profile and subsequently aggregated. The mean absolute error (MAE) [79] is given by

$$\text{MAE}_i = \frac{1}{L}\sum_{j=1}^{L}|e_{i,j}|, \tag{61}$$

providing a robust measure of average accuracy. The root mean square error (RMSE) [80],

$$\text{RMSE}_i = \sqrt{\frac{1}{L}\sum_{j=1}^{L}e_{i,j}^2}, \tag{62}$$

emphasizes larger local deviations, while the bias [81],

$$\text{Bias}_i = \frac{1}{L}\sum_{j=1}^{L}e_{i,j}, \tag{63}$$

quantifies systematic over- or underestimation. In addition, a robust percentile-based metric,

$$\text{P90}_i = \text{percentile}_{90}(|e_{i,1}|, \ldots, |e_{i,L}|), \tag{64}$$

was used to characterize the upper range of typical errors.

All metrics were evaluated separately for flat, urban–hilly, and mountainous terrain types. For each terrain category, the final performance indicators were obtained by averaging the corresponding profile-level metrics over the fixed test set:

$$\overline{\text{MAE}} = \frac{1}{N_{\text{test}}}\sum_{i=1}^{N_{\text{test}}}\text{MAE}_i, \tag{65}$$

with analogous expressions for RMSE, Bias, and P90.

To analyze learning behavior, model performance was further evaluated as a function of the number of training profiles. For each training set size, repeated random selections of training profiles were used, while the test set remained fixed. The resulting mean values and dispersions of the profile-level metrics form the basis of the learning curves presented in the Section 6.3.

6.3. Results

This subsection presents the empirical evaluation of the proposed delta learning framework based on ridge regression. The results quantify the extent to which the learned correction reduces the systematic deviation of the baseline propagation models with respect to the deterministic PEM reference.

The analysis is performed separately for the three terrain categories (flat, urban–hilly, and mountainous environments). Model performance is evaluated through learning curves that describe the evolution of the error metrics as a function of the number of training terrain profiles. The curves for Budapest region are shown in Figure 9.

To analyze the data efficiency and generalization behavior of the proposed delta learning framework, learning curves were constructed by evaluating the model performance as a function of the number of training terrain profiles. For each terrain scenario, a dataset

consisting of 5000 terrain profiles was available for training and model selection, while an independent test set of 1000 profiles was kept fixed throughout the experiments. Each terrain profile contained $L = 601$ distance samples corresponding to a propagation path of 30 km with a spatial resolution of 50 m.

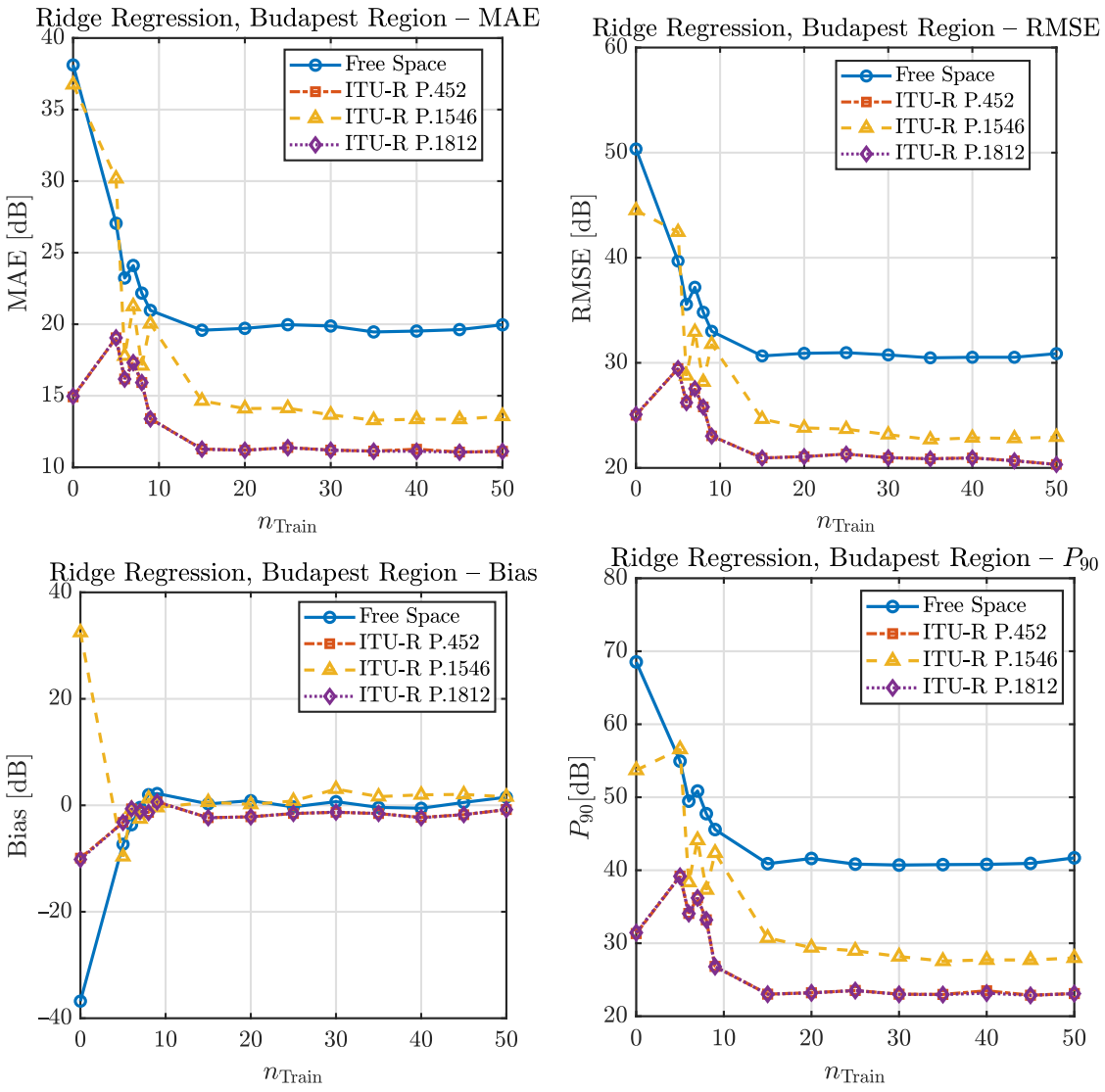


Figure 9. Error metrics as a function of the number of training profiles, Budapest region.

The learning curves were generated by progressively increasing the number of training profiles n_{Train}. For a given training size, a subset of terrain profiles was randomly selected from the training pool. To reduce the effect of random sampling variability, the selection was repeated multiple times, and the resulting error metrics were averaged over the repetitions. Within each training subset, the regularization parameter of the ridge regression model was determined using profile-level K-fold cross-validation. In this procedure, complete terrain profiles were assigned to folds, ensuring that all distance samples belonging to a given profile were kept within the same fold. This prevents information leakage caused by the strong spatial correlation between neighboring samples along a terrain profile. For each candidate value of the regularization parameter, the model was trained on $K - 1$ folds and validated on the remaining fold, and the average cross-validation error was used to select the optimal parameter.

After determining the optimal regularization parameter, the ridge regression model was retrained on the entire training subset of size n_{Train}. The resulting model was then

evaluated on the fixed test set of 1000 terrain profiles. Performance metrics were computed at the profile level according to Equations (61)–(64), and the reported values correspond to averages over the full test set. This procedure was repeated for all considered baseline propagation models (free-space, ITU-R P.1546, ITU-R P.1812, and ITU-R P.452), resulting in the learning curves shown in Figure 9 for the urban–hilly Budapest scenario. The same evaluation methodology was applied to the other terrain categories.

A clear saturation behavior can be observed: the error metrics reach their minimum values already at relatively small training set sizes of approximately $n_{\mathrm{Train}} \sim 30$–$40$ terrain profiles. Beyond this point, increasing the number of training samples results only in marginal changes in the error metrics. This indicates that the proposed feature representation captures the dominant terrain-dependent correction patterns efficiently, and that the regression model can be trained successfully even with a limited number of reference profiles.

The magnitude of the improvement provided by the learned delta correction depends on the baseline propagation model. The largest gain is observed for the simplest models. In the case of the free-space model, the MAE decreases from approximately 38 dB to about 19 dB after correction, corresponding to roughly a twofold reduction in the average error. A similarly pronounced improvement can be observed for the empirical ITU-R P.1546 model, where the error decreases substantially once the correction model is trained.

For the more accurate hybrid propagation models (ITU-R P.452 and ITU-R P.1812), the achievable improvement is naturally smaller, since these models already incorporate terrain-dependent physical mechanisms. Nevertheless, the results confirm that the learned correction further reduces the residual error. It is also notable that the curves of the ITU-R P.452 and ITU-R P.1812 models are almost identical across all error metrics, suggesting that the two models exhibit very similar residual structures relative to the PEM reference.

The bias curves provide additional insight into the behavior of the correction model. For all baseline models, the initial systematic over- or underestimation relative to the PEM reference gradually disappears as the number of training profiles increases. The bias converges toward zero, indicating that the learned correction successfully compensates for the systematic component of the baseline error. As a result, the remaining deviation between the corrected model and the reference becomes approximately noise-like, which is consistent with the expected behavior of the delta-learning framework.

Figure 10 presents representative examples of corrected propagation profiles for the Budapest scenario.

In each subplot, the baseline model prediction, the PEM reference, and the ridge-regression-based estimates are shown together with the terrain profile. Two corrected curves are illustrated: one obtained using a very small training set ($n_{\mathrm{Train}} = 25$) and another using a much larger training set ($n_{\mathrm{Train}} = 1000$). The figure confirms the trends observed in the learning curves. Even when the regression model is trained using only a small number of terrain profiles, the corrected estimate already follows the PEM reference considerably better than the original baseline prediction. Increasing the training set size further improves the agreement slightly, but the difference between the two corrected curves remains relatively small, which is consistent with the early saturation of the learning curves.

The effect of the correction is particularly visible for the simpler baseline models, where the original predictions deviate strongly from the PEM reference. In contrast, for the hybrid models (ITU-R P.452 and ITU-R P.1812), the baseline predictions are already closer to the reference, and the learned correction mainly refines the remaining local deviations along the propagation path.

The results obtained for the Kecskemét region (Figures 11 and 12) show a behavior that is largely consistent with the trends observed in the Budapest scenario.

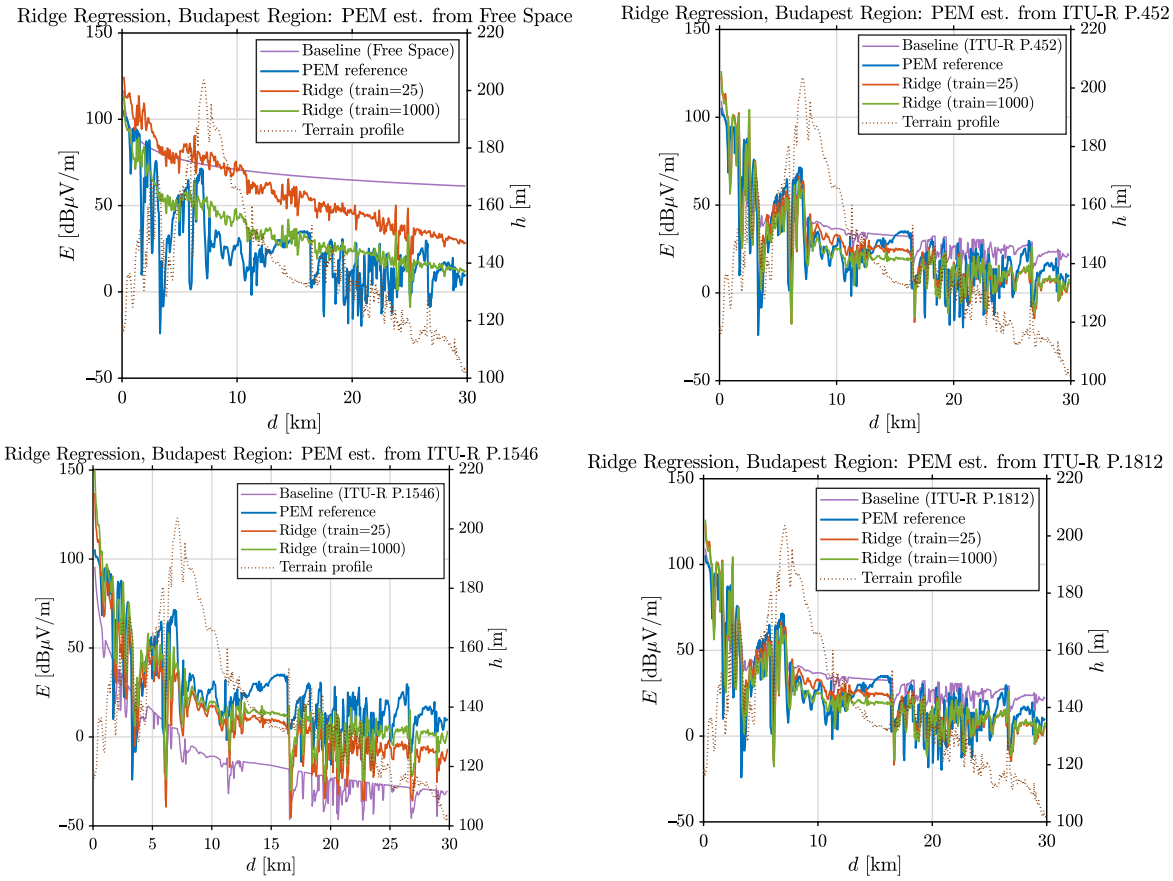


Figure 10. Examples for calculated profiles with ridge regression, Budapest region.

As shown in Figure 11, the learning curves again exhibit a rapid saturation with respect to the number of training profiles. Similar to the Budapest case, the error metrics reach their minimum values already at approximately n_{Train}∼30 profiles, indicating that only a limited number of reference samples is required to train the correction model effectively. However, the absolute error levels are generally lower than in the Budapest region, which can be attributed to the less complex terrain morphology and the consequently simpler propagation conditions.

The largest relative improvement is again seen for the simplest baselines. The empirical ITU-R P.1546 model shows a substantial reduction in all error metrics after correction, while the free-space model also improves, though its remaining errors stay higher due to its inherent limitations.

For the hybrid models (ITU-R P.452 and ITU-R P.1812), the improvement is smaller but still clear. As in the Budapest case, their curves nearly overlap across all metrics, indicating very similar residual error structures relative to the PEM reference.

The bias curves show that systematic errors decrease with increasing training set size and approach zero for all models, confirming that the learned delta correction removes systematic error. Figure 12 also shows that corrected profiles follow the PEM reference much better than the original baselines, even with limited training data.

The Mátra learning curves, representing a mountainous region, are shown in Figure 13.

Compared to the previous two terrain types, a slower saturation effect can be observed (n_{Train}∼70–80). At the same time, the initial error value at $n_{\text{Train}} = 0$, which represents the error between the baseline model result and the PEM reference result, is the most significant in the case of the free space model.

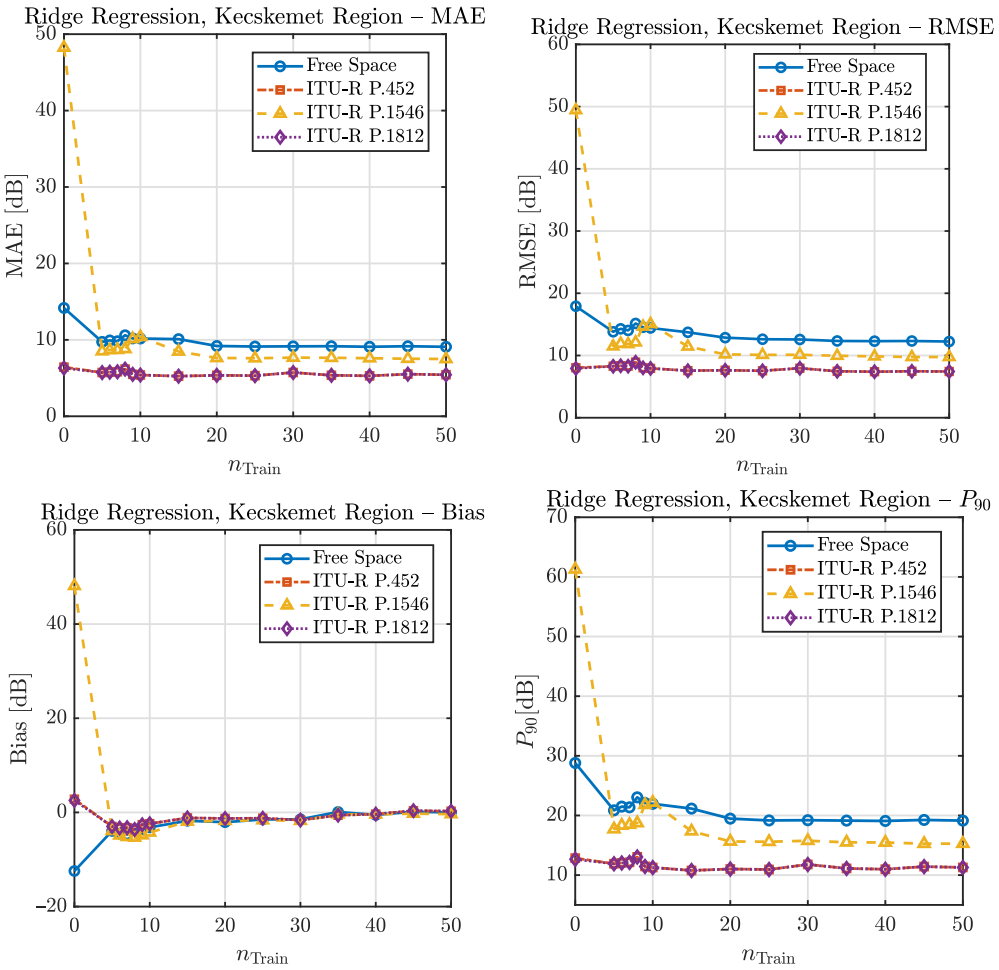


Figure 11. Error metrics as a function of the number of training profiles, Kecskemét region.

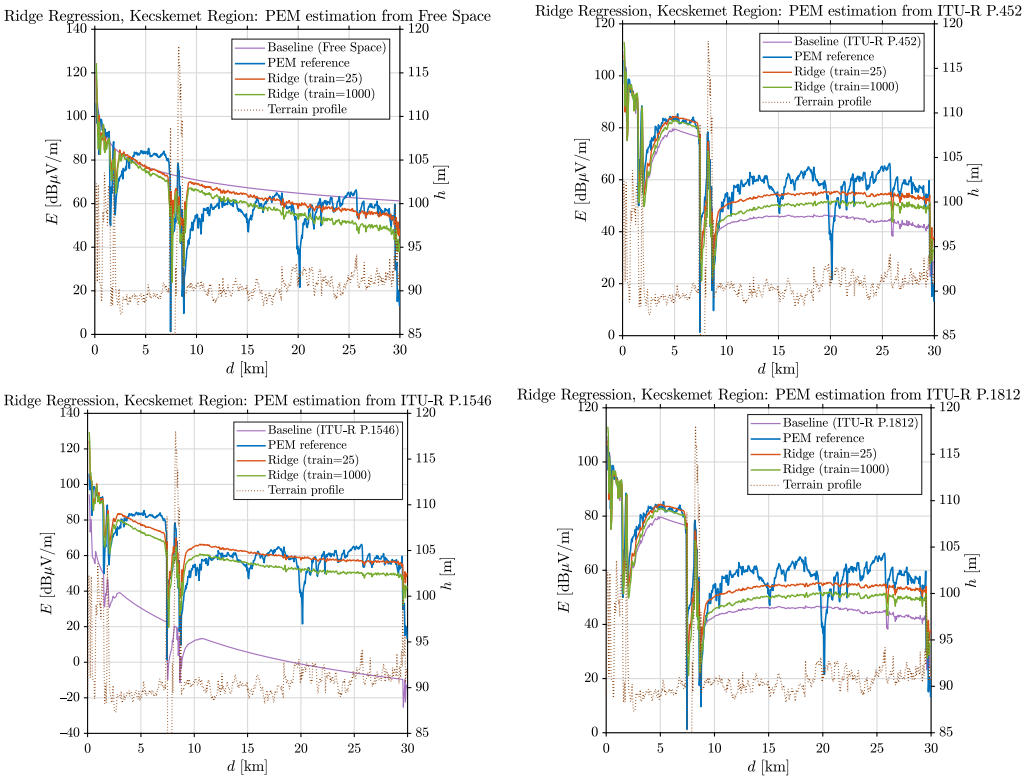


Figure 12. Examples for calculated profiles with ridge regression, Kecskemét region.

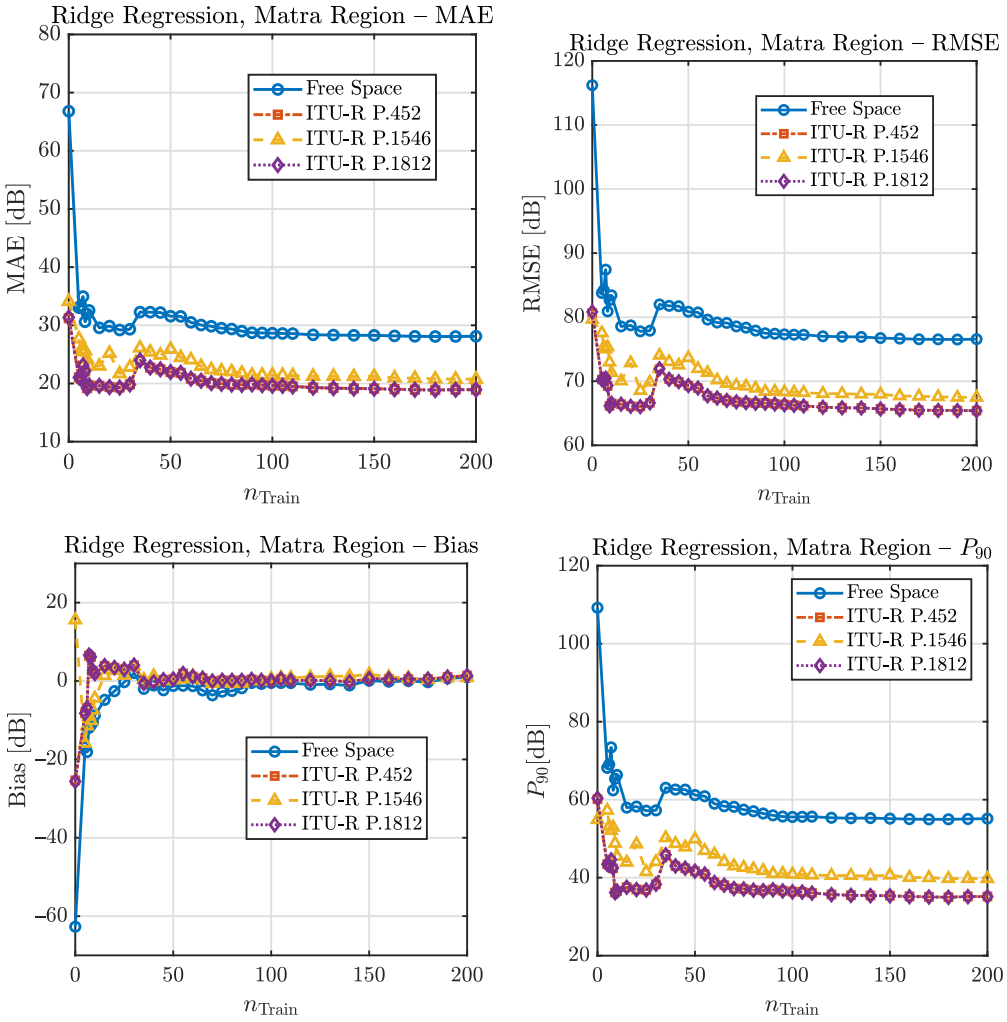


Figure 13. Error metrics as a function of the number of training profiles, Mátra region.

Here, the model provides a substantial improvement of 35 dB in the MAE value, while for the ITU models the improvement is smaller but still significant. The consistent under- and overestimation of the models is again smoothed by the ridge regression delta estimation, making it noise-like and close to 0 dB.

Overall, although a larger training sample is required for this terrain type, with 100 samples the model already reaches the limit of its performance, providing a significant improvement compared to the baseline results.

Figure 14 shows examples of the estimation of field strength level curves using 25 and 1000 training samples.

It can be clearly observed how challenging it is for our linear ridge regression model to operate in an environment fragmented by mountains and hills for all baseline models, with particular difficulty in the case of the free space model. The curves clearly demonstrate both this difficulty and the nevertheless considerable improvement in error.

Although the model faces the greatest challenge here among the three terrain types, a significant improvement can still be achieved even in this context. Despite the previously discussed limitations, this makes its practical applicability possible.

Overall, the ridge-regression-based delta correction provides a consistent and substantial improvement across all investigated terrain categories and baseline propagation models. In the urban–hilly Budapest scenario, the largest gain is observed for the free-space baseline, where the MAE decreases from approximately 38 dB to about 19 dB, corresponding to roughly a 50% reduction in the average error. The empirical ITU-R P.1546 model

also exhibits a strong improvement, with typical error reductions on the order of 30–40%, while the hybrid ITU-R P.1812 and ITU-R P.452 models show smaller but still consistent improvements of approximately 10–20%, reflecting their already higher physical fidelity. In the flat Kecskemét region, where baseline models already provide relatively accurate predictions, the achievable improvement is naturally more moderate but still significant, typically ranging between 20–40% depending on the baseline model. In the mountainous Mátra scenario, which represents the most challenging propagation environment, the learned correction remains highly effective: for the free-space baseline the MAE improvement reaches roughly 35 dB, corresponding to an error reduction of about 45–55%, while the ITU-based models still benefit from improvements in the range of approximately 15–30%. Across all terrain types, the systematic bias of the baseline models is effectively eliminated, with the residual error converging toward a noise-like distribution around zero.

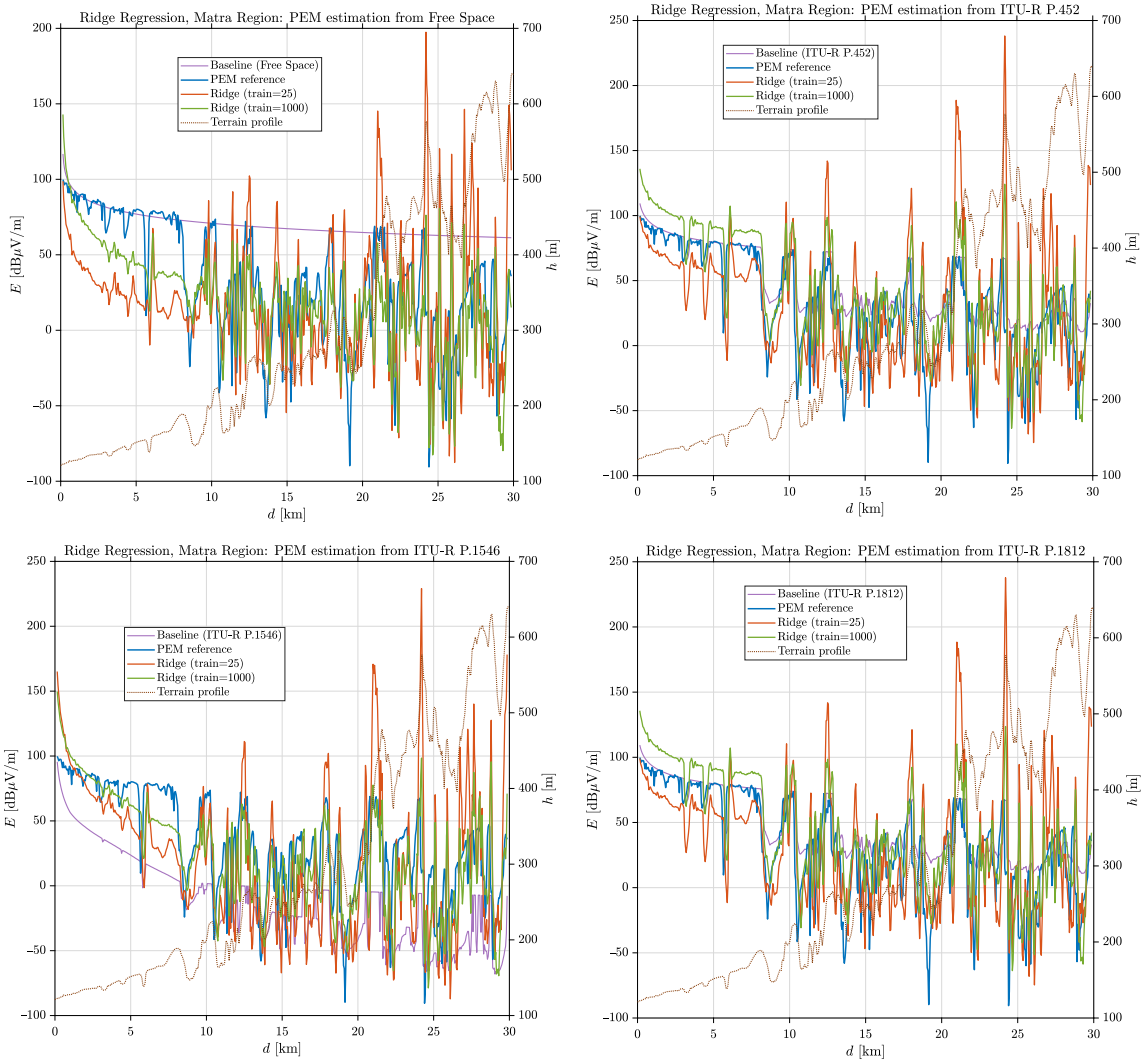


Figure 14. Examples for calculated profiles with ridge regression, Mátra region.

These results demonstrate that the proposed ridge-regression delta-learning framework can reliably capture the systematic discrepancy between fast empirical or hybrid propagation models and the deterministic PEM reference, enabling substantial accuracy improvements while preserving the computational efficiency of the baseline approaches.

7. Delta Correction with Kernel Ridge Regression

This section extends the previously introduced ridge regression framework by incorporating nonlinear feature transformations through kernel methods. While the linear ridge regression model assumes that the residual correction term can be approximated as a linear function of the feature vector, this assumption may be restrictive when the relationship between terrain descriptors and propagation-model errors is inherently nonlinear.

To address this limitation, we introduce kernel ridge regression, which enables nonlinear regression while preserving the regularized least-squares formulation of ridge regression. The section first presents the theoretical foundations of the kernel-based formulation, including the representer theorem and the resulting dual problem. Subsequently, we describe the use of the radial basis function (RBF) kernel and the landmark-based approximation employed to make the method computationally tractable for large terrain-profile datasets.

7.1. Theoretical Overview and Algorithm Description

While ridge regression assumes a globally linear relationship between the feature representation and the residual field, the systematic error of propagation models may exhibit nonlinear dependencies on terrain geometry and baseline responses. In particular, diffraction effects, terrain shadowing, and the interaction of local and global terrain descriptors may introduce nonlinear structures that cannot be captured by a linear model in the original feature space. For this reason, we can assume that the linear ridge regression framework introduced in the previous subsection can be extended to a nonlinear regression model through the application of kernel methods.

Let

$$\psi : \mathbb{R}^{18} \to \mathcal{H} \tag{66}$$

denote a nonlinear mapping that transforms the original feature vector $\boldsymbol{\phi_{i,j}}$ into a (possibly high-dimensional) Hilbert space $\mathcal{H}$. In this transformed space, the regression model takes the form

$$\delta_i(d_j) = \mathbf{w}^\top \psi(\boldsymbol{\phi}_{i,j}) + b + \varepsilon_i(d_j), \tag{67}$$

where $\mathbf{w} \in \mathcal{H}$ denotes the regression coefficient vector in the transformed feature space, b is a bias term, and $\varepsilon_i(d_j)$ represents the residual modeling error [82].

Similarly to ridge regression, the parameters of the model are obtained by minimizing a regularized least-squares objective:

$$(\hat{\mathbf{w}}, \hat{b}) = \arg\min_{\mathbf{w},b} \left[\sum_{i=1}^{N} \sum_{j=1}^{L} \left(\delta_i(d_j) - \mathbf{w}^\top \psi(\boldsymbol{\phi}_{i,j}) - b \right)^2 + \lambda \|\mathbf{w}\|^2 \right], \tag{68}$$

where $\lambda > 0$ is the regularization parameter controlling the trade-off between data fidelity and model complexity [83].

Direct computation of the nonlinear mapping $\psi(\cdot)$ may be computationally impractical if the dimensionality of the feature space $\mathcal{H}$ is large. Kernel methods overcome this difficulty through the so-called kernel trick. Instead of explicitly evaluating $\psi(\cdot)$, the regression model is expressed entirely through inner products in the transformed feature space. For two profile-distance samples (i, j) and (k, l), the similarity between their feature vectors is measured using a kernel function

$$K(\boldsymbol{\phi}_{i,j}, \boldsymbol{\phi}_{k,l}) = \psi(\boldsymbol{\phi}_{i,j})^\top \psi(\boldsymbol{\phi}_{k,l}). \tag{69}$$

According to the representer theorem, the optimal solution of the regularized regression problem defined in Equation (68) can be expressed as a linear combination of the mapped training samples. Consequently, the regression function can be written as

$$\hat{g}(\boldsymbol{\phi}) = \sum_{k=1}^{N} \sum_{l=1}^{L} \alpha_{k,l} K(\boldsymbol{\phi}_{k,l}, \boldsymbol{\phi}) + b, \tag{70}$$

where $\alpha_{k,l}$ are scalar coefficients associated with the training samples indexed by (k, l).

Let the kernel matrix be defined as

$$\mathbf{K}_{(i,j),(k,l)} = K(\boldsymbol{\phi}_{i,j}, \boldsymbol{\phi}_{k,l}), \tag{71}$$

which contains the pairwise kernel evaluations between all training samples. Using the composite index $n = (i, j)$ introduced earlier, the kernel matrix $\mathbf{K} \in \mathbb{R}^{NL \times NL}$ can be constructed from these elements [84].

The coefficients can then be obtained by solving the linear system

$$(\mathbf{K} + \lambda \mathbf{I})\boldsymbol{\alpha} = \boldsymbol{\delta}, \tag{72}$$

where $\boldsymbol{\alpha} \in \mathbb{R}^{NL}$ is the vector of kernel coefficients and $\boldsymbol{\delta} \in \mathbb{R}^{NL}$ is the target vector defined in Equation (49).

For a new terrain-profile point characterized by the feature vector $\boldsymbol{\phi}_{i,j}$, the predicted residual value becomes

$$\hat{\delta}_i(d_j) = \sum_{k=1}^{N} \sum_{l=1}^{L} \alpha_{k,l} K(\boldsymbol{\phi}_{k,l}, \boldsymbol{\phi}_{i,j}). \tag{73}$$

Within the delta learning framework introduced earlier, the reconstructed reference field is obtained by additively correcting the baseline model output

$$\hat{y}_i(d_j) = b_i(d_j) + \hat{\delta}_i(d_j), \tag{74}$$

which preserves the additive correction structure defined in Equation (8).

In this study, nonlinear similarity between feature vectors is modeled using the radial basis function (RBF) kernel

$$K(\boldsymbol{\phi}_{i,j}, \boldsymbol{\phi}_{k,l}) = \exp\left(-\frac{\|\boldsymbol{\phi}_{i,j} - \boldsymbol{\phi}_{k,l}\|^2}{2\sigma^2}\right), \tag{75}$$

where σ is the kernel bandwidth parameter controlling the smoothness of the regression function [85].

Because each terrain profile contains L samples and the total number of training samples equals NL, the kernel matrix $\mathbf{K} \in \mathbb{R}^{NL \times NL}$ introduced in Equation (72) may become very large. Consequently, solving the full kernel ridge regression system can be computationally expensive both in terms of memory usage and computational complexity.

To address this limitation, we employ a landmark-based low-rank approximation of the kernel matrix. Instead of using all training samples as kernel basis functions, a smaller subset of representative feature vectors is selected from the training data. These vectors are referred to as landmark points.

Let

$$\mathbf{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\} \tag{76}$$

denote a set of m landmark feature vectors, where each $\mathbf{u}_k$ is selected from the training feature vectors $\boldsymbol{\phi}_{i,j}$ and $m \ll NL$. These landmark vectors serve as representative samples that approximate the kernel interactions of the full training set.

Using these landmark points, kernel evaluations are computed only between the training samples and the landmark vectors. This results in the rectangular kernel matrix

$$\mathbf{K}_{nm} = \left[K(\boldsymbol{\phi}_{i,j}, \mathbf{u}_k)\right], \tag{77}$$

which contains the similarities between all training samples and the landmark points. In addition, the pairwise kernel evaluations between the landmark vectors form the matrix

$$\mathbf{K}_{mm} = [K(\mathbf{u}_k, \mathbf{u}_l)]_{k,l=1}^{m}, \tag{78}$$

where $\mathbf{K}_{mm} \in \mathbb{R}^{m \times m}$ contains the kernel similarities among the landmark feature vectors.

Using this low-rank approximation, the kernel ridge regression problem can be reformulated as

$$(\mathbf{K}_{nm}^{\top}\mathbf{K}_{nm} + \lambda \mathbf{K}_{mm})\mathbf{a} = \mathbf{K}_{nm}^{\top}\boldsymbol{\delta}, \tag{79}$$

where $\mathbf{a} \in \mathbb{R}^m$ is the vector of regression coefficients associated with the landmark points. This approximation is closely related to Nyström-type kernel approximations commonly used for scalable kernel methods.

Once the coefficients $\mathbf{a}$ are obtained, the predicted residual for a new feature vector $\boldsymbol{\phi}_{i,j}$ can be evaluated as

$$\hat{\delta}_i(d_j) = \sum_{k=1}^{m} a_k K(\boldsymbol{\phi}_{i,j}, \mathbf{u}_k). \tag{80}$$

This formulation replaces the full kernel expansion over all NL training samples with a reduced expansion over only m landmark points. As a result, the computational complexity of kernel ridge regression is significantly reduced, while the model retains the ability to capture nonlinear relationships between terrain features and the systematic errors of the baseline propagation model [86].

7.2. Results

The performance of the proposed delta-learning framework was further evaluated using kernel ridge regression (KRR). The evaluation protocol follows the same profile-level methodology described previously in order to ensure direct comparability with the ridge regression results.

For each terrain scenario, the available terrain profiles were partitioned into a training pool and a fixed test set using a profile-level split. Approximately 20% of the profiles were reserved for testing, while the remaining profiles were used for model selection and training. Each terrain profile contains $L = 601$ distance samples corresponding to a 30 km propagation path with a spatial resolution of 50 m.

Similarly to the ridge regression experiments, learning curves were constructed by progressively increasing the number of training terrain profiles while keeping the test set fixed. For each training set size, the model was trained on the selected profiles and evaluated on the same independent test set.

The learning curves obtained for the Budapest region are shown in Figure 15.

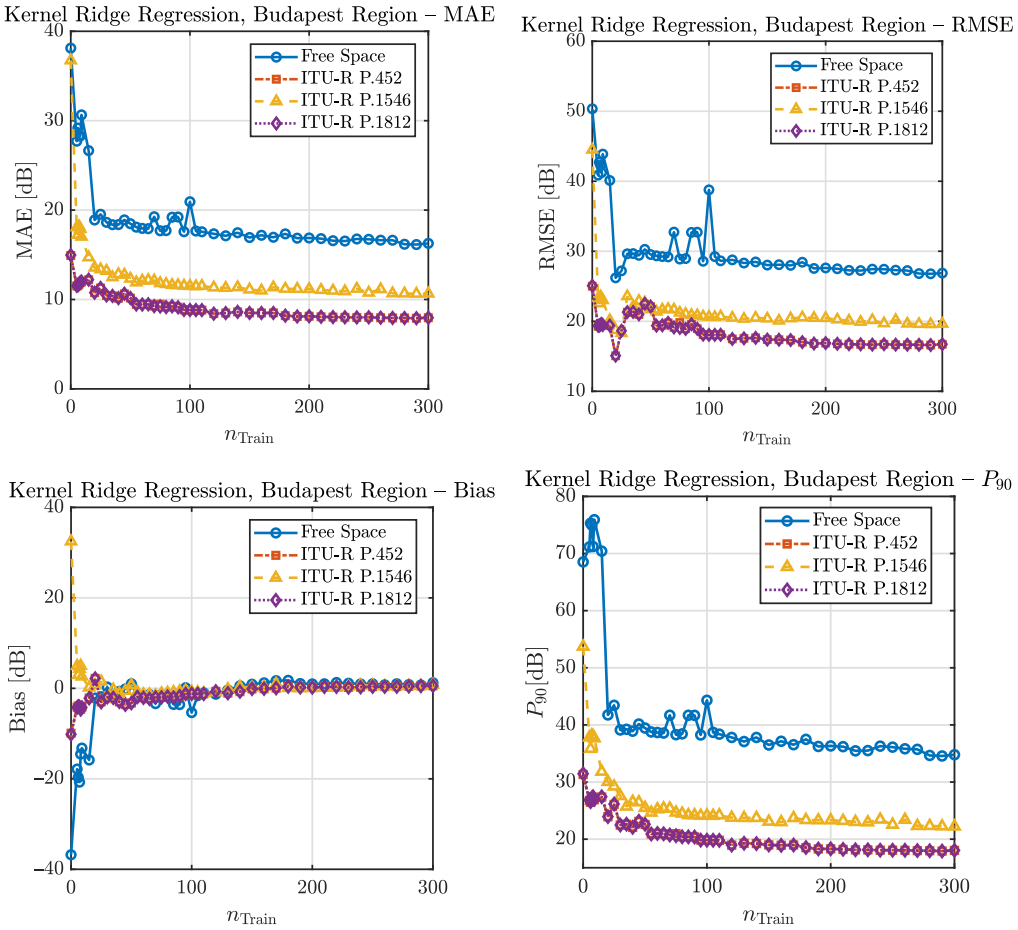


Figure 15. Examples for calculated profiles with kernel ridge regression, Budapest region.

These curves illustrate the evolution of the prediction error as a function of the number of training terrain profiles used to train the kernel ridge regression model.

A behavior similar to that observed for the linear ridge regression model can be identified. The error metrics decrease rapidly when the first few training profiles are introduced, indicating that the model quickly learns the dominant terrain-dependent correction patterns. The most pronounced error reduction occurs within the range of approximately 10–30 training profiles. Beyond this range, the curves gradually approach a stable level and further increases in the training set size lead only to marginal improvements.

This saturation behavior suggests that the feature representation introduced in Section 4 captures the most important terrain-dependent structures of the baseline model residual with relatively small training datasets.

Figure 16 presents representative examples of corrected propagation profiles obtained with ridge regression and kernel ridge regression for the Budapest region.

The plots illustrate the baseline model predictions, the PEM reference solution, and the corrected estimates for both a small training set ($n_{\text{Train}} = 25$) and a large training set ($n_{\text{Train}} = 1000$).

In all cases, both learning approaches significantly reduce the deviation of the baseline propagation models from the deterministic PEM reference. The improvement is particularly pronounced for the simpler baseline models, such as free-space propagation and the empirical ITU-R P.1546 model, where the original baseline curves differ substantially from the PEM reference. After correction, both regression methods shift the predicted profiles toward the reference solution and capture the main large-scale propagation trends along the path.

When only a small number of training profiles is available, the kernel ridge regression model tends to follow local fluctuations of the reference curve more closely, whereas the ridge regression solution produces smoother estimates that primarily capture the dominant propagation trend. As the number of training profiles increases, the corrected curves produced by both approaches converge and the overall differences between the two models become small.

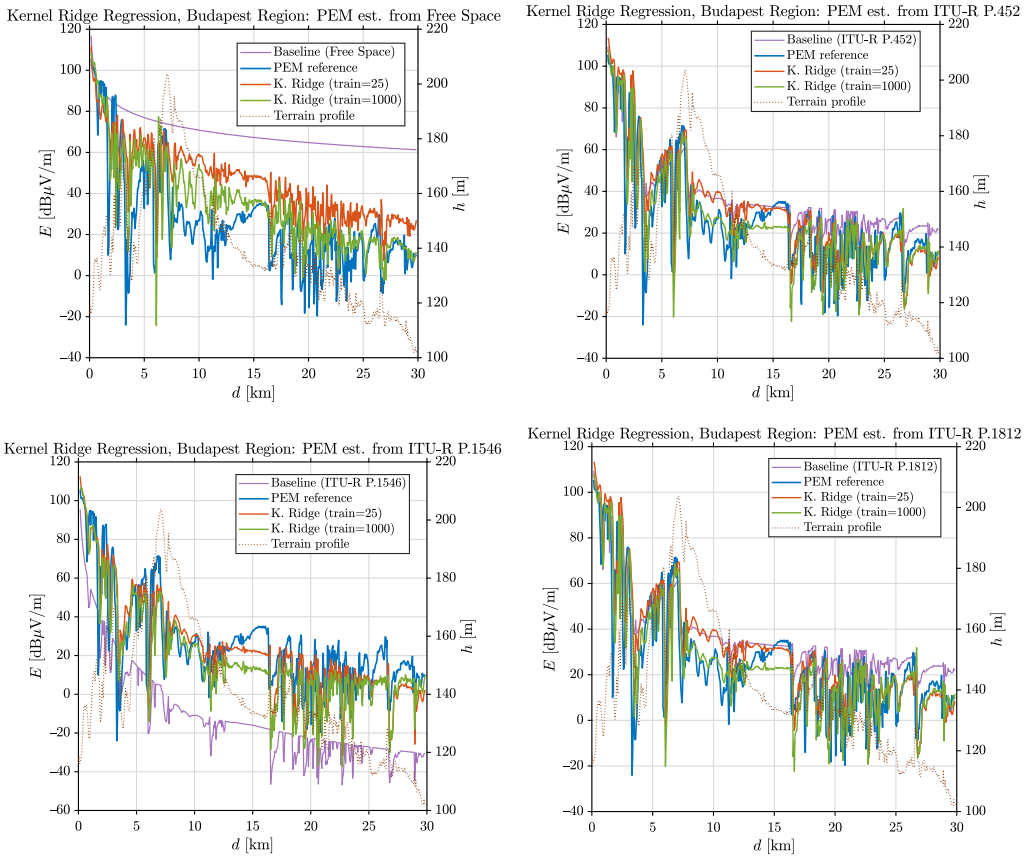


Figure 16. Examples for calculated profiles with kernel ridge regression, Budapest region.

For the hybrid baseline models (ITU-R P.452 and ITU-R P.1812), the baseline predictions are already closer to the PEM reference, and therefore the achievable correction is more moderate. In these cases, both regression models produce very similar corrected profiles, indicating that the residual structure of the hybrid models relative to the PEM solution can be captured effectively even by the linear ridge formulation.

Overall, the profile-level examples for the Budapest region indicate that both ridge regression and kernel ridge regression are capable of substantially reducing the systematic deviation of the baseline propagation models from the PEM reference solution. In particular, the corrected profiles follow the dominant large-scale propagation trends along the path, while the remaining discrepancies are mainly associated with small-scale fluctuations of the reference field strength. The visual comparison further suggests that the two regression approaches produce very similar corrected profiles once sufficient training data are available. However, the learning curves indicate that the kernel ridge regression model typically requires a somewhat larger number of training profiles before the error metrics stabilize. This behavior is consistent with the higher model flexibility of the kernel formulation, which allows the regression to capture local variations of the reference curve but also increases the data requirement for stable generalization.

The evolution of error metrics in the case of Kecskemét Region is illustrated in Figure 17.

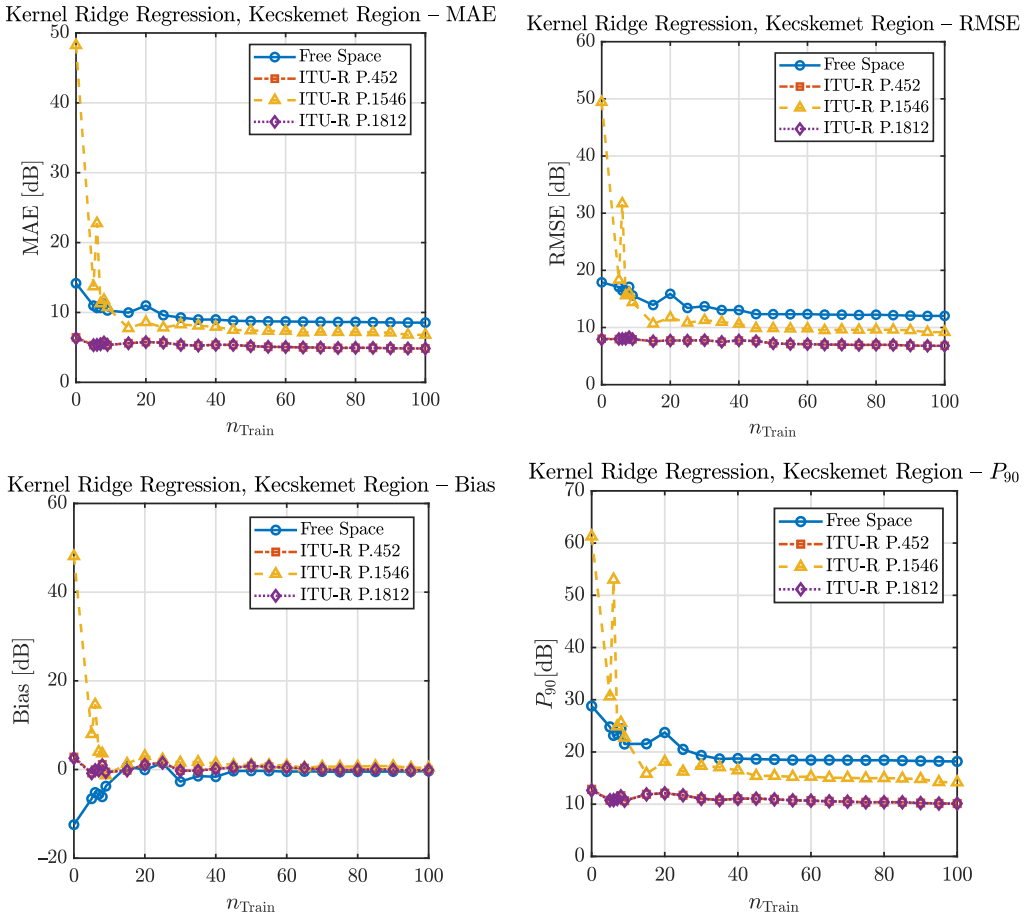


Figure 17. Error metrics as a function of the number of training profiles, Kecskemét region.

The learning curves clearly indicate a rapid convergence of the error metrics as the number of training profiles increases. For all baseline propagation models, the largest improvement occurs within approximately the first 10–20 training profiles, after which the curves gradually approach a stable error level. This behaviour is consistent across all error metrics (MAE, RMSE, Bias and P_{90}) and reflects the relatively simple propagation conditions of the investigated flat-terrain environment. Compared to the more complex urban and hilly terrain of the Budapest region, the propagation behaviour over the Kecskemét region is less affected by terrain-induced diffraction and shadowing effects, which leads to smoother and more predictable residual structures that can be learned with relatively small training datasets.

The magnitude of the achievable improvement depends strongly on the baseline propagation model. The largest reduction in error is observed for the empirical ITU-R P.1546 model, which exhibits a large initial deviation from the PEM reference. In contrast, the hybrid models ITU-R P.452 and ITU-R P.1812 already provide relatively accurate baseline predictions for this terrain type, and therefore only moderate improvements can be achieved through the learning-based correction.

A comparison with the ridge regression results presented earlier also indicates that the kernel ridge formulation does not provide a measurable performance advantage in this scenario. The final error levels obtained with kernel ridge regression are nearly identical to those achieved with the linear ridge regression model. This observation suggests that the residual error between the baseline models and the PEM reference can already be well represented within the chosen feature space using a linear mapping.

Figure 18 presents representative propagation profiles obtained with kernel ridge regression for the Kecskemét region.

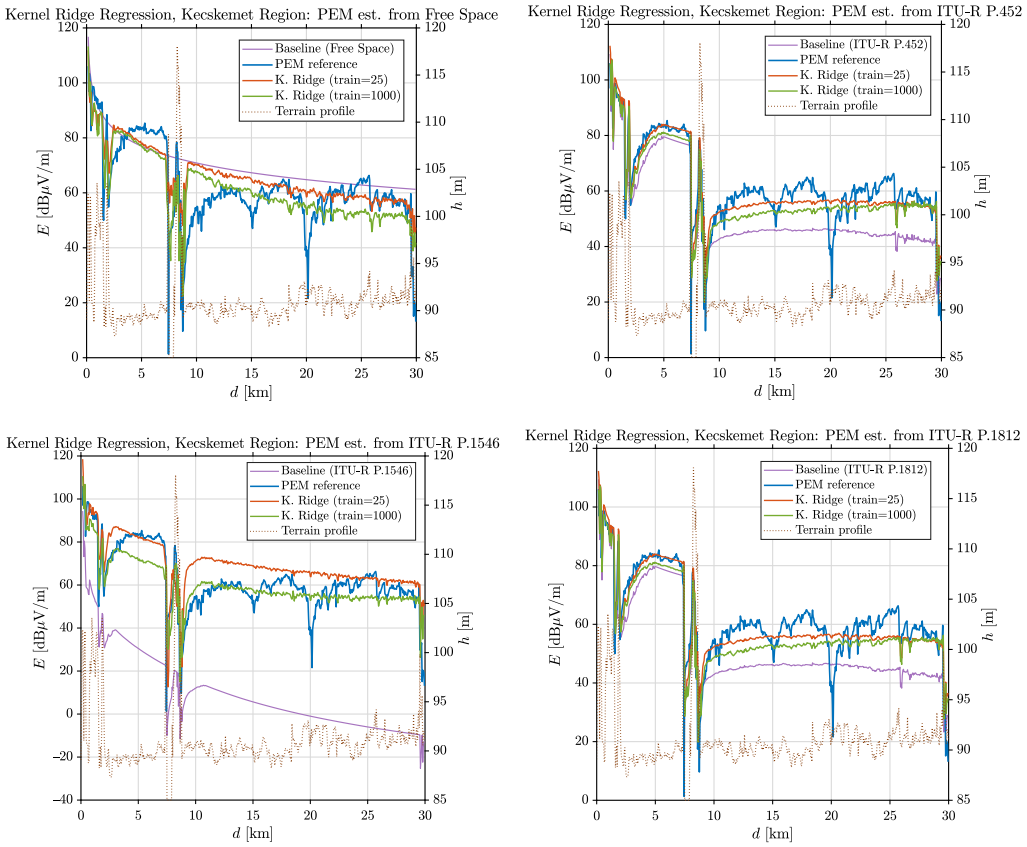


Figure 18. Examples for calculated profiles with kernel ridge regression, Kecskemét region.

The plots illustrate the baseline model predictions, the PEM reference solution, and the corrected estimates obtained using two different training set sizes. Similarly to the ridge regression results presented earlier, the learning-based correction substantially reduces the deviation of the baseline propagation models from the deterministic PEM reference. This effect is particularly visible for the empirical ITU-R P.1546 model and the free-space baseline, where the original predictions differ considerably from the reference solution. After applying the correction, the predicted curves shift towards the PEM reference and reproduce the dominant large-scale propagation behaviour along the path.

A comparison with the ridge regression results indicates that the kernel ridge formulation does not provide a noticeable qualitative improvement in the corrected profiles. For both small and large training sets, the estimated curves closely resemble those obtained with the linear ridge regression model. The remaining discrepancies relative to the PEM reference are mainly associated with small-scale fluctuations in the field strength, which are difficult to reproduce exactly by either regression approach.

This behaviour is consistent with the learning-curve analysis presented earlier. The similarity between the ridge and kernel ridge results suggests that the residual error between the baseline models and the PEM reference can already be represented effectively within the chosen feature space using a linear mapping. Consequently, the additional flexibility of the kernel formulation does not translate into improved prediction accuracy for this terrain type.

Last, but not least, Figure 19 shows the evolution of the prediction errors as a function of the number of training profiles for the Mátra region using kernel ridge regression.

The error metrics decrease rapidly as the number of training profiles increases and gradually converge towards a stable level. Compared to the flat-terrain case, the convergence is slightly slower, reflecting the increased propagation complexity of the mountainous environment, where diffraction and terrain shadowing lead to more complex residual structures.

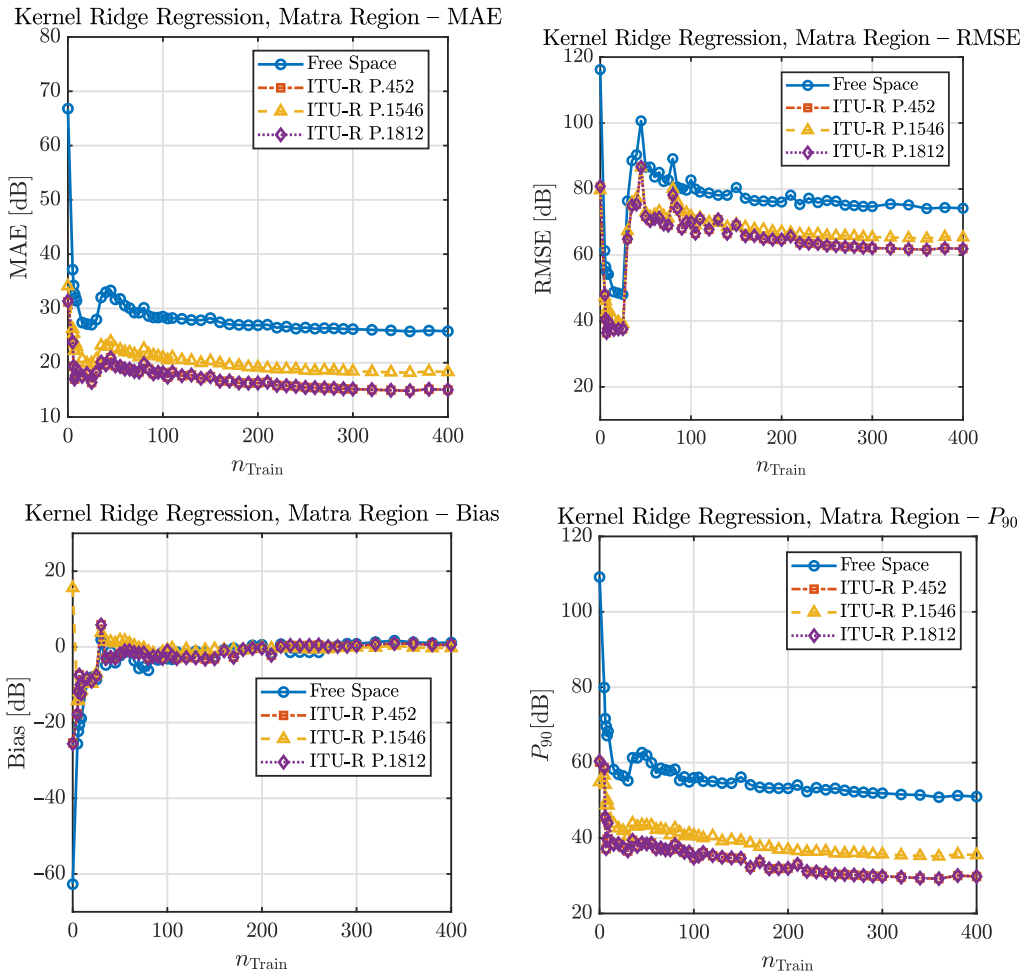


Figure 19. Error metrics as a function of the number of training profiles, Mátra region.

The magnitude of the correction depends strongly on the baseline model. The largest improvement is observed for the free-space and ITU-R P.1546 models, which exhibit large baseline deviations from the PEM reference. In contrast, the hybrid models ITU-R P.452 and ITU-R P.1812 already provide relatively accurate baseline predictions, and therefore only moderate improvements can be achieved through the learning-based correction.

A comparison with the ridge regression results shows that kernel ridge regression does not provide a measurable performance advantage in this region. The final error levels obtained with both approaches are nearly identical, indicating that the residual error between the baseline models and the PEM reference can already be represented effectively within the selected feature space using a linear mapping.

Figure 20 shows representative propagation profiles obtained with kernel ridge regression for the Mátra region.

The corrected profiles follow the dominant large-scale behaviour of the PEM reference and substantially reduce the deviation of the baseline models, particularly for the free-space and ITU-R P.1546 cases where the baseline errors are large in mountainous terrain. A visual comparison with the ridge regression results presented earlier indicates that the kernel ridge formulation produces very similar corrected profiles. The remaining discrepancies are mainly associated with strong local fluctuations caused by terrain-induced diffraction and shadowing, which are difficult to reproduce exactly by either regression approach.

Overall, the kernel ridge regression model provides stable and accurate delta corrections across all three terrain scenarios. The corrected predictions closely follow the deterministic PEM reference, confirming that the residual structure between baseline mod-

els and the deterministic solution can be effectively learned from terrain-profile-derived features. The magnitude of the correction increases with terrain complexity, with the largest deviations observed in the mountainous scenario, where diffraction and terrain-induced shadowing effects are strongest.

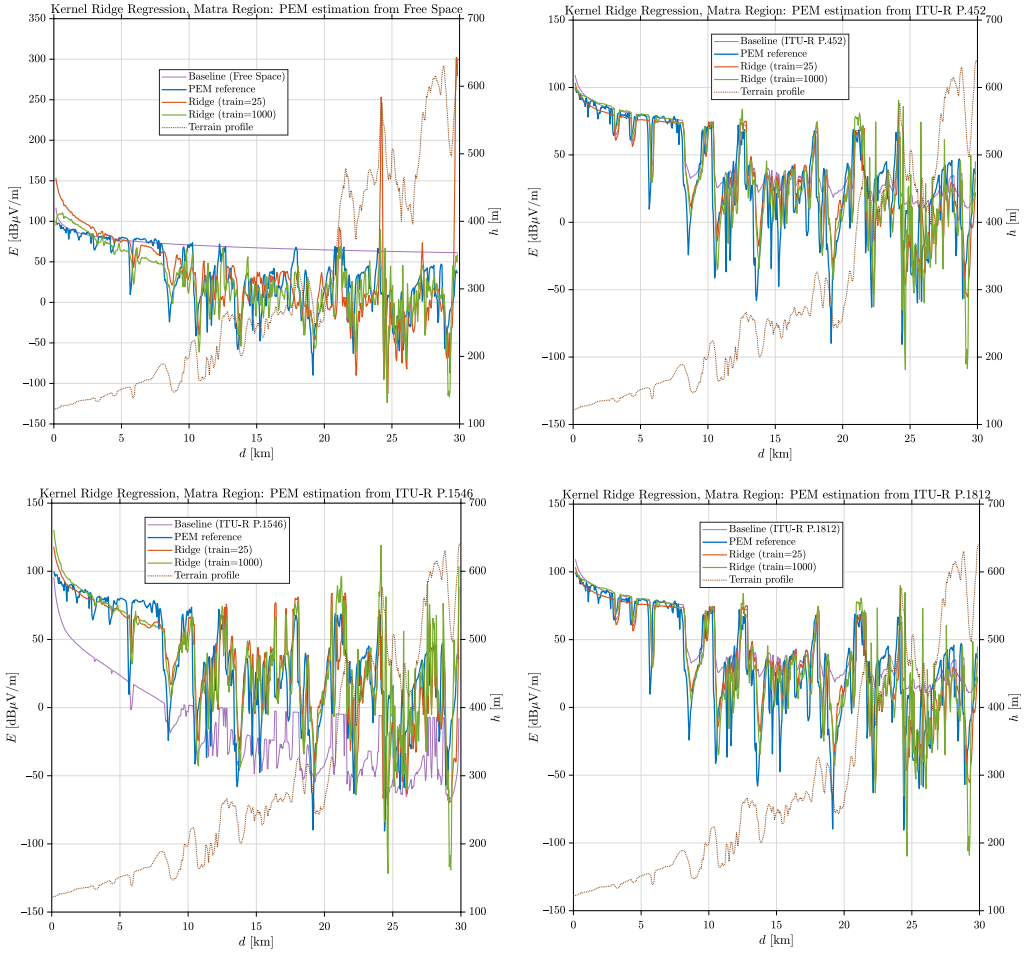


Figure 20. Examples for calculated profiles with kernel ridge regression, Mátra region.

At the same time, the kernelized formulation does not yield a substantial improvement compared to the linear ridge regression model. This indicates that the relationship between the engineered terrain features and the residual correction is largely linear within the chosen feature space. Consequently, most of the systematic baseline-model error can already be captured by the simpler linear model, while the remaining deviations are likely dominated by stochastic or small-scale propagation effects that cannot be represented by the available terrain descriptors.

8. Delta Correction with Gradient Boosting Regression

This section introduces the theoretical framework of the delta correction approach based on gradient boosting regression. In contrast to ridge regression models, which assume a global linear relationship between the feature vector and the residual correction term, gradient boosting constructs a flexible nonlinear model through an additive ensemble of weak learners.

Following the theoretical overview, we examine the performance of this model in addressing the investigated wave propagation delta-correction problem, using the analyses and metrics presented in the previous chapters.

8.1. Theoretical Overview and Algorithm Description

The gradient boosting model approximates the target mapping as a stage-wise additive expansion of the form

$$g_M(\phi) \;=\; \sum_{m=1}^{M} \nu\, h_m(\phi), \tag{81}$$

where $h_m : \mathbb{R}^{18} \to \mathbb{R}$ denotes the m-th weak learner, M is the number of boosting iterations, and $\nu \in (0, 1]$ is the learning rate controlling the contribution of each component [87]. In the present formulation, the weak learners are regression trees, which implement piecewise-constant approximations by recursively partitioning the feature space. Each tree therefore acts as a nonlinear basis function, enabling the model to capture interactions and local dependencies between the input features [88].

The additive structure implies that the model is constructed iteratively. Starting from an initial model

$$g_0(\phi) \;=\; \arg\min_{c} \sum_{n=1}^{NL} \ell(\delta_n, c), \tag{82}$$

the model is updated in a stage-wise manner as

$$g_m(\phi) \;=\; g_{m-1}(\phi) + \nu\, h_m(\phi). \tag{83}$$

The learning procedure can be interpreted as functional gradient descent in function space. At iteration m, the pseudo-residuals are defined as

$$r_n^{(m)} \;=\; -\left.\frac{\partial \ell(\delta_n, g(\phi_n))}{\partial g(\phi_n)}\right|_{g\,=\,g_{m-1}}. \tag{84}$$

For the squared error loss function

$$\ell(\delta_n, g(\phi_n)) \;=\; (\delta_n - g(\phi_n))^2, \tag{85}$$

the pseudo-residuals reduce to

$$r_n^{(m)} \;=\; \delta_n - g_{m-1}(\phi_n). \tag{86}$$

Thus, at each iteration, the next weak learner is obtained by solving

$$h_m \;=\; \arg\min_{h} \sum_{n=1}^{NL} \left(r_n^{(m)} - h(\phi_n)\right)^2. \tag{87}$$

In this work, $h_m(\cdot)$ is implemented as a regression tree [89]. Such a tree partitions the feature space $\mathbb{R}^{18}$ into a finite number of disjoint regions,

$$\mathbb{R}^{18} \;=\; \bigcup_{k=1}^{K_m} R_k^{(m)}, \quad R_k^{(m)} \cap R_l^{(m)} \;=\; \varnothing \quad \text{for } k \neq l, \tag{88}$$

and assigns a constant value to each region. The resulting function can be written as

$$h_m(\phi) \;=\; \sum_{k=1}^{K_m} c_k^{(m)}\, \mathbf{1}_{\{\phi \in R_k^{(m)}\}}, \tag{89}$$

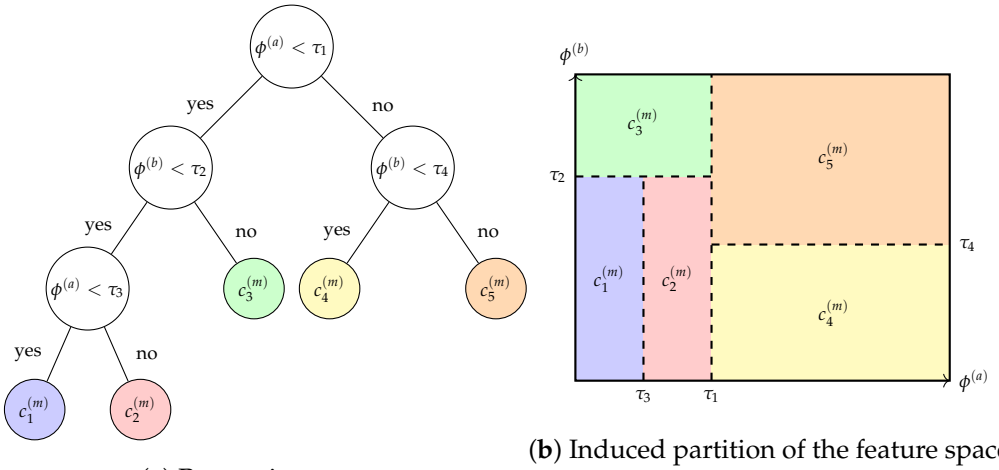
where $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function. The constants are computed as

$$c_k^{(m)} = \frac{1}{|R_k^{(m)}|} \sum_{\phi_n \in R_k^{(m)}} r_n^{(m)}. \tag{90}$$

This formulation shows that each regression tree defines a piecewise-constant approximation of the residual field. Geometrically, this corresponds to a partitioning of the feature space into axis-aligned regions, where each region is associated with a constant correction value. Each split is performed along one component of the feature vector ϕ, as illustrated in Figure 21. Here, $\phi^{(a)}$ denotes the a-th component of the feature vector $\phi_{i,j}$.

By combining multiple such trees, the final model can be expressed as

$$g_M(\phi) = \sum_{m=1}^{M} \nu \sum_{k=1}^{K_m} c_k^{(m)} \mathbf{1}_{\{\phi \in R_k^{(m)}\}}. \tag{91}$$



(**a**) Regression tree

(**b**) Induced partition of the feature space

Figure 21. Interpretation of a regression tree used as a weak learner in gradient boosting. (**a**) Tree structure defining threshold-based splits on selected components of the feature vector ϕ, with decision directions indicated on the edges. (**b**) Corresponding partition of a two-dimensional feature space spanned by $\phi^{(a)}$ and $\phi^{(b)}$, where each region is associated with a constant correction value. Colors indicate the correspondence between leaf nodes and regions.

This representation highlights that the gradient boosting model is a superposition of simple, locally defined corrections. Each individual tree provides a coarse approximation, while their superposition results in a progressively refined representation of the residual field [90].

The learning rate ν controls the contribution of each tree and therefore affects both convergence and generalization. The overall model complexity is determined by the number of boosting iterations M, the learning rate ν, and the complexity of the individual regression trees.

In the context of delta correction, the model is trained to approximate

$$\delta = E_{\text{ref}} - E_{\text{base}}, \tag{92}$$

and the final corrected prediction is obtained as

$$E_{\text{corr}} = E_{\text{base}} + g_M(\phi). \tag{93}$$

8.2. Results

The learning curves and the examples of the gradient boosting regression model (Figures 22 and 23) for the Budapest region show a consistently strong and well-balanced performance compared to both ridge and kernel ridge regression.

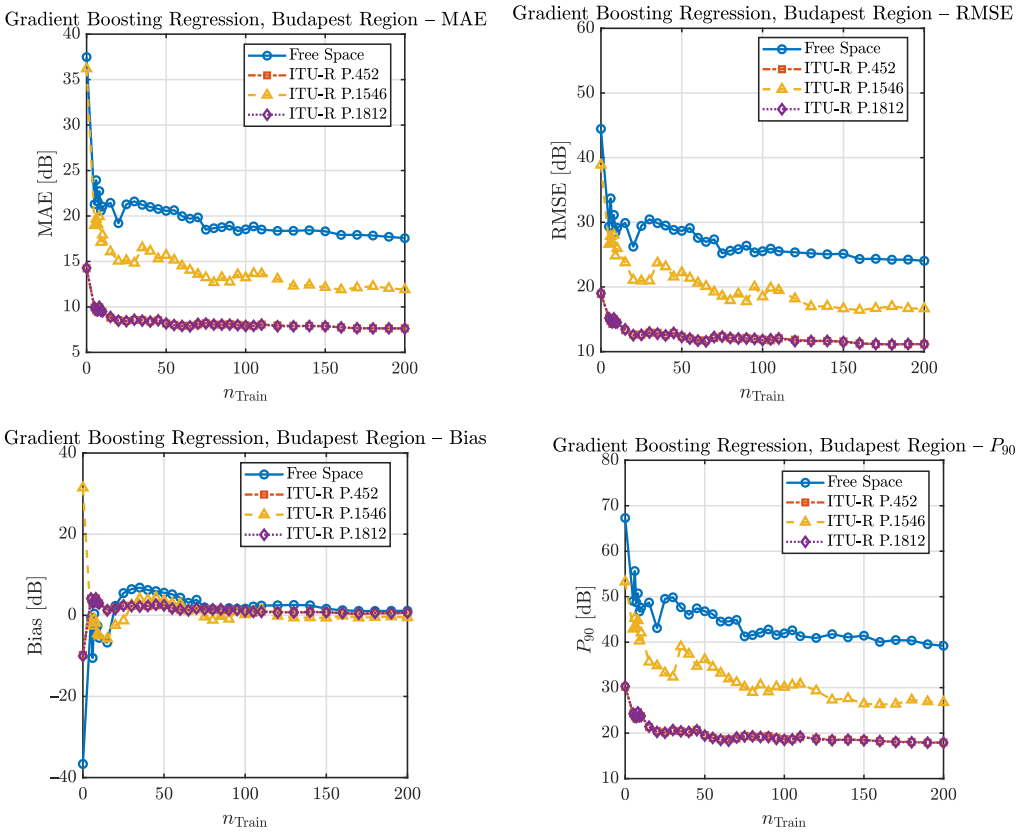


Figure 22. Error metrics as a function of the number of training profiles, Budapest region.

While ridge regression reaches a plateau already at low training sample sizes (around 10–15 samples) with MAE values of approximately 11 dB, and kernel ridge regression gradually improves to about 7–8 dB, gradient boosting achieves similar MAE levels (approximately 7–8 dB for ITU-R P.452 and ITU-R P.1812) but reaches this accuracy significantly earlier. This indicates a clear advantage in data efficiency.

More pronounced differences are observed in RMSE, where ridge regression remains in the range of 20–30 dB and kernel ridge regression reduces this to approximately 16–20 dB, while gradient boosting further decreases it to around 11–12 dB. This substantial improvement suggests a more effective handling of large prediction errors.

The primary benefit of gradient boosting lies instead in its faster convergence and lower RMSE, reflecting better control over error variance. This behaviour can be attributed to its locally adaptive structure, which allows it to model spatially varying and discontinuous propagation effects more effectively than both the global linear ridge model and the globally nonlinear kernel ridge approach.

The results show that gradient boosting provides the closest agreement with the PEM reference among the evaluated models.

The reconstructed profiles capture both the global decay trend and a significant portion of the local variations along the propagation path. In particular, rapid signal fluctuations and attenuation regions, often associated with terrain features, are reproduced with higher fidelity compared to the other methods.

In comparison, ridge regression mainly corrects the global trend but fails to represent local variations, while kernel ridge regression captures part of the nonlinear behaviour but still produces relatively smooth profiles. Gradient boosting differs by exhibiting clear local adaptivity, allowing it to follow sharp signal changes more closely.

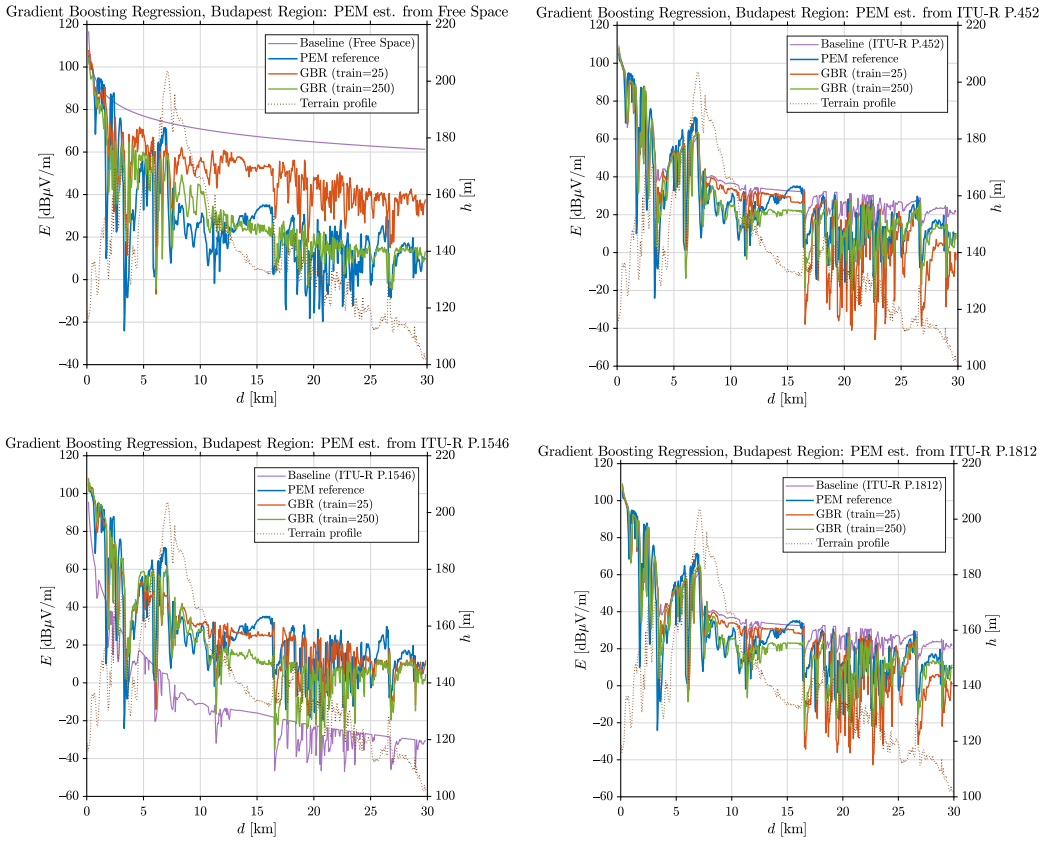


Figure 23. Examples for calculated profiles with gradient boosting regression, Budapest region.

Figure 24 presents the error metrics of the gradient boosting regression model for the Kecskemét region, while Figure 25 shows the profile examples.

In contrast to the Budapest results, all three models exhibit very rapid convergence, reaching stable performance already at low training sample sizes (approximately 10–15 samples).

The gradient boosting model achieves MAE values of approximately 4–6 dB for the ITU-based inputs, which are nearly identical to those obtained with ridge and kernel ridge regression. Similar behaviour is observed for RMSE and P_{90}, indicating that in the relatively homogeneous Kecskemét region, simple linear correction is already sufficient to capture the dominant propagation behaviour.

A small local increase in error is observed in the Free Space case around $n_{\text{train}} = 50$–60. This effect appears across all models and can be attributed to the sensitivity to training sample composition, as the lack of environment-specific information makes the correction more dependent on sample representativeness.

Overall, the results suggest that in this environment, more complex nonlinear and locally adaptive models provide only marginal improvement over linear approaches. This claim is supported by a comparison of the results of the depicted profiles with those of ridge regression.

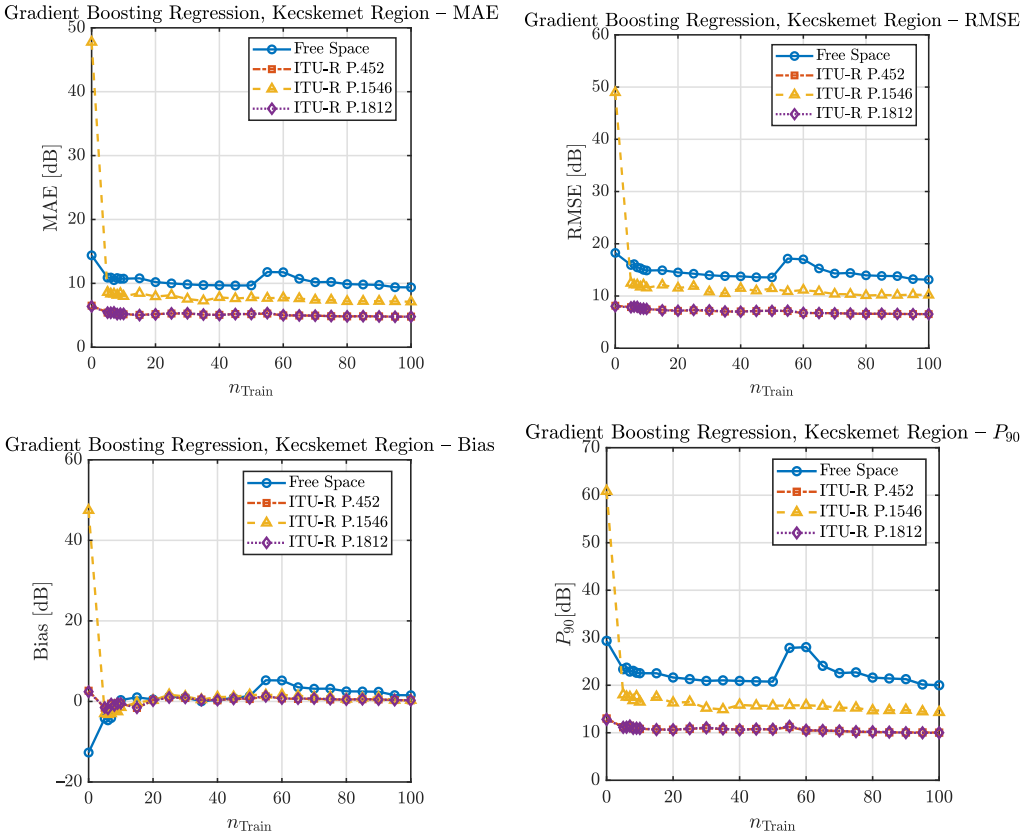


Figure 24. Error metrics as a function of the number of training profiles, Kecskemét region.

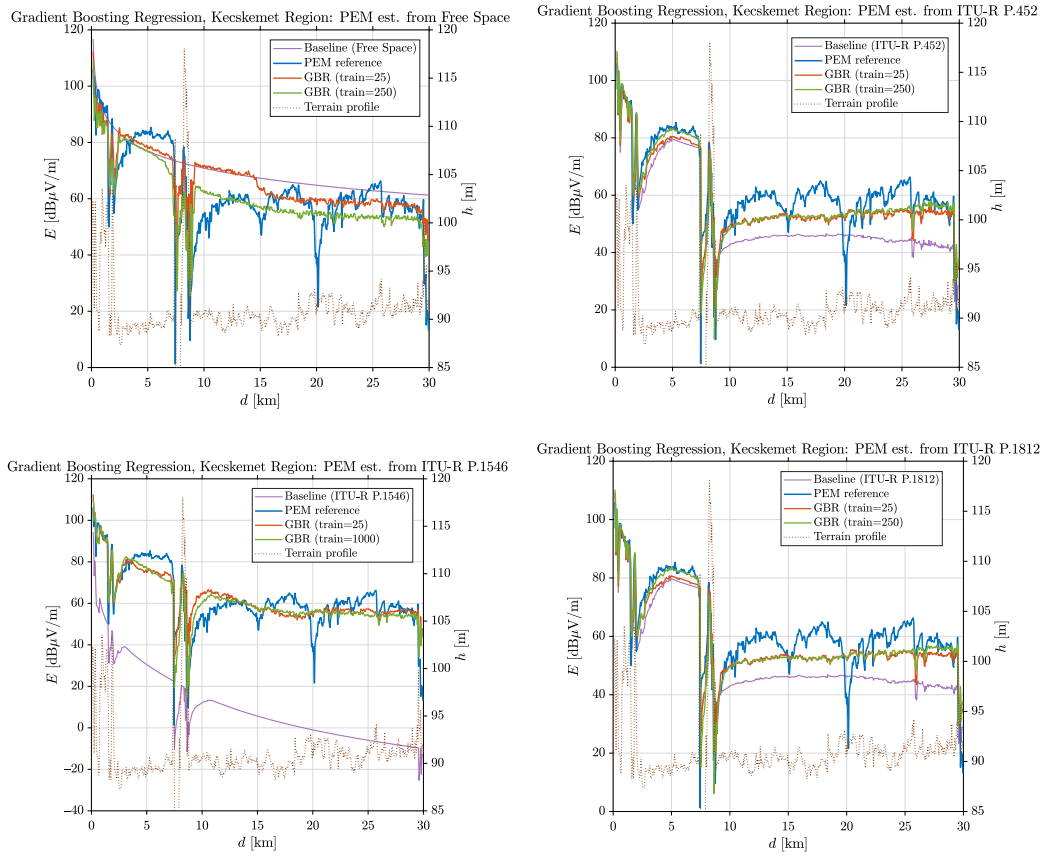


Figure 25. Examples for calculated profiles with gradient boosting regression, Kecskemét region.

https://doi.org/10.3390/technologies14060363

The performance of the Gradient Boosting Regression (GBR) model in the mountainous Mátra region shows that the error metrics stabilize already at relatively low training sample sizes. As illustrated in Figure 26, both MAE and RMSE exhibit a rapid initial decrease, followed by a plateau, indicating that the model is able to efficiently capture the dominant global trends with a limited number of training profiles. A similar behavior can be observed for the bias and percentile-based metrics, which also converge quickly toward stable values.

At the same time, the profile-level results in Figure 27 reveal that, despite the overall stability of the global metrics, significant local deviations remain present. The reconstructed field strength profiles follow the general trend of the PEM reference well, particularly when physically more consistent baseline models (e.g., ITU-R P.452 and P.1812) are used.

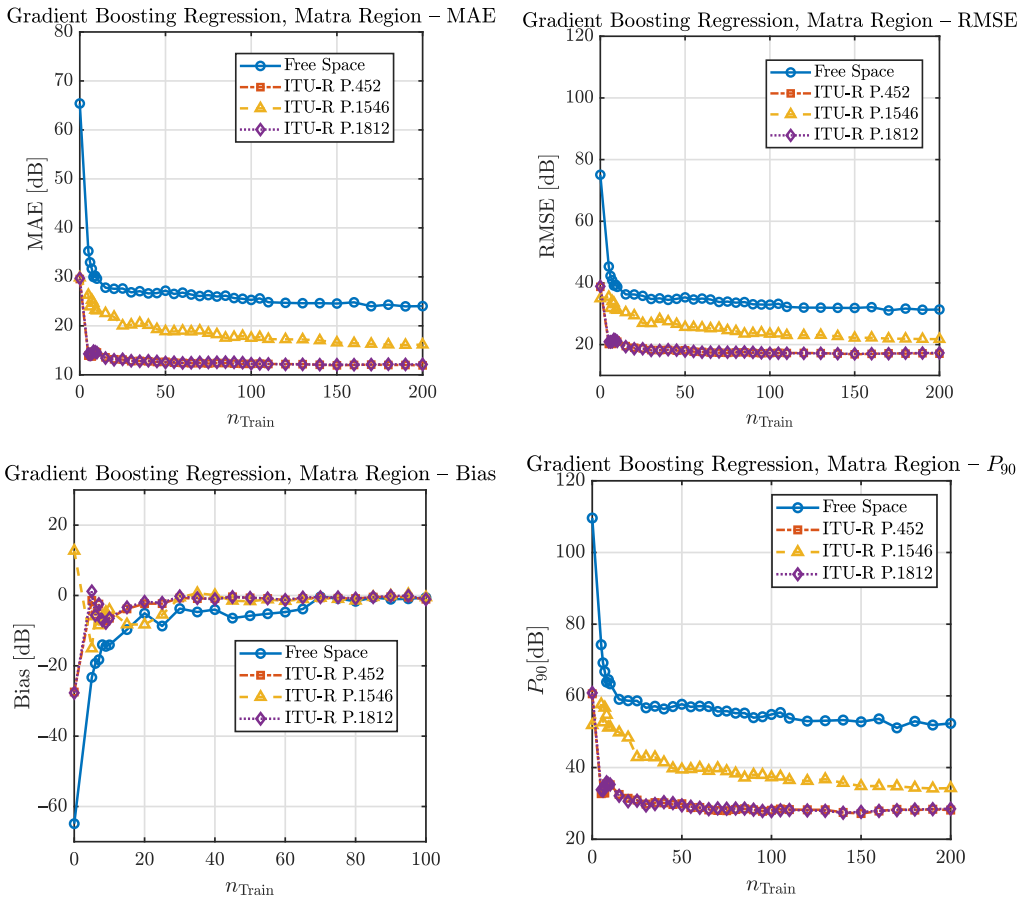


Figure 26. Error metrics as a function of the number of training profiles, Mátra region.

However, localized phenomena such as diffraction-induced minima and shadowing effects are only partially captured, leading to noticeable discrepancies in specific sections of the profiles.

These observations indicate that the GBR model provides a robust approximation of the overall propagation behavior, but its performance is fundamentally constrained by the complexity of the underlying physical processes. The remaining errors are primarily associated with localized terrain-induced effects, which are not fully represented in the input features. Consequently, further improvements are expected to depend less on the regression method itself and more on the incorporation of physically meaningful descriptors that better characterize the propagation environment.

To sum up: although GBR provides a stable approximation of the global behavior and reaches this stable regime already at relatively low training sample sizes, it does not yield a significant improvement over simpler regression methods, indicating that model

performance is primarily limited by the physical complexity of the propagation process rather than the choice of the regression algorithm.

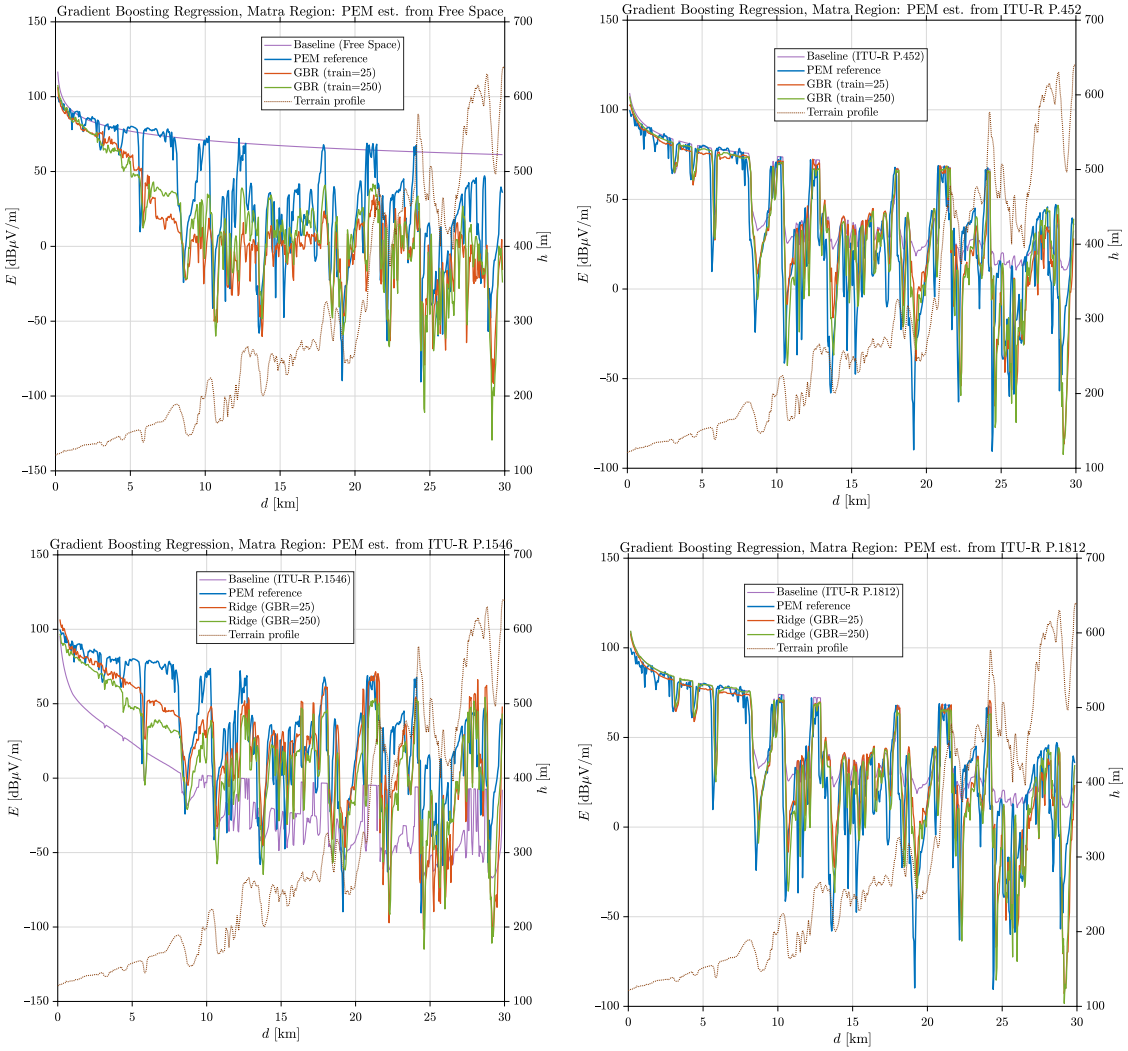


Figure 27. Examples for calculated profiles with gradient boosting regression, Mátra region.

9. Delta Correction with Multilayer Perceptron Regression

This section presents the short theoretical framework of the delta correction approach based on multilayer perceptron (MLP) regression. The aim of the implementation of this model is to capture complex nonlinear relationships between the input features and the residual correction term through a layered neural network structure.

Following the theoretical overview, the performance of the MLP-based correction is evaluated for the investigated propagation scenarios, using the same metrics and analysis framework introduced in the previous chapters.

9.1. Theoretical Overview and Algorithm Description

The nonlinear residual correction is formulated as the approximation of the mapping from the extended feature vector to the baseline error. Reusing the notation introduced in Section 4, for profile i and distance sample d_j, the residual is defined as

$$\delta_{i,j} \;=\; y_i(d_j) - b_i(d_j), \tag{94}$$

where $b_i(d_j)$ is the baseline prediction and $y_i(d_j)$ is the reference value. The learning task is therefore not the direct approximation of the reference field itself, but the approximation of the residual operator

$$f_\theta : \mathbb{R}^{18} \to \mathbb{R}, \qquad \hat{\delta}_{i,j} \;=\; f_\theta(\boldsymbol{\phi}_{i,j}), \tag{95}$$

with $\boldsymbol{\phi}_{i,j} \in \mathbb{R}^{18}$ denoting the extended feature vector defined previously as the concatenation of the local descriptor vector $\boldsymbol{\varphi}_{i,j} \in \mathbb{R}^8$ and the global profile descriptor $\mathbf{g}_i \in \mathbb{R}^{10}$. In this way, the model operates entirely within the residual-learning framework introduced earlier, and remains fully consistent with the additive reconstruction

$$\hat{y}_i(d_j) \;=\; b_i(d_j) + \hat{\delta}_{i,j}. \tag{96}$$

This additive structure is also shown explicitly in Figure 28, where the output of the neural network is interpreted as a correction term rather than a standalone propagation estimate [91].

To define the training set in a compact form, a composite index $n \;=\; (i, j)$ is introduced, with $n \;=\; 1, \ldots, NL$, yielding

$$\mathcal{D} \;=\; \{(\boldsymbol{\phi}_n, \delta_n)\}_{n=1}^{NL}, \tag{97}$$

where $\boldsymbol{\phi}_n \;=\; \boldsymbol{\phi}_{i,j}$ and $\delta_n \;=\; \delta_{i,j}$. Since each terrain profile contains L strongly correlated spatial samples, the regression problem is pointwise only in its algebraic form; statistically, the samples are structured and grouped by profile. Therefore, data partitioning is performed at the profile level; i.e., entire terrain profiles are assigned either to the training pool or to the fixed test set.

The regression function f_θ is implemented as a feedforward multilayer perceptron with two hidden layers. Let the standardized input vector be $\mathbf{z}_{i,j} \in \mathbb{R}^{18}$, obtained from $\boldsymbol{\phi}_{i,j}$ by feature-wise normalization:

$$z_{n,m} \;=\; \frac{\phi_{n,m} - \mu_m}{\sigma_m}, \qquad m \;=\; 1, \ldots, 18, \tag{98}$$

or in vector form

$$\mathbf{z}_n \;=\; \mathbf{D}_\sigma^{-1}(\boldsymbol{\phi}_n - \boldsymbol{\mu}), \tag{99}$$

where $\boldsymbol{\mu} \in \mathbb{R}^{18}$ is the vector of feature means and $\mathbf{D}_\sigma \;=\; \mathrm{diag}(\sigma_1, \ldots, \sigma_{18})$ is the diagonal matrix of standard deviations [92].

The forward propagation through the network is defined as

$$\mathbf{h}_n^{(1)} = \tanh\Big(\mathbf{W}^{(1)}\mathbf{z}_n + \mathbf{b}^{(1)}\Big), \tag{100}$$

$$\mathbf{h}_n^{(2)} = \tanh\Big(\mathbf{W}^{(2)}\mathbf{h}_n^{(1)} + \mathbf{b}^{(2)}\Big), \tag{101}$$

$$\hat{\delta}_n = \mathbf{W}^{(3)}\mathbf{h}_n^{(2)} + b^{(3)}, \tag{102}$$

with parameter dimensions

$$\mathbf{W}^{(1)} \in \mathbb{R}^{40\times 18}, \qquad \mathbf{b}^{(1)} \in \mathbb{R}^{40}, \tag{103}$$

$$\mathbf{W}^{(2)} \in \mathbb{R}^{20\times 40}, \qquad \mathbf{b}^{(2)} \in \mathbb{R}^{20}, \tag{104}$$

$$\mathbf{W}^{(3)} \in \mathbb{R}^{1\times 20}, \qquad b^{(3)} \in \mathbb{R}. \tag{105}$$

In compact form, the network realizes the mapping

$$f_\theta(\mathbf{z}_n) = \mathbf{W}^{(3)} \tanh\Big(\mathbf{W}^{(2)} \tanh\Big(\mathbf{W}^{(1)}\mathbf{z}_n + \mathbf{b}^{(1)}\Big) + \mathbf{b}^{(2)}\Big) + b^{(3)}. \tag{106}$$

The hidden layers use the hyperbolic tangent activation function, corresponding to the `tansig` transfer function in MATLAB 9.7.0.1190202 (R2019b), while the output layer is linear (`purelin`). This structure is consistent with Figure 28, where nonlinear transformations are applied in the hidden layers and a linear mapping is used at the output [93,94].

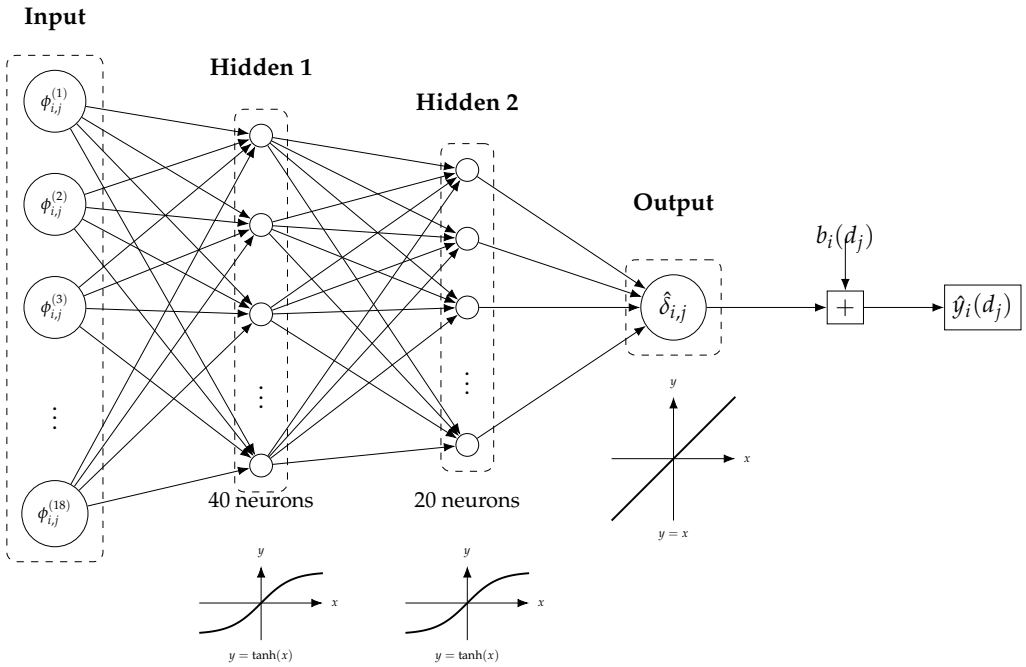


Figure 28. Multilayer perceptron used for delta correction. The network maps the feature vector $\boldsymbol{\phi}_{i,j}$ to the residual $\hat{\delta}_{i,j}$, which is added to the baseline model output $b_i(d_j)$ to obtain the corrected prediction $\hat{y}_i(d_j)$. The inset plots illustrate the activation functions used in the hidden and output layers.

Let $\mathcal{I}_{\text{tr}} \subset \{1, \ldots, N\}$ and $\mathcal{I}_{\text{te}} \subset \{1, \ldots, N\}$ denote the sets of terrain profile indices assigned to the training and test sets, respectively, with $\mathcal{I}_{\text{tr}} \cap \mathcal{I}_{\text{te}} = \emptyset$. The corresponding set of training samples is defined as

$$\mathcal{T} = \{(i, j) \mid i \in \mathcal{I}_{\text{tr}},\ j = 1, \ldots, L\}, \tag{107}$$

so that $|\mathcal{T}| = N_{\text{tr}} L$, where $N_{\text{tr}} = |\mathcal{I}_{\text{tr}}|$. This formulation reflects the profile-level data split, while the loss is evaluated over all distance samples within each training profile.

The model parameters θ are obtained by minimizing the empirical mean squared error

$$\mathcal{L}(\theta) = \frac{1}{|\mathcal{T}|} \sum_{n \in \mathcal{T}} (\delta_n - f_\theta(\mathbf{z}_n))^2, \tag{108}$$

or equivalently,

$$\mathcal{L}(\theta) = \frac{1}{N_{\text{tr}} L} \sum_{i \in \mathcal{I}_{\text{tr}}} \sum_{j=1}^{L} \big(y_i(d_j) - \big[b_i(d_j) + f_\theta(\mathbf{z}_{i,j})\big]\big)^2. \tag{109}$$

The optimization is performed using scaled conjugate gradient (SCG) backpropagation [95]. Denoting the parameter vector at iteration t by $\theta^{(t)}$, the update rule is

$$\theta^{(t+1)} = \theta^{(t)} + \alpha_t \mathbf{p}_t, \tag{110}$$

where $\mathbf{p}_t$ is a conjugate search direction and α_t is the SCG scaling coefficient determined internally by the algorithm [96].

After training, the corrected prediction on the test set is obtained as

$$\hat{y}_i(d_j) \;=\; b_i(d_j) + \hat{\delta}_{i,j}, \qquad i \in \mathcal{I}_{\text{te}},\; j = 1, \ldots, L, \tag{111}$$

and evaluated using the same profile-level metrics as the baseline model.

9.2. Results

The error metrics of the MLP regression model as a function of the training set size for the Budapest region and the sample results are shown in Figures 29 and 30.

Looking at the results, two key observations can be made. First, for small training sample sizes, the model's error metrics are unstable and fluctuate between extreme values, especially compared to the previously examined models. As the sample size increases, the metrics clearly converge toward their minimum values. With sufficient data, already around 500 terrain profiles, the model yields significantly better results than the earlier approaches. For example, the ITU-R P.452 and ITU-R P.1812 models achieve MAE values below 10 dB. Even the free space propagation model shows notable improvement, with MAE dropping well below 20 dB. Although gradient boosting regression produces similar error metrics, the MLP proves more robust and less sensitive to outliers or locally erroneous input data.

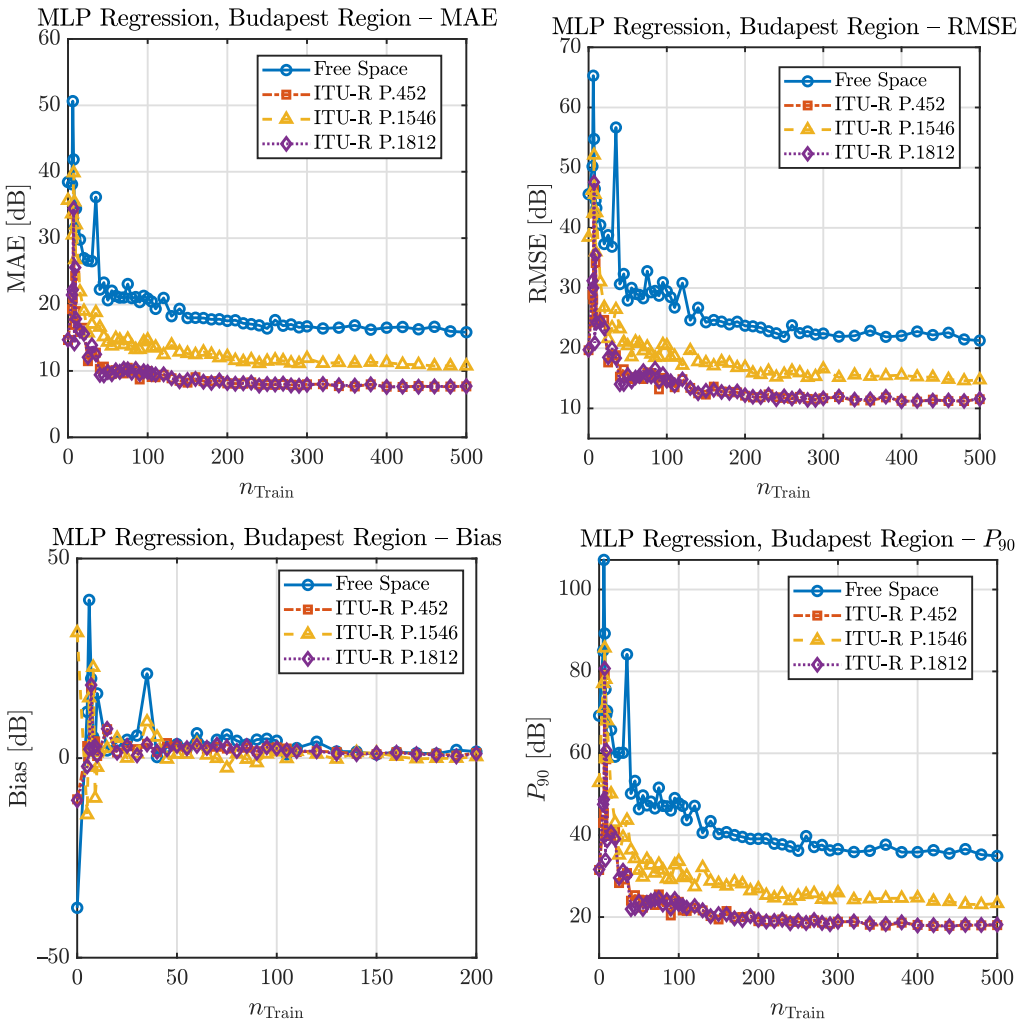


Figure 29. Error metrics as a function of the number of training profiles, Budapest region.

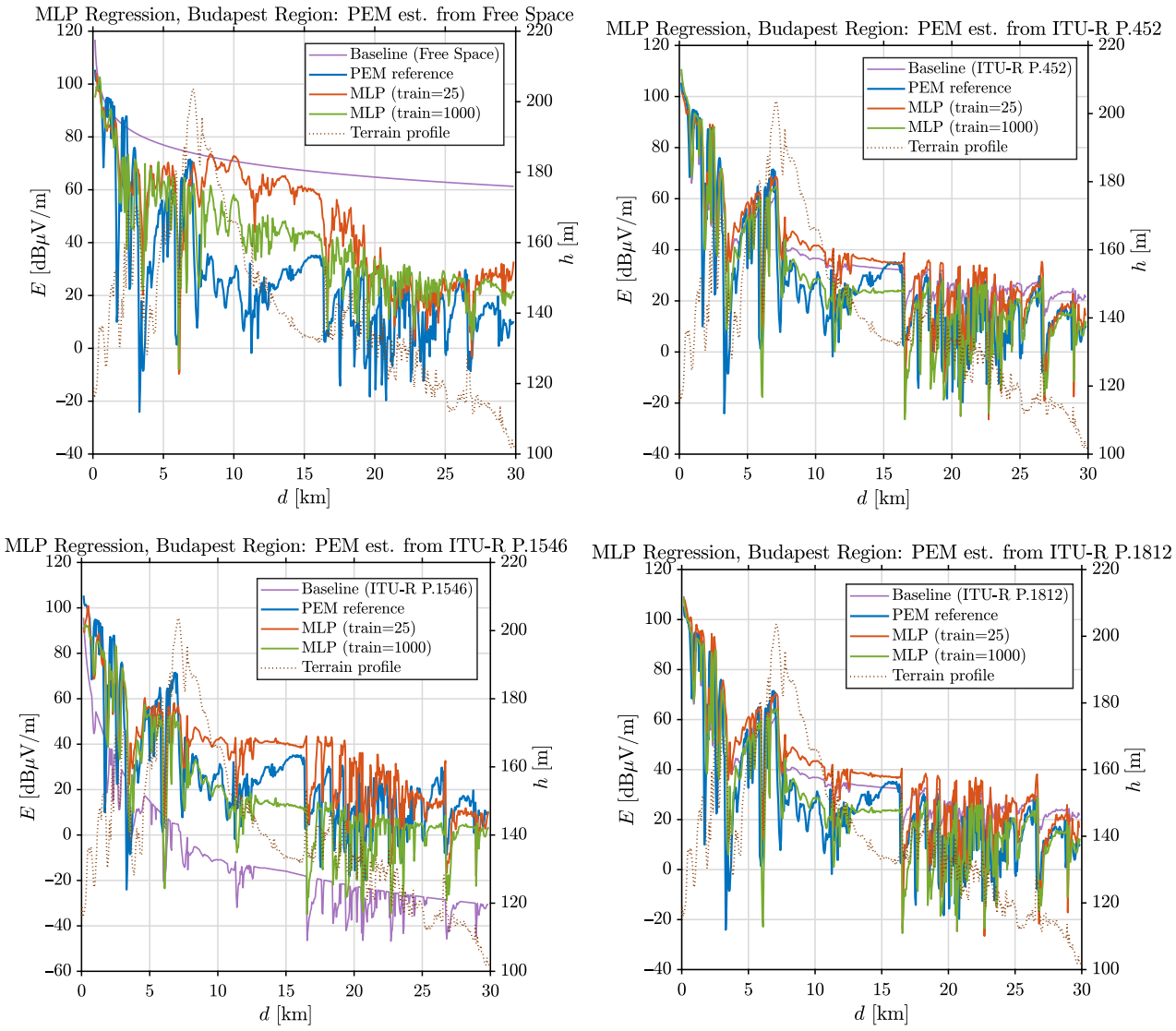


Figure 30. Examples for calculated profiles with MLP regression, Budapest region.

In all cases, the baseline predictions exhibit systematic deviations from the reference and fail to capture local terrain-induced variations along the propagation path, particularly in regions with rapid changes in signal level.

The MLP-based correction significantly improves the agreement with the PEM results across all models. Even with a small training set (train = 25), the dominant bias is reduced and the overall trend of the reference solution is more accurately followed, although fine-scale variations are still only partially reproduced. With a larger training set (train = 1000), the corrected curves closely track the PEM reference over the entire distance range, capturing both the global behavior and a substantial portion of the local fluctuations associated with terrain features.

The improvement is consistent across all examined baseline models, including both empirical and hybrid approaches, indicating that the proposed method is robust with respect to the underlying propagation model. These visual observations are in good agreement with the previously reported improvements in the error metrics.

In the Kecskemét region, the MLP regression demonstrates strong performance, with relatively fast convergence as the number of training samples increases, as illustrated in Figure 31. The model stabilises already at moderate training sizes (∼30–50 samples),

achieving MAE values of approximately 5–7 dB and RMSE in the range of 8–10 dB. The bias rapidly approaches zero, while the P_{90} metric indicates low extreme errors (∼10–12 dB).

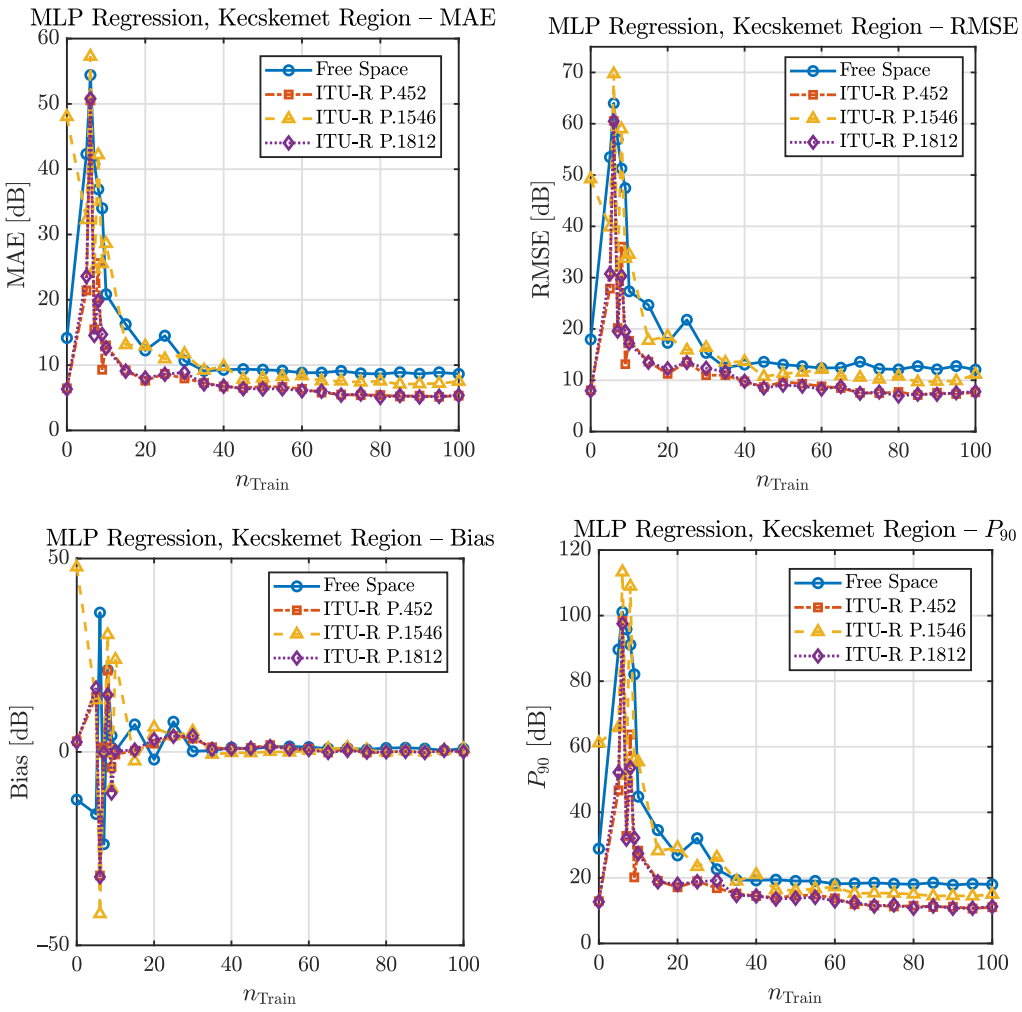


Figure 31. Error metrics as a function of the number of training profiles, Kecskemét region.

Compared to the other models, the MLP provides a clear improvement over linear and kernel-based approaches, while its performance remains comparable to gradient boosting. Although gradient boosting is slightly more stable at very small training sizes, the difference becomes negligible as the training set grows.

The qualitative illustration example between predicted and reference propagation profiles further supports these findings, as shown in Figure 32, where the MLP accurately captures the main variations in the signal strength along the path.

Overall, the MLP can be considered a top-performing model for the Kecskemét region; however, it does not consistently outperform gradient boosting.

In the Mátra region, the multilayer perceptron (MLP) regression exhibits stable but noticeably slower convergence compared to the flatter Kecskemét case, as the number of training samples increases (Figure 33). This behaviour can be attributed to the increased variability and structural complexity of the terrain, which requires a richer representation to be learned by the model. As a result, the MLP requires a larger training set (∼80–120 samples) to reach a stable regime, where further increases in the training size yield only marginal improvements.

In the converged regime, the error metrics reach MAE values of approximately 12–14 dB and RMSE around 18–20 dB. While these values are higher than those observed in less complex regions, they remain consistent and stable, indicating that the model success-

fully captures the dominant propagation characteristics. The bias gradually approaches zero, suggesting that systematic errors are effectively mitigated. At the same time, the P_{90} metric remains relatively high (∼28–32 dB), reflecting the presence of strong local variations and extreme propagation conditions that are more difficult to model accurately in mountainous environments.

Compared to the other models, the MLP provides a clear improvement over linear ridge regression and outperforms kernel ridge regression in terms of overall accuracy. Its performance is again comparable to that of gradient boosting, with only minor differences observed in the converged regime, where both models reach a similar level of prediction error. However, a key distinction emerges in the learning dynamics: gradient boosting demonstrates more stable behaviour at smaller training sizes, indicating better robustness in the low-data regime, whereas the MLP requires a larger dataset to achieve a comparable level of generalisation. This suggests that the MLP is more sensitive to the availability of training data and benefits more strongly from increased sample size.

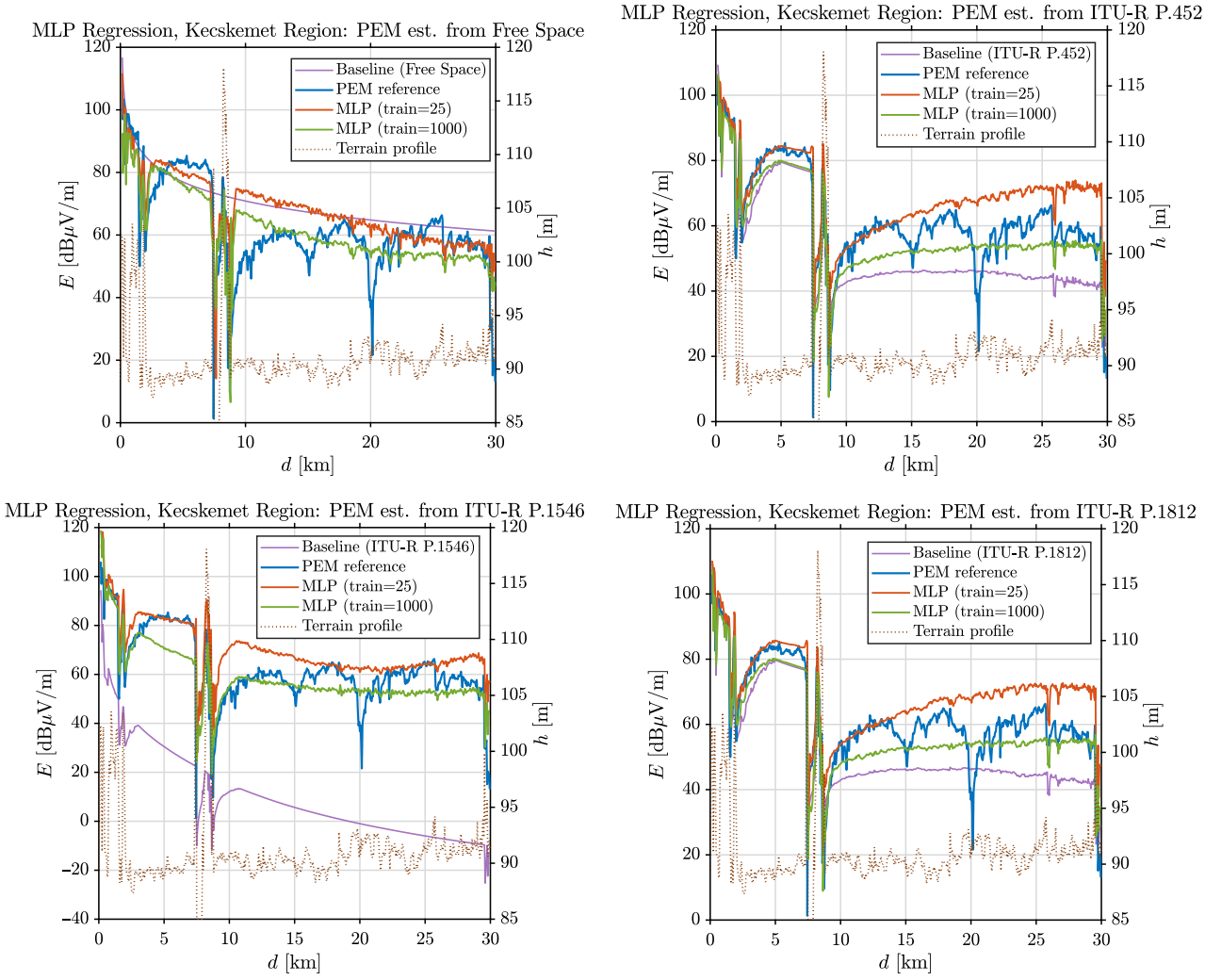


Figure 32. Examples for calculated profiles with MLP regression, Kecskemét region.

The qualitative difference between predicted and reference propagation profiles further supports these observations (Figure 34).

The MLP is able to capture the dominant large-scale variations in the signal along the propagation path, including general attenuation trends and major terrain-induced effects. At the same time, larger local deviations can still be observed in sections with highly irregular terrain, where rapid elevation changes and complex diffraction effects

occur. In such regions, the predictions tend to exhibit slight smoothing, indicating that fine-scale variations are not fully captured.

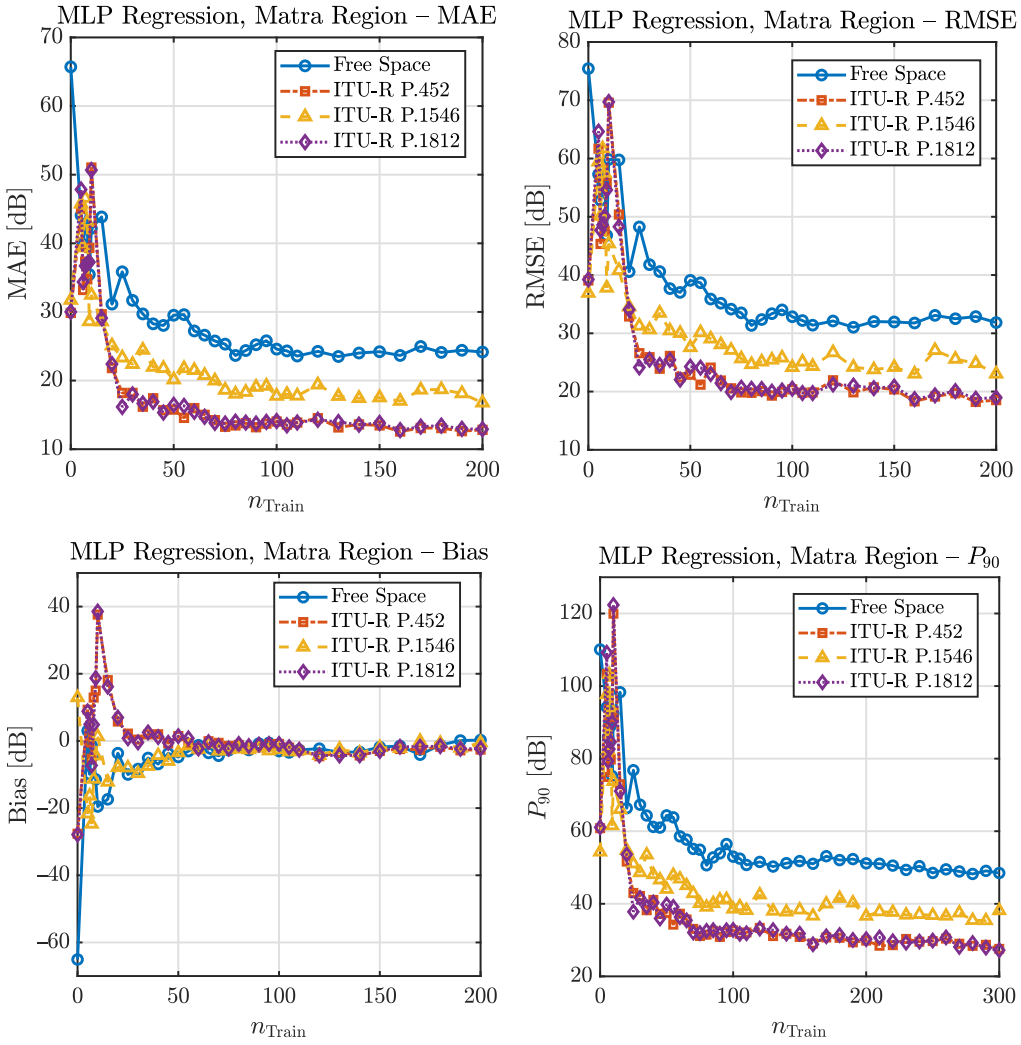


Figure 33. Error metrics as a function of the number of training profiles, Mátra region.

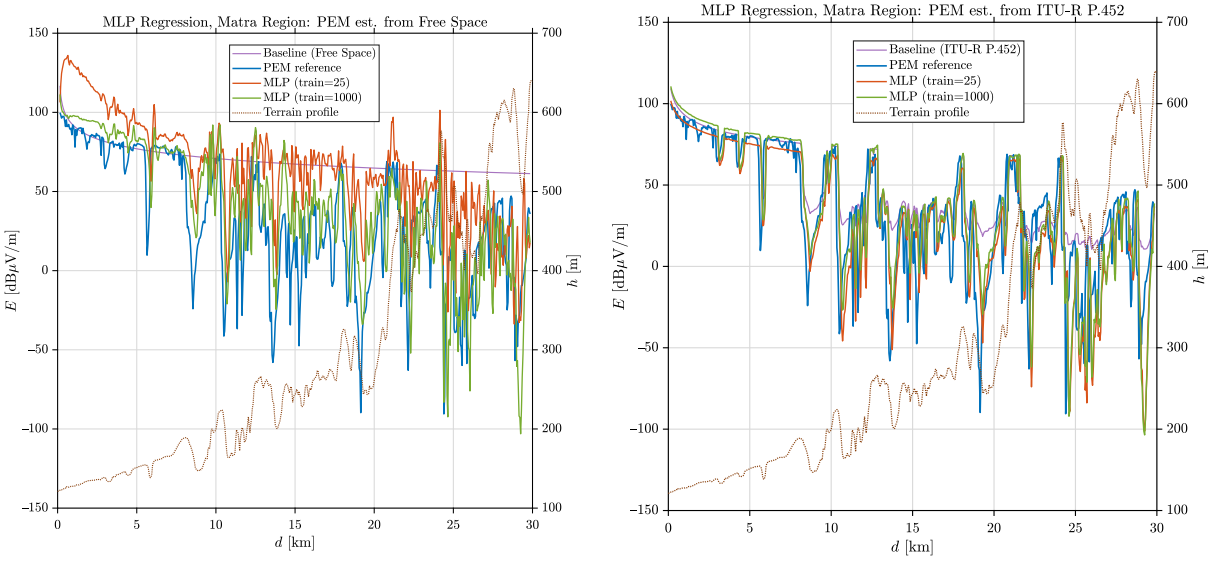


Figure 34. *Cont.*

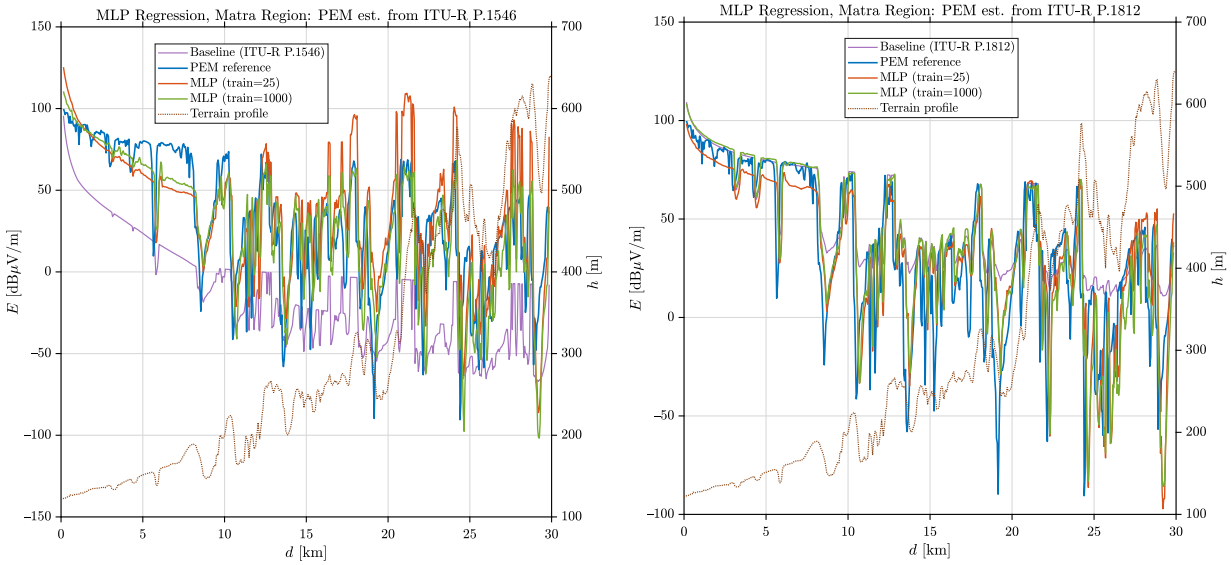


Figure 34. Examples for calculated profiles with MLP regression, Mátra region.

Overall, the MLP can be considered a robust and competitive model for the Mátra region, particularly when sufficient training data are available. Nevertheless, similarly to the other investigated regions, it does not demonstrate a consistent or systematic advantage over gradient boosting. The results therefore suggest that while neural network-based approaches are capable of achieving high accuracy, tree-based ensemble methods remain highly competitive, especially in scenarios with limited training data or increased terrain complexity.

10. Conclusions

Within the scope of this study, we conducted a detailed evaluation of the performance of four machine learning regression algorithms for delta-correction of empirical and hybrid wave propagation models at a frequency of 3.6 GHz. We developed a terrain representation methodology incorporating both local and global descriptors, enabling the regression algorithms to be effectively trained for delta learning. The proposed approach, together with the four regression models, was implemented and tailored to the specific problem, and its performance was subsequently validated using real-world measurement data.

Table 5 presents the best achievable values of the error metrics based on the learning curves, with the top-performing results for each terrain category highlighted in bold. Several conclusions can be drawn from these results.

Although this can be anticipated a priori, the results also confirm that regression models provide more accurate corrections for hybrid wave propagation models; therefore, whenever available, these models are recommended as the preferred baseline.

It can also be observed that under simpler, flat terrain conditions, the implementation and application of more complex regression models are unnecessary: even linear ridge regression yields strong performance, with kernel ridge regression providing only marginal improvement. In contrast, for more complex terrains, the use of gradient boosting regression and/or multilayer perceptron (MLP) models is recommended, as they achieve significantly better performance than ridge and kernel ridge regression.

It is also worth noting that an input feature importance analysis was conducted for each model, allowing us to investigate how the individual components of the 18-dimensional feature vector influence the model behavior and the accuracy of the predictions. This dependence is not solely determined by the regression model itself, but is also strongly affected by the type of baseline model and the terrain category under consideration. Here,

we only briefly remark that our results confirmed the relevance of all components of the feature vector, as each of them is utilized by the regression models, albeit with different weights. A more detailed analysis of this aspect could be carried out; however, we consider it to be beyond the scope of the present study.

Table 5. Stabilization behaviour of machine learning models across different regions (approximate values).

Propagation Model	Budapest			Kecskemét			Mátra		
	RMSE (dB)	**MAE (dB)**	**P90 (dB)**	**RMSE (dB)**	**MAE (dB)**	**P90 (dB)**	**RMSE (dB)**	**MAE (dB)**	**P90 (dB)**
Ridge (n_{Train}: ∼50/∼50/∼100)									
Free Space	30.0	19.5	40.5	12.0	9.0	19.0	76.5	28.0	54.5
ITU-R P.452	20.5	11.5	23.0	7.0	5.0	11.0	65.0	19.0	35.0
ITU-R P.1546	22.5	13.5	27.5	9.5	7.5	15.0	67.0	20.5	39.5
ITU-R P.1812	20.5	11.5	23.0	7.0	5.0	11.0	65.0	19.0	35.0
Kernel Ridge (n_{Train}: ∼300/∼300/∼500)									
Free Space	28.0	16.5	36.5	11.0	8.0	17.5	74.0	26.0	50.5
ITU-R P.452	16.5	8.0	18.5	**6.0**	**4.5**	**9.5**	62.5	15.0	29.5
ITU-R P.1546	20.5	11.0	22.5	8.0	6.0	12.5	65.5	18.0	35.0
ITU-R P.1812	18.5	8.0	19.0	**6.0**	**4.5**	**9.5**	62.0	15.0	30.0
Gradient Boost (n_{Train}: ∼700/∼700/∼800)									
Free Space	22.5	16.5	36.0	12.0	8.5	18.5	32.0	24.0	52.0
ITU-R P.452	**10.5**	**7.5**	**17.0**	6.5	**4.5**	**9.5**	**16.5**	**11.5**	27.0
ITU-R P.1546	14.5	10.5	23.5	9.0	6.5	13.0	20.5	15.0	33.0
ITU-R P.1812	**10.5**	**7.5**	**17.0**	6.5	**4.5**	**9.5**	**16.5**	**11.5**	27.0
MLP (n_{Train}: ∼900/∼900/∼1000)									
Free Space	20.5	15.5	33.5	10.5	7.5	16.5	29.0	22.0	47.5
ITU-R P.452	**10.5**	**7.5**	17.5	**6.0**	**4.5**	**9.5**	**16.5**	**11.5**	26.5
ITU-R P.1546	14.0	10.0	22.0	8.0	6.0	12.5	20.0	15.0	32.0
ITU-R P.1812	**10.5**	**7.5**	17.5	**6.0**	**4.5**	**9.5**	**16.5**	**11.5**	**26.0**

Another limitation of the present study is that the proposed framework has not yet been validated against independent real-world field measurements. Although PEM provides a physically grounded deterministic reference and is suitable for assessing systematic deviations of faster empirical and hybrid models under controlled and reproducible conditions, measurement-based validation would be required to quantify the absolute accuracy of the corrected predictions in operational deployment scenarios. Such validation, including comparisons against measured field-strength or path-loss data, is therefore identified as an important direction for future work.

The framework presented in this paper therefore offers a practically applicable toolkit for the fast and efficient delta-correction of wave propagation models. At the same time, the proposed framework extends beyond the specific scope defined in this study: given the availability of more accurate, higher-resolution terrain data and measurement results, it enables the alignment of fast computational models with empirical observations. This direction also defines the continuation of our research, where future work will focus on integrating measurement data and on the development and validation of improved modeling approaches.

The proposed method should not be interpreted as a universally transferable correction model trained in one terrain type or geographical region and directly applicable to all others. Its demonstrated strength lies in morphology-aware, profile-level generalization,

where the correction model is trained and applied within a terrain class identified from available topographical or morphological information. Cross-terrain and cross-country transferability remain important topics for future investigation.

Author Contributions: Conceptualization, T.I.U.; methodology, T.I.U.; software, T.I.U.; validation, T.I.U.; formal analysis, T.I.U.; investigation, T.I.U.; resources, T.I.U.; data curation, T.I.U.; writing—original draft preparation, T.I.U.; writing—review and editing, T.I.U.; visualization, T.I.U.; supervision, M.K.; project administration, T.I.U.; funding acquisition, T.I.U. 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 study are available within the article. Further inquiries can be directed to the corresponding authors.

Acknowledgments: The author, Tamás István Unger, gratefully acknowledges the National Media and Infocommunications Authority of Hungary (NMHH) for providing a calm, stable, and supportive working environment that enables sustained and in-depth professional research. The author also expresses sincere appreciation for the opportunity to utilize the available software resources in support of his research and scientific work. Particular thanks are extended to Irén Bálint, for her continuous personal and departmental support, as well as for her professional recognition of the author's research activities. The atmosphere of stability, tranquility, and encouragement provided within the unit has significantly contributed to the author's ability to engage deeply in his field of research, ultimately supporting the completion of this paper. The author is especially grateful to Péter Vári, for his active and unwavering support of the author's professional and academic development from the very beginning, spanning more than a decade. The author would like to express his heartfelt gratitude to his wife for her patience, understanding, and constant support, and for bearing with the many hours dedicated to this work, often drawn from evenings and weekends. Finally, the author expresses his sincere gratitude to Széchenyi István University, in particular to the Doctoral School of Multidisciplinary Engineering Sciences (MMTDI), for providing the institutional framework for his research, and for their financial support, which made the publication of this paper possible.

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

References

1. Lim, C.; Ranaweera, C. Radio access network convergence. In *Handbook of Radio and Optical Networks Convergence*; Springer: Singapore, 2024; pp. 277–308. [CrossRef]
2. Dzemydienė, D. Wireless Communication Network Infrastructure Analysis for Smart Service System Functioning. In *International Series in Operations Research and Management Science*; Springer: Cham, Switzerland, 2022; Volume 330, pp. 29–60. [CrossRef]
3. Lopez, J.; Alcaraz, C.; Roman, R. On the protection and technologies of critical information infrastructures. In *Proceedings of the International Conference on Critical Information Infrastructures Security*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2007; Volume 4677, pp. 160–182. [CrossRef]
4. Bengi, U.; Kosu, S. Smart IoT metering applications in next-generation communication for critical infrastructure resilience. In Proceedings of the 33rd IEEE Conference on Signal Processing and Communications Applications (SIU), Istanbul, Turkiye, 25–28 June 2025. [CrossRef]
5. Rawat, D.B.; Bista, B.B.; Yan, G. Introduction to mobile and wireless communications networks. In *Security, Privacy, Trust, and Resource Management in Mobile and Wireless Communications*; IGI Global: Hershey, PA, USA, 2013; pp. 1–10. [CrossRef]
6. Djehaiche, R.; Aidel, S.; Sawalmeh, A.; Saeed, N.; Alenezi, A.H. Adaptive control of IoT/M2M devices in smart buildings using heterogeneous wireless networks. *IEEE Sens. J.* **2023**, *23*, 7836–7849. [CrossRef]
7. Goswami, S.S.; Mondal, S. The role of 5G in enhancing IoT connectivity: A systematic review on applications, challenges, and future prospects. *Big Data Comput. Vis.* **2024**, *4*, 314–326. [CrossRef]

8. Mohseni, M.; Kanwer, B.; Singh, N.; Krishna, M.H.; Rajalakshmi, B.; Alkhafaji, M.A. A comprehensive analysis of the Internet of Things (IoT) in the context of 5G wireless systems. In Proceedings of the 4th International Conference on Innovative Practices in Technology and Management (ICIPTM), Noida, India, 21–23 February 2024. [CrossRef]
9. Kar, S.; Mishra, P.; Wang, K.-C. Efficient resource management using 5G multi-connectivity for high throughput and reliable low latency communication. *EURASIP J. Wirel. Commun. Netw.* **2025**, *2025*, 58. [CrossRef]
10. Mishra, P.; Kar, S.; Bollapragada, V.; Wang, K.-C. Multi-connectivity using NR-DC for high throughput and ultra-reliable low latency communication in 5G networks. In Proceedings of the 2021 IEEE 4th 5G World Forum (5GWF), Montreal, QC, Canada, 13–15 October 2021; pp. 36–40. [CrossRef]
11. Sameer Baddour, L.; Ghazi Abdiwi, F.; Ularbek Duishobekovich, M.; Jawad Abu-AlShaeer, M.; Magbol Alwan, G. Optimizing telecommunications network performance through big data analytics: A comprehensive evaluation. *Iran. J. Inf. Process. Manag.* **2025**, *40*, 1149–1177. [CrossRef]
12. Arnold, P.; von Hugo, D. Future integrated communication network architectures enabling heterogeneous service provision. *Adv. Radio Sci.* **2018**, *16*, 59–66. [CrossRef]
13. Yao, Q.; Liu, Z.; Guo, L. Research on Radio Wave Propagation Characteristics of Urban Environment based on Gaussian-Support Vector Regression Algorithm. In Proceedings of the IEEE 11th Asia-Pacific Conference on Antennas and Propagation (APCAP), Guangzhou, China, 19–22 November 2023. [CrossRef]
14. Sun, S.; Rappaport, T.S.; Shafi, M.; Tang, P.; Zhang, J.; Smith, P.J. Propagation Models and Performance Evaluation for 5G Millimeter-Wave Bands. *IEEE Trans. Veh. Technol.* **2018**, *67*, 8422–8439. [CrossRef]
15. Ethier, J.; Chateauvert, M.; Dempsey, R.G.; Bose, A. Environmental Feature Engineering and Statistical Validation for ML-Based Path Loss Prediction. *IEEE Antennas Wirel. Propag. Lett.* **2026**, *25*, 453–457. [CrossRef]
16. Bakirtzis, S.; Yapar, C.; Fiore, M.; Zhang, J.; Wassell, I. Empowering Wireless Network Applications with Deep Learning-Based Radio Propagation Models. *IEEE Wirel. Commun.* **2025**, *32*, 124–131. [CrossRef]
17. Louie, A.H. Solus mappings. In *Anticipation Science*; Springer International Publishing: Cham, Switzerland, 2017; Volume 2, pp. 45–62. [CrossRef]
18. Widada, W.; Herawati, A.; Fata, R.; Nurhasanah, S.; Yanty, E.P.; Suharno, A.S. Students' understanding of the concept of function and mapping. *J. Phys. Conf. Ser.* **2020**, *1657*, 012072. [CrossRef]
19. Unger, T.I.; Kuczmann, M. Comparison of outdoor radiowave propagation models for land mobile systems in the 3.6 GHz and 6 GHz frequency bands. *Telecom* **2025**, *6*, 42. [CrossRef]
20. Hossain, F.; Geok, T.K.; Aktar, H.; Li, L.L.; Liew, C.P. A comprehensive review: Radio propagation prediction models for 5G and beyond. *Int. J. Commun. Antenna Propag.* **2024**, *14*, 138–153. [CrossRef]
21. Lim, S.Y.; Rafie, I.F.M.; Soo, Q.P.; Yun, Z.; Iskander, M.F. Models of radio propagation in indoor/outdoor environments: Past efforts and future challenges. In *The Advancing World of Applied Electromagnetics: In Honor and Appreciation of Magdy Fahmy Iskander*; Springer: Cham, Switzerland, 2024; pp. 281–311. [CrossRef]
22. Lytaev, M.S.; Vladyko, A.G. Comparative analysis of parabolic equation method and Longley–Rice propagation model. In Proceedings of the International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), Dublin, Ireland, 28–30 October 2019. [CrossRef]
23. Zakaria, Y.; Hosek, J.; Misurec, J. Path loss measurements for wireless communication in urban and rural environments. *Am. J. Eng. Appl. Sci.* **2015**, *8*, 94–99. [CrossRef]
24. Fathy, A.; Newagy, F.; Anis, W.R. Performance evaluation of UWB massive MIMO channels with favorable propagation features. *IEEE Access* **2019**, *7*, 147010–147020. [CrossRef]
25. Hong, J.Y.; Kim, C.-S.; Kim, H.-J.; Lim, J.-S.; Chong, Y.-J. An extended approach of a prediction method for the deterministic propagation path loss model. In Proceedings of the 10th International Conference on ICT Convergence (ICTC), Jeju Island, Republic of Korea, 16–18 October 2019; pp. 1073–1075. [CrossRef]
26. Rahaman, F. Electromagnetic Waves. In *UNITEXT–La Matematica per il 3 + 2*; Springer: Singapore, 2022; Volume 136, pp. 269–309. [CrossRef]
27. Maxwell, J.C. *A Treatise on Electricity and Magnetism*; Cambridge University Press: Cambridge, UK, 2010; pp. 1–442. [CrossRef]
28. Sangster, A.J. Classical Radiation Theory. In *Green Energy and Technology*; Springer: Cham, Switzerland, 2014; Volume 194, pp. 51–72. [CrossRef]
29. Firdaus, R.A.; Khoiro, M.; Asnawi, A.; Bustomi, M.A.; Annovasho, J. Electromagnetic wave equation approximation using FDTD method on conductivity material. *J. Phys. Conf. Ser.* **2021**, *2110*, 012032. [CrossRef]
30. Stolk, C.C. A dispersion minimizing scheme for the 3-D Helmholtz equation based on ray theory. *J. Comput. Phys.* **2016**, *314*, 618–646. [CrossRef]
31. Zhu, J.; Hu, Y. Efficient marching solver for the three-dimensional Helmholtz equation in the cuboid waveguide. *Opt. Contin.* **2025**, *4*, 170–197. [CrossRef]

32. Da Silva, M.A.N.; Costa, E.; Liniger, M. Analysis of the effects from lateral variations of irregular terrain based on a three-dimensional parabolic equation. In Proceedings of the EuCAP 2010–The 4th European Conference on Antennas and Propagation, Barcelona, Spain, 12–16 April 2010.
33. Fathy, A.; Yahya, A.; Ragai, H. Deterministic UWB channel modeling using ray tracing approach. In *Proceedings of ICET 2014–2nd International Conference on Engineering and Technology*; IEEE: New York, NY, USA, 2015. [CrossRef]
34. ITU-R. Recommendation P.1546-6: Method for Point-to-Area Predictions for Terrestrial Services in the Frequency Range 30 MHz to 4000 MHz. Available online: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.1546-6-201908-I!!PDF-E.pdf (accessed on 3 February 2026).
35. ITU-R. Recommendation P.1812-8: A Path-Specific Propagation Prediction Method for Point-to-Area Terrestrial Services in the Frequency Range 30 MHz to 50 GHz. Available online: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.1812-8-202509-I!!PDF-E.pdf (accessed on 3 February 2026).
36. ITU-R. Recommendation P.452-18: Prediction Procedure for the Evaluation of Interference Between Stations on the Surface of the Earth at Frequencies above about 0.1 GHz. Available online: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.452-18-202310-I!!PDF-E.pdf (accessed on 3 February 2026).
37. Zhang, H.; Li, Q.; Guo, X.; Hou, C. Research on radio wave propagation prediction model using 3-D parabolic equation over rough sea surface. In Proceedings of the 2021 International Applied Computational Electromagnetics Society Symposium (ACES-China 2021), Chengdu, China, 28–31 July 2021. [CrossRef]
38. Zhang, R.; Lu, G.; Abomakhleb, G. Finite-difference solution of the parabolic equation under horizontal polar coordinates. *IEEE Antennas Wirel. Propag. Lett.* **2017**, *16*, 2931–2934. [CrossRef]
39. Zhang, X.; Xiong, X.; Liao, C.; Deng, X. Hybrid algorithm of radio wave propagation based on parabolic equation in cylindrical coordinates and method of moments. *High Power Laser Part. Beams* **2020**, *32*, 053004. [CrossRef]
40. Shaw, J.A. Radiometry and the Friis transmission equation. *Am. J. Phys.* **2012**, *81*, 33–37. [CrossRef]
41. Halim, M.A.; Razak, M.S.; Yasin, M.N.M.; Khairunizam, W.; Fareq, M.; Fitra, M. Optimization on wireless power transfer. *J. Teknol.* **2015**, *77*, 17–21. [CrossRef]
42. Unger, T.I.; Kuczmann, M. The impact of terrain sampling density on 5G NR-V2X downlink channel modeling using various propagation models at the 3.6 GHz band. *Radioengineering* **2025**, *34*, 603–623. [CrossRef]
43. Iliev, I.; Velchev, Y.; Petkov, P.Z.; Bonev, B.; Iliev, G.; Nachev, I. A machine learning approach for path loss prediction using combination of regression and classification models. *Sensors* **2024**, *24*, 5855. [CrossRef]
44. López-Ramírez, G.A.; Aragón-Zavala, A. Enhancing indoor mmWave communication with ML-based propagation models. *IEEE Access* **2025**, *13*, 13748–13769. [CrossRef]
45. Bakirtzis, S.; Qiu, K.; Chen, J.; Song, H.; Zhang, J.; Wassell, I. Rigorous indoor wireless communication system simulations with deep learning-based radio propagation models. *IEEE J. Multiscale Multiphys. Comput. Tech.* **2025**, *10*, 58–68. [CrossRef]
46. Bocus, M.Z.; Lodhi, A. Application of machine learning for radiowave propagation modeling below 6 GHz. *IEEE Access* **2025**, *13*, 9755–9765. [CrossRef]
47. Ahmad, K.; Hussain, S. Machine learning approaches for radio propagation modeling in urban vehicular channels. *IEEE Access* **2022**, *10*, 113690–113698. [CrossRef]
48. Kuno, N.; Takatori, Y. Prediction method by deep-learning for path loss characteristics in an open-square environment. In Proceedings of the ISAP 2018 International Symposium on Antennas and Propagation, Busan, Republic of Korea, 23–26 October 2018.
49. Yang, M.; Stavrou, S.; Brown, A.K. Hybrid ray-tracing model for radio wave propagation through periodic building structures. *IET Microw. Antennas Propag.* **2011**, *5*, 340–348. [CrossRef]
50. El Ahdab, Z.; Akleman, F. An efficient 3-D FDTD-PE hybrid model for radio wave propagation with near-source obstacles. *IEEE Trans. Antennas Propag.* **2019**, *67*, 346–355. [CrossRef]
51. Wang, Y.; Tu, L.; Guo, J.; Wang, Z. Residual learning based RF signal denoising. In Proceedings of the 4th IEEE International Conference on Applied System Innovation (ICASI 2018), Chiba, Japan, 13–17 April 2018; pp. 15–18. [CrossRef]
52. Mehtaj, N.; Banerjee, S. Scientific machine learning for elastic and acoustic wave propagation: Neural operator and physics-guided neural network. *Sensors* **2025**, *25*, 3588. [CrossRef] [PubMed]
53. Wang, S.; Huang, S.; Zhang, X. Radio wave propagation modeling based on convolutional neural networks. In Proceedings of the International Conference on Communication Technology (ICCT 2022), Nanjing, China, 11–14 November 2022; pp. 255–259. [CrossRef]
54. Moghadam, H.H.; Foudazi, A.; Kouki, A.B. A comparison of satellite signal simulation in street canyon in 2D and 3D deterministic methods. In *Proceedings of the IEEE Antennas and Propagation Society International Symposium*; IEEE: New York, NY, USA, 2017; pp. 599–600. [CrossRef]
55. Buddendick, H.; Eibert, T.; Hasch, J. Bistatic scattering center models for the simulation of wave propagation in automotive radar systems. In *Proceedings of the German Microwave Conference (GeMIC 2010)*, Berlin, Germany, 15–17 March 2010; pp. 288–291.

56. Huang, S.; Qin, H.; Zhang, X. Efficient parabolic equation-driven CNN propagation model in tunnels based on frequency conversion. *IEEE Trans. Antennas Propag.* **2024**, *72*, 6024–6031. [CrossRef]
57. Barrios-Ulloa, A.; Cama-Pinto, A.; De-la-Hoz-Franco, E.; Ramírez-Velarde, R.; Cama-Pinto, D. Modeling of path loss for radio wave propagation in wireless sensor networks in cassava crops using machine learning. *Agriculture* **2023**, *13*, 2046. [CrossRef]
58. Boga, K.M.; Pichhika, H.C.; Gokanakonda, J.K.; Kandimalla, D.; Yerra, R.V.P. Predictive path loss modeling for wireless sensor networks in structured mango orchards: Using machine learning models. In *Proceedings of the IEEE 21st India Council International Conference (INDICON 2024)*; IEEE: New York, NY, USA, 2024. [CrossRef]
59. Seretis, A.; Sarris, C.D. An overview of machine learning techniques for radiowave propagation modeling. *IEEE Trans. Antennas Propag.* **2022**, *70*, 3970–3985. [CrossRef]
60. European Commission. Commission Implementing Decision 2014/276/EU of 2 May 2014 Amending Decision 2008/411/EC on the Harmonisation of the 3400–3800 MHz Frequency Band for Terrestrial Systems Capable of Providing Electronic Communications Services in the Community. Available online: https://eur-lex.europa.eu/eli/dec_impl/2014/276/oj (accessed on 21 July 2025).
61. Electronic Communications Committee (ECC). ECC Decision (11)06 of 9 December 2011 on Harmonised Frequency Arrangements for Mobile/Fixed Communications Networks (MFCN) Operating in the Band 3400–3800 MHz. Available online: https://docdb.cept.org/document/433 (accessed on 5 February 2026).
62. Radiocommunication Sector of the International Telecommunication Union (ITU-R). Radio Regulations. 2024. Available online: http://handle.itu.int/11.1002/pub/8229633e-en (accessed on 5 February 2026).
63. Electronic Communications Committee (ECC) Within the European Conference of Postal and Telecommunications Administrations (CEPT). The European Table of Frequency Allocations in the Frequency Range 8.3 kHz to 3000 GHz (ECA table), Approved January 2025. Available online: https://efis.cept.org/sitecontent.jsp?sitecontent=ecatable (accessed on 6 February 2026).
64. European Space Agency. Copernicus Global Digital Elevation Model [Dataset]. Distributed by OpenTopography. 2024. Available online: https://portal.opentopography.org/datasetMetadata?otCollectionID=OT.032021.4326.1 (accessed on 5 February 2026).
65. Ozgun, O.; Sahin, V.; Erguden, M.E.; Apaydin, G.; Yilmaz, A.E.; Kuzuoglu, M.; Sevgi, L. PETOOL v2.0: Parabolic equation toolbox with evaporation duct models and real environment data. *Comput. Phys. Commun.* **2020**, *256*, 107454. [CrossRef]
66. Huang, Y.; Hou, Y.-F.; Dral, P.O. Active delta-learning for fast construction of interatomic potentials and stable molecular dynamics simulations. *Mach. Learn. Sci. Technol.* **2025**, *6*, 035004. [CrossRef]
67. Huo, J.; Dong, H. Δ-EGNN method accelerates the construction of machine learning potential. *J. Phys. Chem. Lett.* **2025**, *16*, 2080–2088. [CrossRef] [PubMed]
68. Backurs, A.; Indyk, P.; Schmidt, L. On the fine-grained complexity of empirical risk minimization: Kernel methods and neural networks. In *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*; Curran Associates, Inc.: Red Hook, NY, USA, 2017; pp. 4309–4319.
69. Steinwart, I. Measuring the capacity of sets of functions in the analysis of ERM. In *Measures of Complexity: Festschrift for Alexey Chervonenkis*; Springer: Cham, Switzerland, 2015; pp. 217–233. [CrossRef]
70. Menden, V.; Saleh, Y.; Iske, A. Bounds on the generalization error in active learning. In *Proceedings of the 6th Northern Lights Deep Learning Conference*; PMLR: Tromsø, Norway, 2025; Volume 265, pp. 168–175.
71. Venkatesh, K.A.; Mishra, D.; Manimozhi, T. Model selection and regularization. In *Statistical Modeling in Machine Learning: Concepts and Applications*; Academic Press: London, UK, 2022; pp. 159–178. [CrossRef]
72. McDonald, G.C. Ridge regression. *WIREs Comput. Stat.* **2009**, *1*, 93–100. [CrossRef]
73. Xie, S.; Sun, X.; Zhang, S.; Xiong, Z.; Wei, X.; Cui, C. Precipitation forecast correction in South China based on SVD and machine learning. *Yingyong Qixiang Xuebao* **2022**, *33*, 293–304. (in Chinese) [CrossRef]
74. Huang, F.; Li, D.; Xu, J.; Wu, Y.; Xing, Y.; Yang, Z. Ridge regression based on gradient descent method with memory dependent derivative. In Proceedings of the IEEE International Conference on Software Engineering and Service Sciences (ICSESS), Beijing, China, 16–18 October 2020; pp. 463–467. [CrossRef]
75. Shirzadi, M.; Dehghan, M. Generalized regularized least-squares approximation of noisy data with application to stochastic PDEs. *Appl. Math. Lett.* **2021**, *111*, 106598. [CrossRef]
76. Li, D.; Ge, Q.; Zhang, P.; Xing, Y.; Yang, Z.; Nai, W. Ridge regression with high order truncated gradient descent method. In *Proceedings of the 2020 12th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Hangzhou, China, 22–23 August 2020;* IEEE: New York, NY, USA, 2020; Volume 1, pp. 252–255. [CrossRef]
77. Chen, S.; Liu, Y.; Lyu, M.R.; King, I.; Zhang, S. Fast relative-error approximation algorithm for ridge regression. In *Uncertainty in Artificial Intelligence–Proceedings of the 31st Conference (UAI 2015)*; AUAI Press: Arlington, VA, USA, 2015; pp. 201–210.
78. McDonald, G.C. Tracing ridge regression coefficients. *WIREs Comput. Stat.* **2010**, *2*, 695–703. [CrossRef]
79. Hodson, T.O. Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geosci. Model Dev.* **2022**, *15*, 5481–5487. [CrossRef]

80. Robeson, S.M.; Willmott, C.J. Decomposition of the mean absolute error (MAE) into systematic and unsystematic components. *PLoS ONE* **2023**, *18*, e0279774. [CrossRef]
81. Chai, T. Root mean square. In *Encyclopedia of Earth Sciences Series*; Springer: Cham, Switzerland, 2020. [CrossRef]
82. Tanaka, A.; Imai, H.; Kudo, M.; Miyakoshi, M. Mathematical interpretations of kernel ridge regression. In *AIP Conference Proceedings*; AIP Publishing LLC: Melville, NY, USA, 2006; Volume 839, pp. 347–353. [CrossRef]
83. Tuo, R.; Wang, Y.; Wu, C.F.J. On the improved rates of convergence for matern-type kernel ridge regression with application to calibration of computer models. *SIAM-ASA J. Uncertain. Quantif.* **2021**, *8*, 1522–1547. [CrossRef]
84. Exterkate, P. Model selection in kernel ridge regression. *Comput. Stat. Data Anal.* **2013**, *68*, 1–16. [CrossRef]
85. Singh, H. Machine Learning Application of Generalized Gaussian Radial Basis Function and Its Reproducing Kernel Theory. *Mathematics* **2024**, *12*, 829. [CrossRef]
86. Yang, Y.; Pilanci, M.; Wainwright, M.J. Randomized sketches for kernels: Fast and optimal nonparametric regression. *Ann. Stat.* **2017**, *45*, 991–1023. [CrossRef]
87. Matsubara, T. Wasserstein Gradient Boosting: A Framework for Distribution-Valued Supervised Learning. In *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*; Curran Associates Inc.: Red Hook, NY, USA, 2024; Volume 37.
88. Guelman, L. Gradient boosting trees for auto insurance loss cost modeling and prediction. *Expert Syst. Appl.* **2012**, *39*, 3659–3667. [CrossRef]
89. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
90. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: New York, NY, USA, 2009. [CrossRef]
91. Yildiz, C.; Guney, K.; Turkmen, M.; Kaya, S. Neural models for quasi-static analysis of conventional and supported coplanar waveguides. *AEU Int. J. Electron. Commun.* **2007**, *61*, 521–527. [CrossRef]
92. Alsmadi, M.K.; Omar, K.B.; Noah, S.A.; Almarashdah, I. Performance comparison of multi-layer perceptron (Back Propagation, Delta Rule and Perceptron) algorithms in neural networks. In Proceedings of the 2009 IEEE International Advance Computing Conference (IACC), Patiala, India, 6–7 March 2009; pp. 296–299. [CrossRef]
93. Haykin, S. *Neural Networks and Learning Machines*, 3rd ed.; Pearson Education: Upper Saddle River, NJ, USA, 2009.
94. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006.
95. Møller, M.F. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Netw.* **1993**, *6*, 525–533. [CrossRef]
96. Kohonen, T. *Self-Organizing Maps*, 3rd ed.; Springer: Berlin, Germany, 2001.