

Article

A Data-Driven Gaussian Process Regression Model for Concrete Complex Dielectric Permittivity Characterization

Giovanni Angiulli ¹, Mario Versaci ², Pietro Burrascano ³ and Filippo Laganá ^{4,*}

¹ Department of Information Engineering, Infrastructures and Sustainable Energy, Mediterranea University, 89121 Reggio Calabria, Italy; giovanni.angiulli@unirc.it

² Department of Civil, Energetic, Environmental and Material Engineering, Mediterranea University, 89121 Reggio Calabria, Italy; mario.versaci@unirc.it

³ Dipartimento di Ingegneria—Università di Perugia, 05100 Terni, Italy; pietro.burrascano@unipg.it

⁴ Laboratory of Biomedical Applications Technologies and Sensors (BATS), Department of Health Science, Magna Græcia University, 88100 Catanzaro, Italy

* Correspondence: filippo.lagana@unicz.it

Abstract

Concrete diagnosis is an important task in making informed decisions about reconstructing or repairing buildings. Among the different approaches for evaluating its characteristics, methods based on electromagnetic waves have been proposed in the literature over the years. In this context, the characterization of concrete complex dielectric permittivity $\epsilon_r(f)$ (where f is the frequency) has received considerable attention, taking into account that its values and its frequency behavior are both sensitive to a series of physical parameters, which in turn can significantly influence the mechanical performance of concrete. Recently, data-driven techniques have emerged as alternatives for modeling material properties due to their regression and generalization potential. Following this research line in this work, we investigated the potential of Gaussian Process Regression to model $\epsilon_r(f)$ by comparing its performance with that of the model most employed to characterize the concrete dielectric permittivity: the universal Jonscher model. The inherent ability to provide predictions accompanied by confidence intervals, which allows the assessment of the reliability of the permittivity estimate across frequency, and the related error metrics demonstrate that GPR can effectively characterize $\epsilon_r(f)$ in an effective manner, outperforming the Jonscher model in terms of accuracy in all the cases considered in our study.

Keywords: concrete; permittivity; Gaussian processes; data-driven modeling



Academic Editor: Mohammad Noori

Received: 30 August 2025

Revised: 25 September 2025

Accepted: 13 October 2025

Published: 14 October 2025

Citation: Angiulli, G.; Versaci, M.;

Burrascano, P.; Laganá, F. A

Data-Driven Gaussian Process Regression Model for Concrete Complex Dielectric Permittivity Characterization. *Sensors* **2025**, *25*, 6350. <https://doi.org/10.3390/s25206350>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Due to environmental factors and dynamic loads, building structures are vulnerable to damage [1,2]. In ordinary cement concrete, cracks are inevitable due to several natural or manufactured factors, resulting in a decline in compressive strength performance and leading to structural deformation, which affects the structural integrity [2]. Accordingly, concrete diagnosis results are of primary importance for practitioners in the field in making informed decisions about reconstructing or repairing buildings [3]. During the decades, monitoring concrete mechanical properties has become an area of cross-disciplinary research, and numerous non-destructive testing evaluation (NDT) methods have been developed for this aim [1,3,4].

Alongside techniques using elastic waves, methods based on electromagnetic waves have gained popularity over the years [5]. In fact, electromagnetic techniques offer several

advantages over other methods, such as the possibility of real-time monitoring of concrete building structures using contactless and wireless techniques [6–8], as well as the ability to achieve in situ measurements [9]. The increasing reliance on the electromagnetic approach for monitoring the properties of concrete is shaped by numerous studies published in the literature that have demonstrated the existence of relationships between electromagnetic field-related quantities and concrete properties, such as water content, porosity, and the presence of chloride [10]. Taking into account the scenario discussed so far, it becomes evident how the characterization of the concrete dielectric permittivity turns out to be essential in modeling its electromagnetic behavior as well as in retrieving its mechanical properties, such as compressive strength, which is the most important parameter to describe concrete mechanical performance [11]. To this end, researchers in the field of non-destructive testing (NDT) have placed particular emphasis on developing sensors and measurement techniques that accurately reconstruct complex permittivity through the precise measurement of concrete scattering parameters [12–15].

In the literature, various dielectric models have been employed over the years to characterize the permittivity of concrete, including the Debye model, the complex refractive index method (CRIM), the first-order exponential model, and the Cole–Cole model [16,17]. Among these, the Jonscher universal dielectric permittivity model is the one that finds the most favor among scholars and professionals in the field [1,18,19]. Recently, it was employed to evaluate the concrete shielding effectiveness, yielding notable results compared to other models, and to characterize the dielectric permittivity of Chinese standard concrete, the most widely produced type of concrete in the world [20]. Although the above models are well-established, we note that data-driven techniques have recently gained traction as alternatives for building models for property prediction and material design, owing to their ability to regress and generalize [21]. Researchers have achieved a series of milestones in developing data-driven methods that can predict electromagnetic properties with high accuracy, yielding results in excellent agreement with experimental measurements. In particular, Gaussian Process Regression (GPR) has been extensively applied to this end [22,23]. Gaussian process regression (GPR) is a powerful non-parametric supervised learning method widely used for regression and probabilistic classification tasks that, unlike the classical parametric models, defines a distribution over functions and derives predictions directly from the observed data [24]. This property makes Gaussian process regression suitable in all those situations where the underlying relationship between input and output variables is not easily expressed through analytical models, as is often the case in characterizing material properties [22,23]. A key advantage of GPR is that it provides not only point estimates, but also confidence intervals around the predictions, which are derived directly from the posterior variance of the Gaussian process and thus quantify the uncertainty associated with the prediction, a feature that distinguishes GPR from deterministic regression methods, by offering a means of assessing the reliability of the estimated function and providing a measure of confidence in the predictions, which is essential for non-destructive testing applications, where measurement errors and variability in material samples are inevitable. Based on the above, and taking into account that, to the best of our knowledge, no study has been devoted to employing GPR to characterize the complex permittivity of concrete until now, in this work, we investigated the GPR ability to describe the concrete complex dielectric permittivity $\epsilon_r(\omega)$ by comparing its performance with that of the universal Jonscher model to evaluate its capability to overcome the Jonscher model limitations [25]. The paper is organized as follows: Section 2 illustrates the materials and the methods employed in our analysis; in particular, Section 2.1 contains the basics on Jonscher’s universal model, while Sections 2.2 and 2.3 discuss the foundation of Gaussian process regression and the selection criteria of its hyperparameters. Section 2.4 briefly

introduces the error metrics adopted in our analysis. In Section 3, the implementation of the data-driven GPR model used to characterize the concrete complex permittivity and the related numerical performance on experimental data compared with those offered by the Jonscher model is discussed. Finally, in Section 4, we draw together some considerations and future perspectives connected with the analysis we conducted.

2. Materials and Methods

2.1. The Jonscher Universal Model

The rationale behind Jonscher's universal dielectric model is based on empirical observations rooted in the identification of commonalities in frequency-dependent dielectric behavior conducted over a wide variety of materials differing in chemical composition and structural characteristics as well as over different frequency ranges [26]. Accordingly, the Jonscher model provides a basis for describing the frequency-dependent behavior of the complex dielectric permittivity, $\epsilon(\omega)$, in matter with remarkable consistency. More precisely, the Jonscher model describes the dielectric behavior of materials by modeling relaxation processes that occur across different time scales with a power-law model, replacing the distribution of relaxation times used in Debye and Cole–Cole models with the concept of a frequency-independent ratio of energy lost to energy stored per cycle. It can be formally introduced starting from the relationship between the dielectric displacement field \mathbf{D} and the electric field \mathbf{E} in an isotropic solid material [26], written as follows:

$$\mathbf{D} = \epsilon(\omega)\mathbf{E} = (\epsilon_0\chi(\omega) + \epsilon_\infty)\mathbf{E} \quad (1)$$

where ϵ_0 is the permittivity of the free space, ϵ_∞ is the $\epsilon(\omega)$ high-frequency limit, $\chi(\omega)$ is the complex electric susceptibility, and $\omega = 2\pi f$ is the angular frequency. As said above, Jonscher derived his key insight from the experimental observations that an extensive range of solid dielectrics have a universal behavior in their susceptibility described by [26]:

$$\chi(\omega) \propto \omega^{n-1} \quad (2)$$

where the exponent $n \in [0, 1]$ controls the material frequency dispersion behavior. The result (2) in conjunction with the causality principle, led Jonscher to conclude that the ratio between $\text{Re}[\chi(\omega)]$, and $\text{Im}[\chi(\omega)]$ must result independent from frequency and equal to the following:

$$\frac{\text{Im}[\chi(\omega)]}{\text{Re}[\chi(\omega)]} = \cot\left(\frac{n\pi}{2}\right). \quad (3)$$

Selecting for the real part the following functional form [26], we obtain the following:

$$\text{Re}[\chi(\omega)] = \chi_r \left(\frac{\omega}{\omega_t}\right)^{n-1} \quad (4)$$

with $\omega_t = 2\pi f_t$, where f_t is the so called reference frequency [26] (which can be arbitrarily chosen within the considered frequency band [25,27]), Jonscher obtained the following expression for dielectric permittivity $\epsilon_r(f)$ [26]:

$$\epsilon_r(f) = \chi_r \left(\frac{f}{f_t}\right)^{n-1} \left[1 - i \cot\left(\frac{n\pi}{2}\right)\right] + \bar{\epsilon}_\infty \quad (5)$$

where χ_r is the value of the real part of $\chi(2\pi f)$ for $f = f_t$ and $\bar{\epsilon}_\infty = \epsilon_\infty/\epsilon_0$.

2.2. Gaussian Process Regression

The aim of the Gaussian Process Regression (GPR) is to exploit a set $\{\mathbf{x}_i, y_i\}_{i=1}^n$, where $\mathbf{x}_i \in \mathbb{R}^d$ are the data sites and $y_i \in \mathbb{R}$ are the data values, for obtaining an estimator that can predict the value of the function $y_\beta = y(\mathbf{x}_\beta)$ in correspondence to \mathbf{x}_β , $\mathbf{x}_\beta \notin \{\mathbf{x}_i\}_{i=1}^n$, quantifying the expected error [28]. The fundamental assumption underlying the GPR approach is that the data values y_i are the realizations of a particular random field, $\tilde{\mathcal{Y}}$, named Gaussian random process, thus defined as follows [24,28]:

Definition 1. A set of random variables is calculated as follows:

$$\tilde{\mathcal{Y}} = \{Y(\mathbf{x}, \xi), \mathbf{x} \in \Omega \text{ and } \xi \in \mathcal{W}\} \quad (6)$$

where \mathcal{W} is a probability space, and Ω is a set called parameter space, is named as random field. If, for any arbitrarily choice of the set $\{\mathbf{x}_i\}_{i=1}^n$, the vector is calculated as follows:

$$\mathbf{Y} = \begin{bmatrix} Y_{\mathbf{x}_1}(\xi) \\ \vdots \\ Y_{\mathbf{x}_n}(\xi) \end{bmatrix}$$

is characterized by a multivariate Gaussian distribution, with mean

$$\boldsymbol{\mu} = 0 \quad (7)$$

and covariance

$$\mathbf{Cov}[Y_{\mathbf{x}}(\xi), Y_{\mathbf{x}'}(\xi)] = \sigma^2 \mathbf{K} \quad (8)$$

the random field $\tilde{\mathcal{Y}}$ is named Gaussian random process, where σ^2 is the process variance and \mathbf{K} , defined as follows:

$$\mathbf{K} = \left[\exp(-\varepsilon^2 \|\mathbf{x}_i - \mathbf{x}_j\|^2) \right]_{i,j \in \{1, \dots, n\}} \quad (9)$$

is the covariance kernel matrix, and $\varepsilon^2 = 1/(2l^2)$, is a kernel-specific parameter inversely related to the square of the characteristic length-scale l , respectively.

The random field $\tilde{\mathcal{Y}}$ can be used for regression as follows; let $\tilde{Y}_{\mathbf{x}}$, written as follows:

$$\tilde{Y}_{\mathbf{x}} = \mathbf{w}(\mathbf{x})^t \cdot \tilde{\mathcal{Y}} \quad (10)$$

a linear predictor belonging to $\tilde{\mathcal{Y}}$, where $\mathbf{w}(\mathbf{x})^t$ are the unknowns weight functions values $w_i(\cdot)$ assumed in \mathbf{x} , which must be determined through the minimization of the mean-squared error of $\tilde{Y}_{\mathbf{x}}$ defined as follows:

$$\epsilon_{\tilde{Y}_{\mathbf{x}}} = \mathbb{E} \left[\left(Y_{\mathbf{x}} - \mathbf{w}(\mathbf{x})^t \cdot \tilde{\mathcal{Y}} \right)^2 \right] \quad (11)$$

Because of the stationarity of $\tilde{\mathcal{Y}}$, it results that $\sigma^2 K(\mathbf{x}, \mathbf{x}') = \mathbb{E}[\tilde{Y}_{\mathbf{x}} \tilde{Y}_{\mathbf{x}'}]$. Using this last relation in (11), we obtain the following:

$$\begin{aligned} \epsilon_{\tilde{Y}_{\mathbf{x}}} &= \sigma^2 K(\mathbf{x}, \mathbf{x}') - 2\mathbf{w}(\mathbf{x})^t (\sigma^2 \mathbf{k}(\mathbf{x})) + \\ &\quad + \mathbf{w}(\mathbf{x})^t \cdot \sigma^2 \mathbf{K} \cdot \mathbf{w}(\mathbf{x}) \end{aligned} \quad (12)$$

where $\mathbf{k}(\mathbf{x}) = [K(\mathbf{x}, \mathbf{x}_1), \dots, K(\mathbf{x}, \mathbf{x}_N)]^T$. To find the optimal weight vector $\mathbf{w}(\mathbf{x})$, we take the derivative of (12), with respect to $\mathbf{w}(\mathbf{x})$, and set it to zero:

$$\frac{\partial \epsilon_{\bar{Y}_x}}{\partial \mathbf{w}(\mathbf{x})} = \sigma^2 (-2\mathbf{k}(\mathbf{x}) + 2\mathbf{K}\mathbf{w}(\mathbf{x})) = 0 \quad (13)$$

Finally, solving for $\mathbf{w}(\mathbf{x})$, we obtain the following stationary point for $\epsilon_{\bar{Y}_x}$:

$$\mathbf{w}(\mathbf{x}) = \mathbf{K}^{-1} \cdot \mathbf{k}(\mathbf{x}) \quad (14)$$

Using (14) in (10), and fixing the value of ϵ^2 , the following realization of the predictor \bar{Y}_x is obtained:

$$\bar{y}_x = \mathbf{k}(\mathbf{x})^t \cdot \mathbf{K}^{-1} \cdot \bar{\mathbf{y}} \quad (15)$$

which provides the GPR regression value in correspondence of \mathbf{x} [24].

2.3. Hyperparameters Selection

The generalization capability and the predictive accuracy of the GPR model are contingent upon the appropriate selection of the process variance σ^2 , the kernel-specific parameter ϵ^2 , which governs the covariance kernel matrix \mathbf{K} defined by (9), and the noise variance σ_n^2 , which accounts for the measurement uncertainty in the observed data. To optimize these terms, named hyperparameters, the marginal likelihood of the observed data $\{\mathbf{x}_i, y_i\}_{i=1}^n$ must be maximized. This approach leverages the nature of the random process, where the log marginal likelihood function for the zero-mean field \vec{Y} is expressed as follows:

$$\log p(\mathbf{y} | \mathbf{X}, \epsilon^2, \sigma^2, \sigma_n^2) = -\frac{1}{2} \mathbf{y}^T (\sigma^2 \mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y} - \frac{1}{2} \log |\sigma^2 \mathbf{K} + \sigma_n^2 \mathbf{I}| - \frac{N}{2} \log(2\pi) \quad (16)$$

where $\mathbf{y} = [y_1, \dots, y_n]^t$ is the vector of observed values, \mathbf{X} is the design matrix of data sites, \mathbf{I} is the identity matrix, and \mathbf{K} depends on ϵ^2 via (9). Gradient-based optimization methods, such as conjugate gradients or quasi-Newton algorithms, are typically employed to maximize the log likelihood (16) with respect to σ^2 , ϵ^2 , and σ_n^2 . The optimized values ϵ_{opt}^2 , σ_{opt}^2 , and $\sigma_{n,opt}^2$ are subsequently substituted into the predictor defined by Equation (15) to minimize regression uncertainty and enhance out-of-sample predictions [24]. Furthermore, this optimization inherently mitigates overfitting by striking a balance between model complexity and data fidelity [24].

2.4. Error Metrics

To compare the performance of the Jonscher model and the GPR model with each other, the following metrics were employed: The mean-squared error (RMSE), calculated as follows:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\epsilon_{ri}^{mod} - \epsilon_{ri}^{meas})^2 \quad (17)$$

The mean absolute percentage error (MAPE), calculated as follows:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\epsilon_{ri}^{mod} - \epsilon_{ri}^{meas}}{\epsilon_{ri}^{meas}} \right| \% \quad (18)$$

The mean absolute error (MAE), calculated as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n \left| \epsilon_{ri}^{mod} - \epsilon_{ri}^{meas} \right| \quad (19)$$

where n is the number of available measured concrete complex permittivity samples, $\epsilon_{r_i}^{mod}$ and $\epsilon_{r_i}^{meas}$, $i \in [1, \dots, n]$, are the predicted and the measured permittivity value at the frequency f_i , respectively.

3. Numerical Results

To implement the data-driven GPR model the built-in functions `fitrgp` and `predict` of MATLAB 2023b © STATISTICS AND MACHINE LEARNING TOOLBOX and the related options have been used [29]. Algorithm 1 reports the MATLAB pseudocode of the implemented procedure. All the datasets considered in our study are of the following form:

$$\mathcal{D} = \{\mathbf{x}_i = f_i, y_i = \epsilon_i^{meas}\}_{i=1}^n$$

where the i -th tuple in \mathcal{D} corresponds to the permittivity measurement ϵ_i^{meas} at the discrete frequency f_i ($\mathbf{x}_i = f_i, \mathbb{R}^d = \mathbb{R}$) for a specific concrete specimen. The Jonscher parametrization for each of the cases considered in this study is reported in Table 1. The kernel function we used was specified as the squared exponential (Gaussian) kernel using the options ('KernelFunction', 'squarexponential') included in the `fitrgp` function [29]. Furthermore, we used the option ('Standardize, true') available in the options of the same function to standardize the input predictors, i.e., in this case the frequency values, by removing the mean and scaling to unit variance before training, thus improving kernel resizing and optimization stability without altering the physical units of frequency or the predicted permittivity [29]. Regarding the setting of the hyperparameters ϵ , σ^2 , and σ_n^2 we optimized them using the predefined options ('Fitmethod', 'exact'), and ('Optimizer', 'quasinewton') included in the `fitrgp` function [29]. The 'exact' fit option was used due to the small size of the data sets considered, allowing for a precise calculation of the covariance matrix, while the 'quasinewton' optimizer was selected for its efficiency in solving the relative unconstrained optimization problem. The involved CPU time for all computations running on a machine based on Intel(R) Core(TM) i5-4570S CPU at 2.90 GHz was a few fractions of a second (Intel Corporation, Santa Clara, CA, USA). All the test and training datasets employed to build the GPR models discussed in the following were created by randomly partitioning the dataset \mathcal{D} across frequencies.

The first case we considered involved two datasets, each composed of $n = 16$ tuples from the experimental data reported in [27]. The first set of data was relevant to concrete characterized by a water-to-cement ratio, W/C , of 0.30, average porosity, AP , of 8.7% and bulk density, BD , of 2.27 ($\text{g}\cdot\text{cm}^{-3}$) labeled as B1, while the second one was relevant to concrete with a W/C of 0.66, AP equal to 13.7%, and BD of 2.18, labeled as B2. The permittivity data were collected in the frequency range $[0, 1]$ GHz [27], see Algorithm 1.

Algorithm 1 MATLAB GPR concrete permittivity modeling pseudocode.

- 1: INPUT: $Data_concrete, dim_training_set$;
 - 2: OUTPUT: $\bar{y}_{x_{Re}}, \bar{y}_{x_{Im}},$
 $MSE_{Re}, MAE_{Re}, MAPE_{Re},$
 $MSE_{Im}, MAE_{Im}, MAPE_{Im}$;
 - 3: **function** $[DTrain_f, DTrain_{\epsilon_{rRe}}, DTrain_{\epsilon_{rIm}}, DTest_f, DTest_{\epsilon_{rRe}}, DTest_{\epsilon_{rIm}}, idx]=$
 $PartitionDATA(Data_concrete, dim_training_set)$;
 - 4: $GPR_Re = fitrgp(DTrain_f, DTrain_{\epsilon_{rRe}})$;
 - 5: $GPR_Im = fitrgp(DTrain_f, DTrain_{\epsilon_{rIm}})$;
 - 6: $\bar{y}_{x_{Re}} = predict(GPR_Re, DTest_f)$;
 - 7: $\bar{y}_{x_{Im}} = predict(GPR_Im, DTest_f)$;
 - 8: **function** $[MSE_{Re}, MAE_{Re}, MAPE_{Re}] = ErrorMETRICS(\bar{y}_{x_{Re}}, DTest_{\epsilon_{rRe}})$;
 - 9: **function** $[MSE_{Re}, MAE_{Re}, MAPE_{Re}] = ErrorMETRICS(\bar{y}_{x_{Im}}, DTest_{\epsilon_{rIm}})$;
 - 10: **end**
-

Table 1. Jonscher parameters for the types of concrete considered in this work (the f_t values are also reported).

Concrete	n	χ_r	$\bar{\epsilon}_\infty$	f_t (GHz)
B1	0.777	3.82	3.56	0.2266
B2	0.683	2.10	3.71	0.2266
LFC	0.901	1.55	1.0	0.2129
ULCC	0.712	3.73	5.0	0.0321

We built up two GPR models: one for characterizing the real part and the other for the imaginary part of the complex permittivity of the concrete at hand. For both considered datasets, the set used to train the GPR model for model fitting and hyperparameter optimization consisted of 70% of the tuples belonging to \mathcal{D} . In contrast, the pertinent test set consisted of the remaining 30%. Figures 1 and 2 report a comparison among the values of the real and the imaginary parts of ϵ^{meas} provided by the Jonscher and the GPR model. The figures also display the confidence intervals, also known as epsilon tubes, for both GPR models (represented by the dashed red and blue lines, respectively). These intervals represent the uncertainty bounds around the GPR predictions, enabling the assessment of the predictive performance and generalization capabilities of both GPR models.

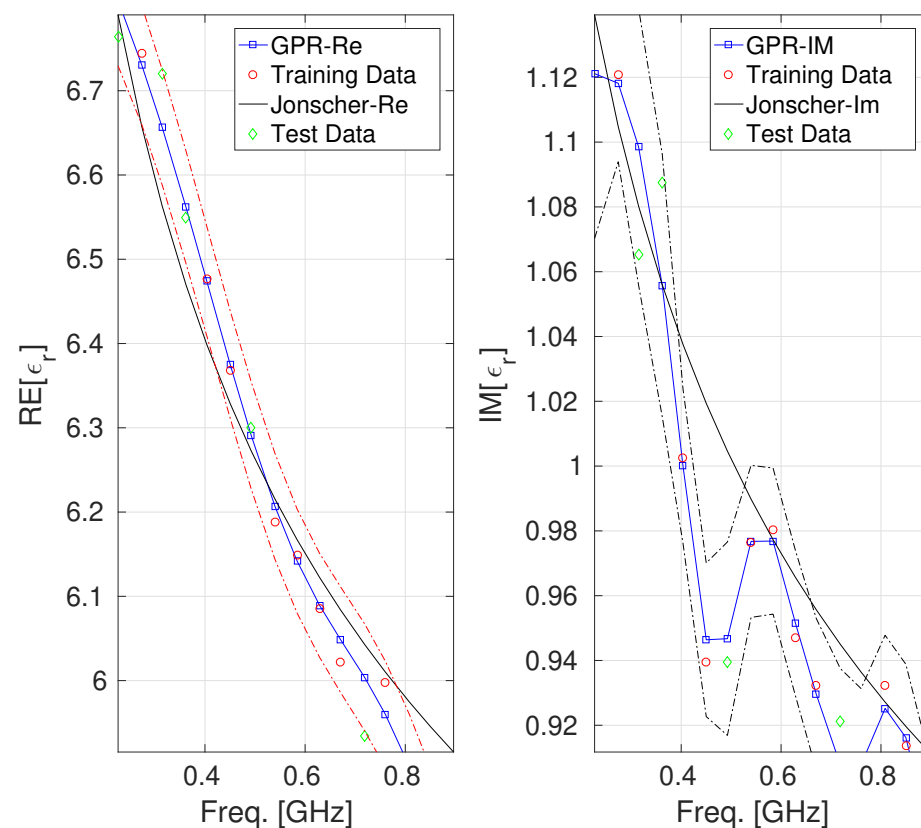


Figure 1. Behavior of the real and the imaginary parts of ϵ_r for the concrete labeled as B1 [27]: GPR model vs. Jonscher parametrization. The uncertainty bands (epsilon-tube) are also reported (identified by the red and black dashed lines).

Regarding the model characterizing the real part of the dielectric permittivity of concrete B1, it can be observed from the left side of Figure 1 that the confidence interval results are symmetric and not too wide, indicating a good characterization of the prediction. Furthermore, the test data all fall within the confidence interval, indicating that the model is faithful and capable of good generalization. This behavior suggests a well-calibrated

GPR model for the real part of the concrete permittivity with limited overfitting and a reliable quantification of uncertainty. About the GPR model characterizing the imaginary part of the dielectric permittivity of the concrete B1 (reported on the right side of Figure 1, we can observe that the confidence interval results in a reasonable width, not too narrow (which would indicate overconfidence) and not too wide (which would indicate high uncertainty) across the frequency band, which suggests that the model is accurate enough in this range. The confidence interval related to the GPR model of the real part of the B2 concrete permittivity (which is reported on the left side of Figure 2) shows that the GPR predictions remain tightly bounded within this interval, with the test data falling entirely within the bound it defines. In contrast, the GPR model of the imaginary part of the B2 concrete permittivity exhibits a broader confidence interval, particularly for frequencies greater than 0.4 GHz (where a test point falls marginally outside the predicted confidence region). Although the GPR model reproduces the general shape of the imaginary permittivity curve, the increasing width of the uncertainty band may point to localized model underfitting or regions of sparse training data, supporting the interpretation that the model may be potentially less accurate in this area, a behavior attributable to the low number of samples composing the dataset under consideration. Tables 2 and 3 provide a quantitative estimation of the performance offered by the different models, using MSE, MAE, and MAPE metrics, which are computed exclusively on the test set to ensure an unbiased comparison [24]. It can be observed that GPR achieves lower error values for all considered metrics.

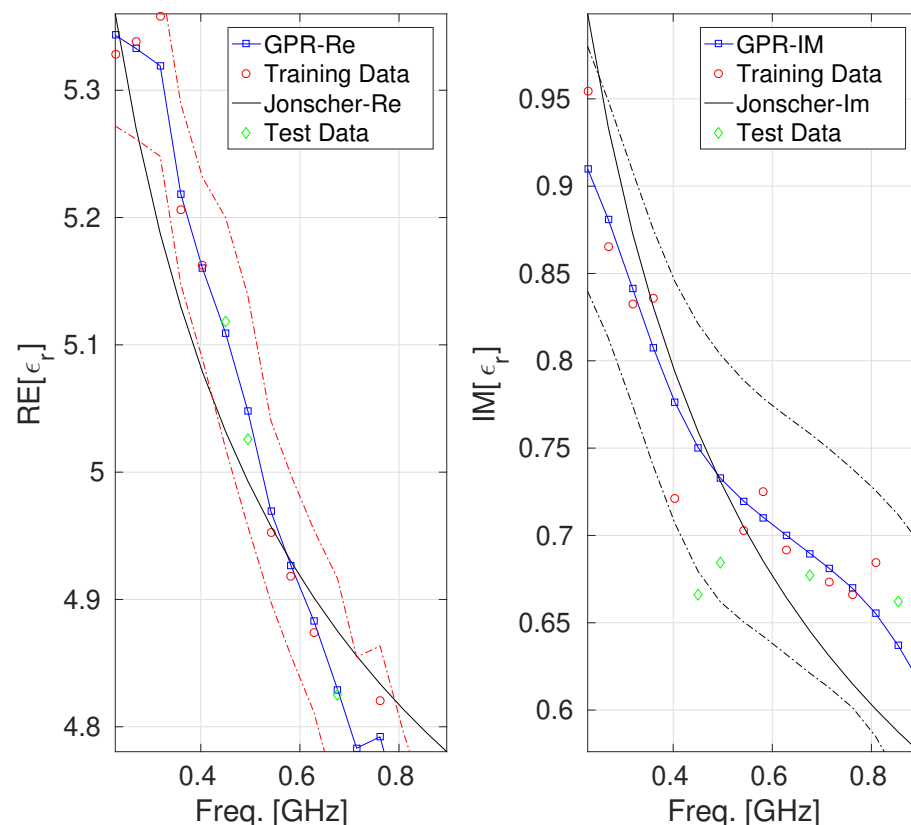


Figure 2. Behavior of the real and the imaginary parts of ϵ_r for the concrete labeled as B2 [27]: GPR model vs. Jonscher parametrization. The uncertainty bands (epsilon-tube) are also reported (identified by the red and black dashed lines).

As a second case, we considered the dataset from [30], which pertains to lightweight foamed concrete (LFC) made of sand, foam, and ordinary Portland cement. The dataset

consisted of $n = 18$ permittivity samples collected in the frequency range of $[0, 8]$ GHz, referred to as the LFC, with a density of 1368 kg/m^3 . Since, in this case, only the values of the real part of the complex permittivity were available, a GPR model was developed to characterize this quantity alone, using 70% of the samples as the training set and the remaining 30% as the test set. Figure 3 and Table 4 report the results obtained for the considered case. The confidence interval characterizing the GPR prediction, within which all test points fall, reveals, once again, that the GPR model has a more than good ability to describe the behavior of the real part of the dielectric permittivity of concrete for the case under consideration. In particular, the superiority of the GPR characterization compared to the Jonscher one is apparent, especially in the frequency range above 5 GHz.

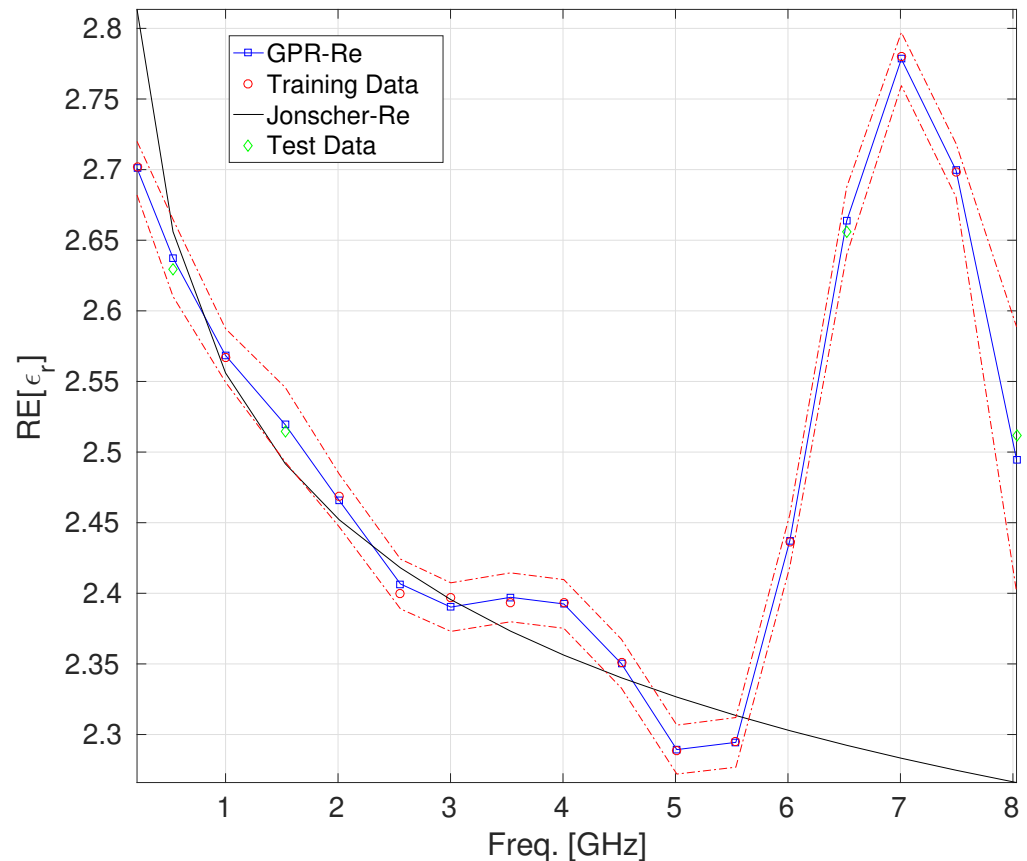


Figure 3. The behavior of the real part of ϵ_r for the lightweight foamed concrete [30]: GPR model vs. Jonscher parametrization. The uncertainty bands (epsilon-tube) are also reported (identified by the red and black dashed lines).

Table 2. Error metrics values for the ϵ_r characterization provided by the Jonscher and by the GPR models reported in Figure 1.

B1	MSE	MAE	MAPE
Jonscher(Re)	0.0422	0.0799	1.24
GPR(Re)	0.021	0.0398	0.622
Jonscher(Im)	0.0175	0.0352	3.45
GPR(Im)	0.0152	0.0289	2.65

The last case we considered regarded the dataset of permittivity from [25] about an ultra-lightweight cement composite (ULCC) fabricated using cement, water, fly ash, cenosphere, uncoiled polyvinyl alcohol (PVA) fibers (2% volume fraction), and high-range water reducers (which are chemical admixtures used in concrete to reduce the amount of water needed while

maintaining workability significantly). For this last case, the exploited dataset consisted of $n = 42$ samples collected in the range $[0, 20]$ GHz. The two GPR models we developed were trained using 65% of the available tuples, while the remaining 35% were used to build the test set. In Figure 4, a comparison among the values of the real and the imaginary parts of $\epsilon_r(f)$ provided by the Jonscher and the GPR model is reported. The confidence intervals, qualitatively slightly wider in the case of GPR predictions on the real part of $\epsilon_r(f)$ and narrower in the case of GPR predictions on the imaginary part, demonstrate, for this last case also, the goodness of the GPR modeling. Table 5 shows a more quantitative assessment of the performance of the different models using MSE, MAE, and MAPE metrics, which reveals the superiority of the GPR modeling over the Jonscher law.

Table 3. Error metrics values for the ϵ_r characterization provided by the Jonscher and by the GPR models reported in Figure 2.

B2	MSE	MAE	MAPE
Jonscher(Re)	0.0488	0.0961	2.02
GPR(Re)	0.00882	0.0167	0.349
Jonscher(Im)	0.0291	0.0614	9.26
GPR(Im)	0.0207	0.0381	5.71

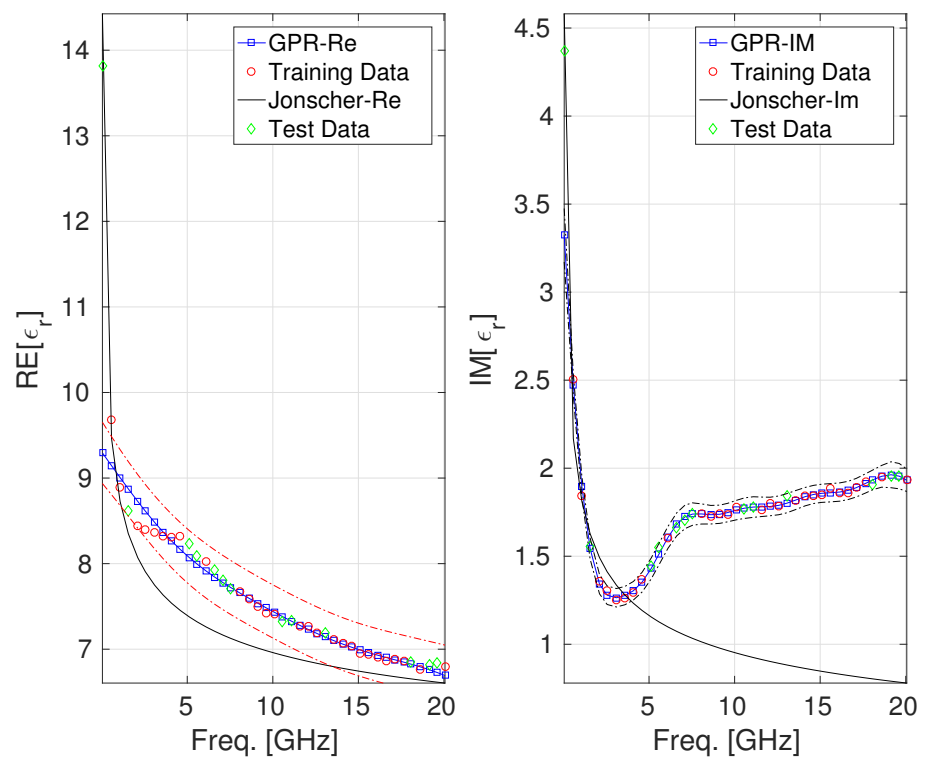


Figure 4. Behavior of the real and the imaginary parts of ϵ_r for the ultra-lightweight cement composite [25]: GPR model vs. Jonscher parametrization. The uncertainty bands (epsilon-tube) are also reported (identified by the red and black dashed lines).

Table 4. Error metrics values for the ϵ_r characterization provided by the Jonscher and by the GPR models reported in Figure 3.

LFC	MSE	MAE	MAPE
Jonscher(Re)	0.11	0.165	6.35
GPR(Re)	0.00514	0.00927	0.361

Table 5. Error metrics values for the ϵ_r characterization provided by the Jonscher and by the GPR models reported in Figure 4.

ULCC	MSE	MAE	MAPE
Jonscher(Re)	0.147	0.481	5.97
GPR(Re)	0.349	0.421	3.46
Jonscher(Im)	0.216	0.695	38.1
GPR(Im)	0.0807	0.0948	2.68

4. Conclusions

This work presents the results of a data-driven Gaussian Process Regression (GPR) modeling approach employed to characterize the complex dielectric permittivity $\epsilon_r(f)$ of concrete. The GPR models were trained on experimental datasets of concrete permittivity, and results were compared with measurements and to those provided by the well-established Jonscher universal dielectric model.

The related error metrics demonstrate that GPR can effectively characterize $\epsilon_r(f)$ in an effective manner, outperforming the Jonscher model in terms of accuracy in all the cases considered in our study. Furthermore, the ϵ -tubes generated by our Gaussian Process Regression models provide direct evidence against overfitting, along with a reliability assessment of the permittivity estimate across frequency. Although our study involved only three sets of data with specific compositions, and despite the small number of samples ($n = 17, 18, 42$) used for our numerical validations, limited by the data available in the literature on the subject, the three types of concrete considered (normal weight B1/B2, lightweight LFC, and ultra-lightweight ULCC) represent strategically different cases covering different ranges of density, composition, and frequency behavior. This peculiar characteristic of the data sets considered in our study provides strong evidence, based also on the fact that the error metrics related to the GPR model exceed those related to the Jonscher model for this spectrum of materials, of the methodological value of the proposed approach to characterize the complex permittivity of concrete.

For completeness, we would like to point out that, although we did not conduct a formal sensitivity analysis in this study, the predictive performance and superiority of GPR over the Jonscher model that we observed were found in several subdivisions of the datasets and default initialization settings, with the quasi-Newton algorithm achieving stable solutions in a few iterations (usually 7–13), thus demonstrating convergence behavior that indicates reasonable insensitivity of the GPR models we developed to specific hyperparameter values within the ranges relevant for characterizing the complex permittivity of concrete.

Finally, it should be noted that the complex permittivity of concrete $\epsilon_r(f)$ does not depend exclusively on the frequency alone, but also on its physical composition (W/C ratio, porosity) [17,18]. This last observation provides a starting point for future work, which will focus on integrating these physical parameters into a more advanced, complex permittivity GPR model capable of handling multiple typologies to estimate compressive strength from in situ electromagnetic measurements. Furthermore, since measurements of concrete dispersion parameters are often only available over a limited frequency band [25], a sequential data splitting strategy will be tested to create training and test sets for the GPR permittivity model, with the aim of evaluating its ability to extrapolate permittivity values beyond the training range.

Author Contributions: Conceptualization, G.A. and M.V.; methodology, G.A.; software, G.A.; validation, P.B., M.V. and F.L.; formal analysis, G.A.; investigation, M.V., G.A., P.B. and F.L.; resources, F.L.; data, G.A. and F.L.; writing—original draft preparation, G.A., M.V., P.B. and F.L.; writing—review

and editing, P.B. and F.L.; visualization, F.L.; supervision, G.A. and M.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been supported by the Italian Ministry of University and Research under the Program PRIN 2022: “Integration of Artificial Intelligence and Ultrasonic Techniques for Monitoring Control and Self-Repair of Civil Concrete Structures (CAIUS)” and by project FAIR “Future Artificial Intelligence Research” (CUP H23C22000860006), funded by the European Union under “NextGenerationEU”, by the Italian Ministry of University and Research (MUR), and by “Italia Domani - National Recovery and Resilience Plan (PNRR)”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

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

Abbreviations

The following symbols and acronyms are used in this manuscript:

GPR	Gaussian Process Regression;
MSE	Mean-Squared Error;
MAPE	Mean Absolute Percentage Error;
MAE	Mean Absolute Error

References

1. Kanagaraj, B.; Anand, N.; Raj, R.S.; Lubloy, E. Techno-socio-economic aspects of portland cement, geopolymer, and limestone calcined clay cement (LC3) composite systems: A-state-of-art-review. *Constr. Build. Mater.* **2023**, *398*, 132484. [\[CrossRef\]](#)
2. Angiulli, G.; Burrascano, P.; Ricci, M.; Versaci, M. Advances in the Integration of Artificial Intelligence and Ultrasonic Techniques for Monitoring Concrete Structures: A Comprehensive Review. *J. Compos. Sci.* **2024**, *8*, 531. [\[CrossRef\]](#)
3. Bouzaffour, K.; Lescop, B.; Talbot, P.; Gallee, F.; Rioual, S. Development of an embedded UHF-RFID corrosion sensor for monitoring corrosion of steel in concrete. *IEEE Sens. J.* **2021**, *21*, 12306–12312. [\[CrossRef\]](#)
4. Angiulli, G.; Calcagno, S.; La Foresta, F.; Versaci, M. Concrete Compressive Strength Prediction Using Combined Non-Destructive Methods: A Calibration Procedure Using Preexisting Conversion Models Based on Gaussian Process Regression. *J. Compos. Sci.* **2024**, *8*, 300. [\[CrossRef\]](#)
5. Mair, D.; Fischer, M.; Konzilia, J.; Renzler, M.; Ussmueller, T. Evolutionary optimization of antennas for structural health monitoring. *IEEE Access* **2023**, *11*, 4905–4913. [\[CrossRef\]](#)
6. Tenório Filho, J.R.; Goethals, J.; Aminzadeh, R.; Abbas, Y.; Valdez Madrid, D.E.; Cnudde, V.; Vermeeren, G.; Plets, D.; Matthys, S. An automated wireless system for monitoring concrete structures based on embedded electrical resistivity sensors: Data transmission and effects on concrete properties. *Sensors* **2023**, *23*, 8775. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Prattico, D.; Laganá, F.; Oliva, G.; Fiorillo, A.S.; Pullano, S.A.; Calcagno, S.; De Carlo, D.; La Foresta, F. Sensors and Integrated Electronic Circuits for Monitoring Machinery On Wastewater Treatment: Artificial Intelligence Approach. In Proceedings of the IEEE Sensors Applications Symposium (SAS), Napoli, Italy, 23–24 July 2024; pp. 1–6. [\[CrossRef\]](#)
8. Shilar, F.A.; Ganachari, S.V.; Patil, V.; Yunus Khan, T.M.; Saddique Shaik, A.; Azam Ali, M. Exploring the potential of promising sensor technologies for concrete structural health monitoring. *Materials* **2024**, *17*, 2410. [\[CrossRef\]](#)
9. Telfah, D.; Al-Mattarneh, H.; Ismail, R.; Rawashdeh, A.; Aljamal, M.; Dahim, M. Development of permittivity sensor for advanced in situ testing and evaluation of building material. In Proceedings of the 21st IEEE International Multi-Conference on Systems, Signals & Devices, Erbil, Iraq, 22–25 April 2024; pp. 164–169.
10. Bouzaffour, K.; Lescop, B.; Talbot, P.; Nguyen-Vien, G.; Gallée, F.; Rioual, S. Decoupling free chloride and water ingress in concrete by a dielectric resonant sensor. *Constr. Build. Mater.* **2023**, *372*, 130806. [\[CrossRef\]](#)
11. Gao S.; Chung, K.L.; Cui, A.; Ghannam, M.; Luo, J.; Wang, L.; Ma, M.; Liao, Z. Accurate strength prediction models of ordinary concrete using early-age complex permittivity. *Mater. Struct.* **2021**, *54*, 172. [\[CrossRef\]](#)
12. Oliveira, J.G.; Junior, J.G.D.; Pinto, E.; Neto, V.P.S.; D’Assunção, A.G. A new planar microwave sensor for building materials complex permittivity characterization. *Sensors* **2020**, *20*, 6328. [\[CrossRef\]](#)

13. Cataldo, A.; Schiavoni, R.; Masciullo, A.; Cannazza, G.; Micelli, F.; De Benedetto, E. Combined punctual and diffused monitoring of concrete structures based on dielectric measurements. *Sensors* **2021**, *21*, 4872. [[CrossRef](#)]
14. Pereira, R.N.; Júnior, J.G.D.; Praxedes, M.E.T.S.; Cabral, K.C.; da Silva Neto, V.P.; D'Assunção, A.G. A planar DGS sensor for moisture analysis in civil construction aggregates. *Sens. Actuators A Phys.* **2024**, *367*, 115042. [[CrossRef](#)]
15. Suchorab, Z.; Szczepaniak, Z.; Tabiś, K.; Łagód, G.; Brzyski, P.; Rogala, T. Comparison of reflectometric and microwave moisture detection techniques. *Aip Conf. Proc.* **2023**, *2801*, 030029.
16. González-López, G.; Blanch, S.; Romeu, J.; Jofre, L. Debye frequency-extended waveguide permittivity extraction for high complex permittivity materials: Concrete setting process characterization. *IEEE Trans. Instrum. Meas.* **2019**, *69*, 5604–5613. [[CrossRef](#)]
17. Zhang, B.; Ni, Y.; Zhong, Y. Theoretical derivation of and experimental investigations into the dielectric properties modeling of concrete. *J. Mater. Civ. Eng.* **2023**, *35*, 04022445. [[CrossRef](#)]
18. Chung, K.L.; Yuan, L.; Ji, S.; Sun, L.; Qu, C.; Zhang, C. Dielectric characterization of Chinese standard concrete for compressive strength evaluation. *Appl. Sci.* **2017**, *7*, 177. [[CrossRef](#)]
19. Angiulli, G.; Versaci, M. Electromagnetic Characterization of Concrete by Jonscher Universal Dielectric Model: Numerical Validation of Different Nonlinear Optimization Algorithms. *IEEE Access* **2025**, *13*, 142635–142650. [[CrossRef](#)]
20. Jo, S.; Cho, K.; Noh, V.; Lee, N.; Kim, S.; Yook, J.G. Electromagnetic shielding effectiveness of the concrete composite as a function of structural parameters of steel fibers. *J. Korean Inst. Electromagn. Eng. Sci.* **2021**, *32*, 517–526. [[CrossRef](#)]
21. Chen, L.; Kim, C.; Batra, R.; Lightstone, J.P.; Wu, C.; Li, Z.; Deshmukh, A.A.; Wang, Y.; Tran, H.D.; Vashishta, P.; et al. Frequency-dependent dielectric constant prediction of polymers using machine learning. *NPJ Comput. Mater.* **2020**, *6*, 61. [[CrossRef](#)]
22. Guo, H.; Song, Y.; Yin, J.; Li, J.; Liu, X. Predicting Dielectric Constant of Polyimide Composites Using Data-Driven Wavelet Kernel Deep Gaussian Process. *IEEE Trans. Dielectr. Electr. Insul.* **2022**, *29*, 1045–1052. [[CrossRef](#)]
23. Güneşer, M.T.; Atasoy, F. Extracting complex permittivity of materials by Gaussian process regression using the transmission parameter at sub-THz. *J. Electron. Mater.* **2020**, *49*, 681–688. [[CrossRef](#)]
24. Gramacy, R.B. *Surrogates: Gaussian Process Modeling, Design, and Optimization for the Applied Sciences*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2020.
25. Huang, S.; Chung, K.L.; Zheng, K.; Liu, X. An Improved Jonscher Model for Broadband Dielectric Characterization of Cement Composites. In Proceedings of the 2023 IEEE International Symposium On Antennas And Propagation (ISAP), Kuala Lumpur, Malaysia, 30 October–2 November 2023; pp. 1–2.
26. Jonscher, A.K. Dielectric relaxation in solids. *J. Phys. D Appl. Phys.* **1999**, *32*, R57. [[CrossRef](#)]
27. Bourdi, T.; Rhazi, J.E.; Boone, F.; Ballivy, G. Application of Jonscher model for the characterization of the dielectric permittivity of concrete. *J. Phys. D Appl. Phys.* **2008**, *41*, 205410. [[CrossRef](#)]
28. Wong, E. *Introduction to Random Processes*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
29. MATLAB Team. *Statistics and Machine Learning Toolbox*; The Mathworks Inc.: Natick, MA, USA, 2019.
30. Lee, K.Y.; Phua, Y.N.; Lim, S.K.; You, K.Y.; Cheng, E.M. Determination of density and compressive strength of lightweight foamed concrete using frequency domain approach. *Microw. Opt. Technol. Lett.* **2021**, *63*, 2153–2159. [[CrossRef](#)]

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