information

Article

Federated Learning of Explainable AI Models in 6G Systems: Towards Secure and Automated Vehicle Networking

Alessandro Renda [1,*], Pietro Ducange [1], Francesco Marcelloni [1], Dario Sabella [2], Miltiadis C. Filippou [3], Giovanni Nardini [1], Giovanni Stea [1], Antonio Virdis [1], Davide Micheli [4], Damiano Rapone [4] and Leonardo Gomes Baltar [3]

[1] Department of Information Engineering, University of Pisa, 56122 Pisa, Italy
[2] Intel Corporation Italia SpA, 20094 Milan, Italy
[3] Intel Deutschland GmbH, 85579 Neubiberg, Germany
[4] Telecom Italia S.p.a., 00198 Roma, Italy
* Correspondence: alessandro.renda@unipi.it

Abstract: This article presents the concept of federated learning (FL) of eXplainable Artificial Intelligence (XAI) models as an enabling technology in advanced 5G towards 6G systems and discusses its applicability to the automated vehicle networking use case. Although the FL of neural networks has been widely investigated exploiting variants of stochastic gradient descent as the optimization method, it has not yet been adequately studied in the context of inherently explainable models. On the one side, XAI permits improving user experience of the offered communication services by helping end users trust (by design) that in-network AI functionality issues appropriate action recommendations. On the other side, FL ensures security and privacy of both vehicular and user data across the whole system. These desiderata are often ignored in existing AI-based solutions for wireless network planning, design and operation. In this perspective, the article provides a detailed description of relevant 6G use cases, with a focus on vehicle-to-everything (V2X) environments: we describe a framework to evaluate the proposed approach involving online training based on real data from live networks. FL of XAI models is expected to bring benefits as a methodology for achieving seamless availability of decentralized, lightweight and communication efficient intelligence. Impacts of the proposed approach (including standardization perspectives) consist in a better trustworthiness of operations, e.g., via explainability of quality of experience (QoE) predictions, along with security and privacy-preserving management of data from sensors, terminals, users and applications.

Keywords: explainable artificial intelligence; federated learning; 6G; vehicle-to-everything (V2X); quality of service; quality of experience

Citation: Renda, A.; Ducange, P.; Marcelloni, F.; Sabella, D.; Filippou, M.C.; Nardini, G.; Stea, G.; Virdis, A.; Micheli, D.; Rapone, D.; et al. Federated Learning of Explainable AI Models in 6G Systems: Towards Secure and Automated Vehicle Networking. *Information* **2022**, *13*, 395. https://doi.org/10.3390/info13080395

Academic Editors: Pierre-Edouard Portier and Gabriele Gianini

Received: 7 July 2022
Accepted: 17 August 2022
Published: 20 August 2022

1. Introduction

Artificial Intelligence (AI), along with Machine Learning (ML) as one of its core building blocks, is entering many market domains at a fast pace and will not only leverage advanced communication networks but also shape the definition of next-generation networks themselves. In particular, AI is expected to play a crucial role in the design, operation and management of future beyond-5G (B5G)/6G networks and in a plethora of applications [1]. However, the introduction of in-network AI comes with growing concerns on privacy, security and trust for citizens and users; for this reason, the adoption of eXplainable AI (XAI) models is an emerging trend considered for the design of transparent AI-based solutions. Moreover, future service scenarios, especially in the automotive domain, will be characterized by the deployment of connected vehicular systems from heterogeneous car manufacturers, connected via different Mobile Network Operators (MNOs) and different technology infrastructures [2]. In such complex setups, it will be imperative for service providers to consider federated network environments including multiple administrative and technical domains as a working assumption for the design of innovative applications. It is worth noting that the automated driving use case of "Teleoperated Driving (ToD) for Remote Steering" [3] requires a throughput of up to 36 Mbps per single stream, along with a positioning accuracy of 0.1 m and a reliability of 99.999% for the service to be considered available to the end customer. Such stringent requirements call for new technical enablers, to be introduced as part of the 6G network design. Considering the above-mentioned challenges, in this article we envision the use of the federated learning (FL) concept applied jointly with XAI models and discuss its applicability to automated vehicle networking use cases to be encountered in B5G/6G setups. In fact, although FL has recently been widely investigated in the context of Neural Networks and Deep Learning models (due to their gradient based optimization strategy), much less attention has been devoted so far to FL of XAI models.

The main contributions of this article can be summarized as follows:

- We propose the integration of FL with XAI for performing quality of experience (QoE) predictions in B5G/6G networks, by providing a detailed discussion about the benefits it can bring and the main challenges that need to be addressed;
- Considering vehicle-to-everything (V2X) applications as relevant use cases, we present the design of a framework to evaluate the benefits of the proposed approach and provide the guidelines to implement a realistic B5G/6G network testbed supporting the training of XAI models in a federated fashion, as well as the issuance of explainable QoE predictions;
- We shed light on the impact that the proposed FL approach with XAI models will have on both the industrial and standardization sectors.

While the following subsections provide an overview of XAI and FL, respectively, Section 2 describes the FL of XAI models applied to advanced 5G systems towards 6G. Section 3 elaborates on some relevant V2X use cases and provides more details on the proposed FED-XAI framework, focusing on a QoE prediction task. In Section 4, the impact of such solutions to the automotive vertical segment is discussed: in particular, the benefits of predicted QoE explanations useful for decision making are detailed for both car Original Equipment Manufacturers (OEMs) and MNOs. Standardization impacts are also analysed to provide interoperable and globally applicable solutions, and some challenges of FL of XAI models are discussed. Finally, Section 5 draws some conclusions.

1.1. The Need for XAI

The adoption of AI techniques cannot disregard the fundamental value of trustworthiness, which, along with inclusiveness and sustainability, represents the three core values of the European Union Flagship Hexa-X (www.hexa-x.eu (accessed on 16 August 2022)) vision for the upcoming 6G era [1]. Trustworthiness has become paramount for both users and government entities, as witnessed by the "right to explanation" described in the General Data Protection Regulation (GDPR) and by the European Commission's (EC) Technical Report on "Ethics guidelines for trustworthy AI" [4]. According to these, explainability represents a key requirement towards trustworthiness. Thus, industry and academia are placing increasing attention on XAI, that is, an AI "that produces details or reasons to make its functioning clear or easy to understand" [5].

In this context, two strategies for achieving explainability can be identified [5]: the adoption of post-hoc explainability techniques (i.e., the "explaining black-box" strategy) and the design of inherently interpretable models (i.e., "transparent box design" strategy). In this article, we focus on this latter class of approaches, noting that certain applications may tolerate a limited performance degradation to achieve fully trustworthy operation. In fact, performance and transparency are typically considered conflicting objectives [5,6]. However, this trade-off holds as long as the target task entails a certain complexity and the data available are many and high quality. In this case, complex models, such as Deep Neural Networks (DNNs), which are hard to interpret due to their huge number of parameters and non-linear modelling, have proven to achieve high levels of accuracy; conversely, decision trees and rule-based models may feature lower modelling capability but are typically considered “highly interpretable”.

The importance of explainability has been recently highlighted in the context of Secure Smart Vehicles [7]: on one hand, explanation is crucial in safety-critical AI-based algorithms, designed to extend some widely available capabilities (e.g., lane-keeping and braking assistants) towards fully automated driving; on the other hand, explainability is needed at the design stage to perform model debugging and knowledge discovery, thus positively impacting system security by reducing model vulnerabilities against external attacks. Explainability of AI models will be crucial for 6G-enabled V2X systems. A prime example is an AI service consumer requesting in-advance notifications on QoS predictions, as studied in Hexa-X [1] and the 5G Automotive Association (5GAA) [8]. Accurate and timely predictions should support very demanding use cases, with a horizon ranging from extremely short to longer time windows. Better explainability of such predictions and any consequent decision will provide benefits not only for technology and service providers (see Section 4), but also for end-customers, who will become more receptive to AI-based solutions.

1.2. Federated Learning

Exploiting data from multiple sources can enhance the performance (i.e., high accuracy based on reduced bias) of AI models. However, wirelessly collecting and storing peripheral data for processing on a centralized server has become increasingly impractical due to two main reasons: first, it typically introduces severe communication and computation overhead due to the transmission and storage of large training data sets, respectively; second, it violates the privacy and security requirements imposed by data owners by expanding the surface of possible over-the-air attacks towards biased decision making. In other words, the preservation of data privacy represents an urgent requirement of today’s AI/ML systems, because data owners are often reluctant to share their data with other parties; in some jurisdictions, users have the ability to consent or not with the sharing of privacy-sensitive data (e.g., per the General Data Protection Regulation—GDPR in European Union). Such a need to preserve privacy of data owners, however, clashes with the need to collect data to train accurate ML models, which are typically data hungry in their learning stage. To overcome these limitations, FL has been proposed as a privacy-preserving paradigm for collaboratively training AI models. In an FL system, participants iteratively learn a shared model by only transferring local model updates and receiving an aggregated shared model update, without sharing raw data.

The main opportunities of FL in the context of Intelligent Transportation Systems (ITS) have been recently discussed in [9]: FL is expected to support both vehicle management (i.e., automated driving) and traffic management (i.e., infotainment and route planning) applications. Furthermore, FL has been applied in the context of Ultra-Reliable Low-Latency Communications for Vehicle-to-Vehicle scenarios, allowing vehicular users to estimate the distribution of extreme events (i.e., network-wide packet queue lengths exceeding a predefined threshold) with a model learned in a decentralized manner [10]. The model parameters are obtained by executing maximum likelihood estimation in a federated fashion, without sharing the local queue state information data. The concept of Federated Vehicular Network (FVN) has been recently introduced [11], as an architecture with decentralized components that natively support applications, such as entertainment at sport venues and distributed ML. However, FVN is a stationary vehicular network and relies on the assumption that vehicles remain at a fixed location, e.g., parking lots, so that the wireless connection is stable.

In most of the work on FL, the strategy for model aggregation was inspired by the federated averaging protocol (FedAvg), which enables collaborative Stochastic Gradient Descent (SGD) optimization in a federated manner. Thus, FL has been extensively investigated for models implementing SGD as their optimization method, such as Neural Networks (NNs), but has not yet been adequately studied in the context of inherently explainable models. The following section introduces how XAI models can be generated by FL, a new approach which appears very promising for future 6G systems.

2. FED-XAI: Bringing together Federated Learning and Explainable AI

Existing AI-based solutions for wireless network planning, design and operation ignore either or both of the following aspects: (i) the need to preserve data privacy at all times, including wireless transfer and storage, and (ii) the explainability of the involved models. Furthermore, latency and reliability requirements of safety-critical automotive communications call for seamless availability of decentralized and lightweight intelligence, where data are generated—and decisions made—anytime and anywhere.

Current FL approaches only address the first requirement. Explainability has been given less attention, having been approached primarily by exploiting post-hoc techniques, e.g., Shapley values to measure feature importance [12]. There is a substantial lack of approaches for FL of inherently explainable models. On the other hand, a federated approach for learning interpretable-by-design models, in which transparency is guaranteed for every decision made, would represent a significant leap towards trustworthy AI. Therefore, we introduce the concept of FL of XAI (FED-XAI) models, as a framework with a twofold objective: first, to leverage FL for privacy preservation during collaborative training of AI models, especially suitable in heterogeneous B5G/6G scenarios; second, to ensure an adequate degree of explainability of the models themselves (including the obtained aggregated model as a result of FL).

In the following, we provide some insights into how inherently explainable models (e.g., decision trees or rule-based) can be learned, employing an FL paradigm. First, it is worth noting that standard algorithms for learning such models typically adopt a heuristic approach; in fact, gradient descent-based optimization methods, widely used in FL, cannot be immediately applied, as they require the formulation of a global objective function. The greedy induction of decision trees, for example, recursively partitions the feature space by selecting for each decision node the most suitable attribute. The major challenge of the FED-XAI approach, therefore, consists in generating XAI models, whose FL is not based on the optimization of a differentiable global objective function.

The proposed FED-XAI approach relies on orchestration by a central entity but ensures that local data are not exposed beyond source devices: each data owner learns a model by elaborating locally acquired raw data and shares such a model with the central server, which merges the received models to produce a global model (Figure 1). Notably, our envisioned approach for federated learning of explainable AI models ensures data privacy regardless of the data sample size. As per the advantages of the FED-XAI approach, we expect that the global aggregated model performs better than the local models because it exploits the overall information stored and managed by all data owners, without compromising model interpretability.

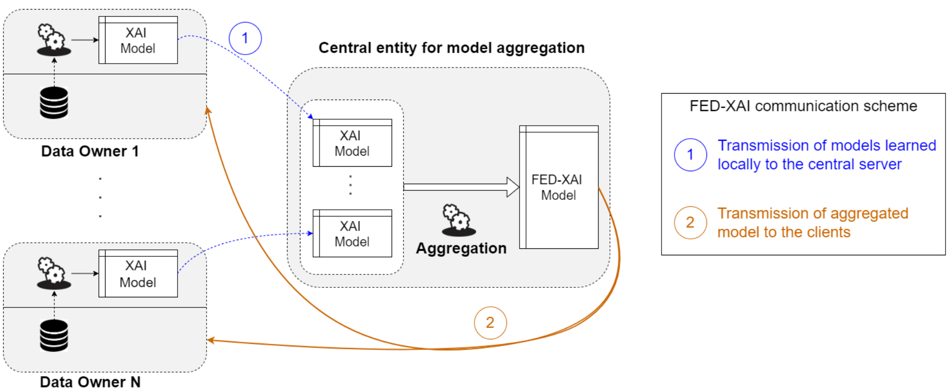


Figure 1. Illustration of federated learning of XAI models.

Our approach differs from classical FL in two aspects: first, it entails a one-shot communication scheme for each model update and not an iterative algorithm. As a consequence, the communication overhead is reduced, and the system is more robust to possible connectivity problems. Second, merging decision trees and rule-based models requires defining appropriate procedures, necessarily different from the simple weighted average of models of the FedAvg protocol applied, for example, to NNs. In more detail, the XAI models we consider can be represented as collections of "IF *antecedent* THEN *consequent*" rules, (natively in a rule-based system, and easily obtainable also from a decision tree). This representation is applicable regardless of the target task (regression or classification) and the type of the attributes (e.g., nominal or numeric). The aggregation procedure consists in juxtaposing rules collected from data owners, and resolving possible *conflicts*, which emerge when rules from different models, having antecedents referring to identical or overlapping regions of the attribute space, have different consequents. In one of our recent works [13], we presented a novel approach for FL of Takagi-Sugeno-Kang (TSK) fuzzy rule based systems [14], which can be considered as XAI models in regression problems. In a TSK model, the antecedent of a rule identifies a specific region of the attribute space, whereas the corresponding consequent allows for the evaluation of the predicted output within such a region as a linear combination of the input variables. When two rules, generated by different clients, share the same antecedent, the aggregation strategy for generating the FED-XAI model involves combining the two rules into a single one with the same antecedent: the coefficients of the linear model of the new consequent are evaluated as the weighted average of the coefficients of the original rules, where the weight of each rule depends on its support and confidence values. Research efforts in the FED-XAI domain, however, are still in their embryonic stage: as for tree-based models, a preliminary investigation of the trade-off between accuracy and interpretability has been recently carried out [15], but learning strategies compliant with the federated setting still need to be sharpened.

The FED-XAI approach may find immediate applicability in the automated vehicle networking domain, and, specifically, within the exemplary scenario described in Section 1: a model for QoS prediction which is explainable by a wide set of service consumers may be learned in a federated manner, reaping the benefits of collaborative training and privacy protection.

Main Challenges of the FED-XAI Approach

There are also challenges related to the FED-XAI approach, especially for time-critical operations in automated driving setups. For example, the computation (and, therefore, energy) footprint of FED-XAI needs to be pre-evaluated before implementation to identify the scalability potential of the solution. A clear distinction should be made between the stages of *training* and *inference*. For most ML models, including decision tree and rule-based systems, the inference time (critical from automated driving service standpoint) is negligible compared to the training time and, in any case, model complexity can be tuned to ensure that time constraints are satisfied. A larger computational overhead is required in the training stage, but it does not affect the application (e.g., learning can be performed in idle state). Another challenge is FED-XAI system resilience to attackers trying to benefit from the access to explanations of QoE predictions (e.g., towards increasing automated driving service outages for all or targeted vehicles). Finally, the approach will also need to address some additional challenges that are typical of FL and are likely to characterize 6G network-based intelligent transportation applications: (i) multi-source data may have different distributions and volumes, (ii) the number of participants can grow fast and their participation to FL may be unstable due to insufficiency of radio and computational resources, and (iii) learned models will need to be agilely updated in scenarios where concept drift alters the characteristics of data distributions over time.

3. The Proposed FED-XAI Framework for QoE Predictions in V2X Environments

This section describes some V2X use cases for which the FED-XAI approach is expected to be beneficial. Furthermore, we discuss a framework for QoE prediction in B5G/6G systems along with realistic sourcing of live data from an MNO network.

3.1. Exemplary 6G Use Cases in V2X Environments

The use of Information and Communication Technology, and especially AI techniques, in the automotive sector is gaining increasing attention [16]. Given the large amount of data generated by multiple, distributed sources, AI is one of the key technologies to enable innovative use cases, such as autonomous driving [17,18], improved safety [19] and platooning [20].

In the Hexa-X project, a general AI-assisted V2X use case is described ([1], Section 4.2.7.3). In future enhanced automotive uses cases or services, collecting a high volume of contextual and sensor data from traffic participants and road infrastructure will be common practice. With these data, a Digital Twin (DT) of the traffic environment can be created in the cloud or the edge of future mobile networks. This DT can be distributed across multiple edge nodes corresponding to a coverage area or to higher hierarchical edge-cloud nodes in different locations of the network and employed to optimize vehicle traffic by generating inputs to traffic management as well as driving or manoeuvring instructions to traffic participants. Moreover, DTs can be used to support ToD, by providing real-time information and predictions of road traffic information, as well as predictions for the QoS of the radio access network (e.g., radio signal quality). Real-time management of DTs is very challenging and requires network capabilities not available today. To guarantee safety, system operation requires extremely low latency, high reliability and ultra-high location accuracy along with efficient and explainable AI algorithms. Multiple edge nodes can be part of MNO or road infrastructure, and, at a given time, geographically proximate nodes might contain similar AI models which are either part of the same learning federation or updated by applying knowledge sharing. Moreover, when it comes to safety-related V2X services for automated and connected vehicles, most of the use cases analysed by 5GAA [2], e.g., See Through, Vulnerable Road User protection, Intersection Movement Assist, or In-Vehicle Entertainment, are evaluated, attributing great significance to security and privacy. In this perspective, the FED-XAI approach provides an intrinsic benefit, compared to its centralized learning counterparts.

In [1] authors proposed to further improve the ToD use case by applying the AI-as-a-Service approach: a driver planning to perform a journey would like to be informed of any V2X service degradations along the planned route by means of in-advance QoS predictions based on a plurality of data, gathered, e.g., from the Uu and PC5 interfaces but also from vehicle sensors (RADAR, LiDAR, etc.). Such notifications will allow for the decision upon activation of other V2X-related functionalities: for instance, automated driving features should be avoided if the predicted QoS in a certain part of the route would not allow such features to be used; hence, the driver should take control of the car until new, favourable QoS predictions suggest switching to automated driving mode. Moreover, QoS predictions could also be used to schedule the execution of non-V2X functionalities, e.g., software over-the-air downloads. In addition, in this scenario, the various data owners are encouraged to participate in the FED-XAI procedure, because the collaborative model will blend the knowledge extracted from all data rather than only from local data.

In the following section, we describe an illustrative automotive scenario, which can be the basis of evaluating the performance of the FED-XAI concept.

3.2. Details of the Proposed FED-XAI Framework

We consider an application where several instances of vehicular User Equipment (UE) connected to a B5G/6G Base Station (BS) receive a video stream whose quality plays a decisive role in the safety of remote driving. The quality of experience (QoE) perceived by UEs depends on the QoS provided by the network. This can be mapped, for instance, to a see-through use-case, where the receiving UE is a car using a live feed from the camera of another car (e.g., to make overtaking safer in the presence of visual impairments for the driver), or to ToD, where the sender is the car being driven and the receiver is the driving operator. In either case, operations may be supported by a DT of the traffic area at the edge. From an implementation perspective, Multi-access Edge Computing (MEC) infrastructure can be leveraged as an intrinsic facilitator for the segregation of data, as requested by international regulation in the matter of ITS services. Furthermore, state-of-the-art MEC technologies are defined by considering MEC security with an end-to-end approach, leveraging existing standards relevant in the area, e.g., ETSI-NFV-SEC (European Telecommunications Standards Institute—Network Functions Virtualisation—Security), TC CYBER (Technical Committee Cybersecurity), 3GPP (Third Generation Partnership Project), carefully selected to be applicable in edge computing systems.

The objective of the envisioned application is to employ XAI models, learnt (and updated) in a federated fashion based on QoS/QoE data, to predict the QoE perceived by UEs in the near future. Notably, it has been recently shown that highly interpretable tree-based models are able to achieve competitive performance in this specific task [21]. In the following, we describe how the FED-XAI approach can be deployed on a MEC-enabled B5G/6G architecture (see Figure 2). A FED-XAI computation engine (CE) instantiated within an edge/cloud node (also called a MEC host) is the central entity responsible for model aggregation. With reference to the see-through scenario in Figure 2, each UE measures QoS and QoE metrics, while receiving the video stream, possibly enhanced by the DT. Examples of relevant metrics are: received throughput, jitter, packet error rate (QoS), startup delay, number of stall events and rebuffering ratio (QoE). The collected values of such QoS/QoE metrics are securely transmitted by each (vehicular) UE to a corresponding MEC application, called FED-XAI manager (FM), possibly together with any other relevant information for the FED-XAI learning algorithm. Each UE communicates with its dedicated FM, which acts as an interface to the CE in the MEC system and provides all the functionalities to allow the UE to participate in the federation, e.g., join/leave, transmit/obtain model to/from the CE, etc. Alternatively, an FM can be hosted directly at the UE (instead of the MEC host).

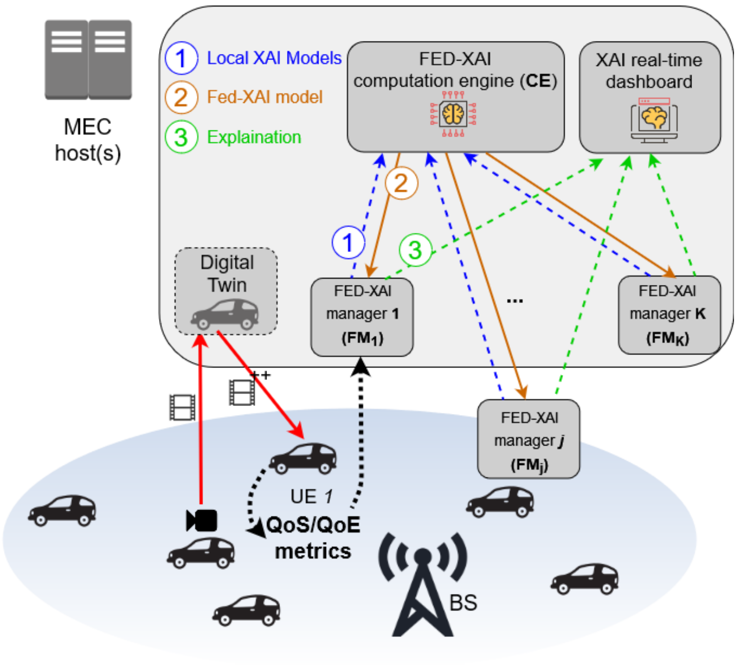


Figure 2. Example of video flow (red arrows) and related QoS/QoE metrics reporting (dashed black arrows) in a MEC-enabled FED-XAI architecture. Interaction among FM, CE and a real time XAI dashboard is also shown.

We divide time in periods, as shown in Figure 3. During each period, a UE measures both QoS and QoE metrics. We call $\mathbf{QOS}_k(i)$ and $\mathbf{QOE}_k(i)$ the vectors of QoS and QoE metrics measured by UE_k during period i. Both vectors are sent to the FM of UE_k, FM_k. At $t = n$, FM_k uses the XAI model obtained by the CE to predict the QoE that UE_k will perceive in the next period, i.e., $QOE_k(n+1)$.

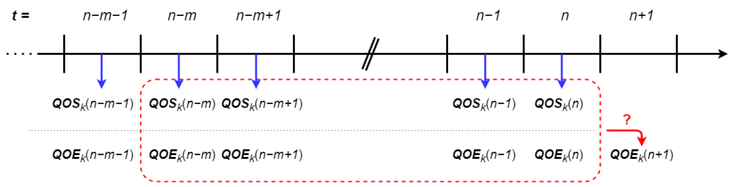


Figure 3. QoE prediction based on previous QoS/QoE samples.

For building (or updating) the FED-XAI model, the involved FMs train (or update) the local model based on recent data ($\mathbf{QOS}(i)$ and $\mathbf{QOE}(i)$ for each $i = n - m, n - m + 1, \ldots, n$, where m is a predefined time window), and share it with the CE. Once the CE produces the aggregated FED-XAI model, the latter is sent back to the FMs that will use it to perform the QoE prediction for their corresponding UE. The results of the prediction feed a dashboard that displays them in real time and explains how they were obtained.

The above scenario will be evaluated in a real-time distributed testbed, which embodies both the communication and computation aspects of the system, as well as the application logic. The communication is realized by Simu5G, a modular simulator of 3GPP-compliant New Radio based on OMNeT++ [22], which also works in real time and interfaces with external applications [23]. The MEC subsystem is realized using Intel's OpenNESS open-source framework (www.openness.org (accessed on 16 August 2022)). Moreover, QoS information is taken in real time from Simu5G, also through its MEC service interface, that can be queried by MEC applications. In order to make the aforementioned testbed more realistic, the network scenario implemented by Simu5G is designed considering data taken from TIM's live network as input, such as base stations position and user data volume, extracted using the techniques described in Section 3.3.

We have recently carried out a preliminary experimental analysis focused on QoE forecasting in B5G/6G networks [24]: we have presented a novel data set (QoE forecasting data set: http://www.iet.unipi.it/g.nardini/ai6g_qoe_dataset.html (accessed on 16 August 2022)) obtained through realistic network simulations and showed how decision trees as an inherently explainable model can be considered a valid baseline for the prediction task. Specifically, the data set consists of time-tagged contextual (e.g., UE position), QoS (e.g., Signal to Interference plus Noise Ratio (SINR) value measured at packet reception) and QoE (e.g., percentage of a frame arrived at the time of its display) metrics from 24 repetitions of a scenario in which 15 instances of UE experiment with a video for approximately 120 s. The prediction task has been formulated as a regression problem. The preprocessing and feature extraction steps are extensively described in [24]; in a nutshell, for each UE we collected the timeseries related to 12 metrics (QoS, QoE and contextual) and obtained any record of the preprocessed data set as follows: for a timestamp t, the input variables consist of 11 statistics (i.e., mean, median, max, min, variance, standard deviation, kurtosis, skewness, Q1 and Q3, number of samples) measured for each metric in the time window $[t - W, t]$ (with W = 10 s), whereas the output variable consists in the mean of the target QoE metric over the time horizon of one second (i.e., in $[t, t + H]$, with H = 1 s). For the preliminary experimental analysis, we considered the *centralized* setting, i.e., all data available on a single node, and resorted to the decision tree for regression available in scikit-learn (https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeRegressor.html (accessed on 16 August 2022)). Figure 4 reports an example of real and predicted timeseries for a given QoE metric. Although our final goal is to learn XAI models in a *federated* fashion, such a preliminary analysis allowed us to set a baseline for the *centralized* setting and to assess the performance of an XAI model in a prediction task on realistic B5G/6G network simulations.

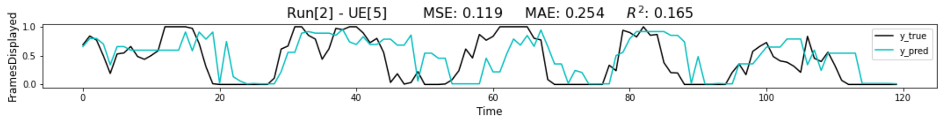
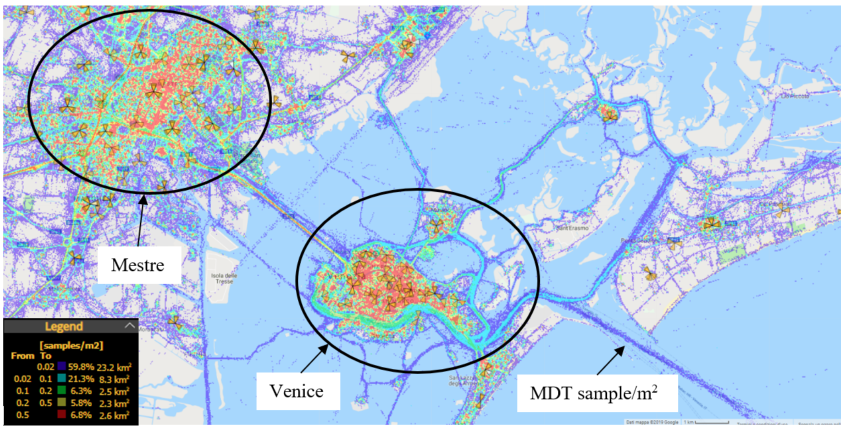


Figure 4. Real and predicted values of QoE for an example UE of the test set (Figure from [24]).

3.3. Feeding Models with Real Network Data

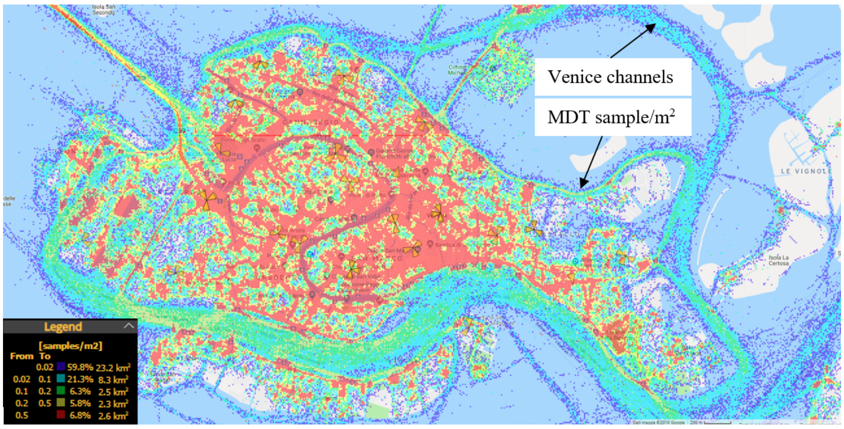
AI-based algorithms need to be fed with real data and collecting live measurements from the MNO network is critical for the reliability of the produced output. In that perspective, the Minimization of Drive Tests (MDT) functionality is applied on TIM Radio Access Network (RAN) to acquire geolocated real data from live RAN. MDT is a 3GPP standard feature (TS 37.320) which allows collecting geolocated radio measurements from UEs in both idle and connected states [25]. This enables UEs to periodically send a large set of measurements from Layer 2—MAC (3GPP TS 36.321), and Layer 3—Radio Resource Control (3GPP TS 36.331). UEs take those measurements for standard procedures, such as cell re-selection, handovers, quality reporting, etc. With MDT, UEs keep measuring the same quantities but share their measurements periodically with the network. If a Global Positioning System (GPS) receiver is enabled, UE measurements are geolocated, which allows anonymous data collection for statistical analysis. MDT paves the way for replacing traditional time-consuming drive tests, made by few test UEs, with thousands or millions of measurements, reported by most devices in the network. Moreover, MDT data come from several customers and UE types, thus allowing a realistic insight into user QoE. MDT measurements in this work mainly refer to geolocated UE throughput and data volume metrics. Examples of MDT georeferenced data from the TIM live network are reported in Figure 5 for the area around Venice, where every pixel represents 1 m^2 of the area, and MDT data in each pixel are averaged and normalized to all MDT data collected in all pixels.

Most MDT data are geolocated in the sea and Venice channels too. This highlights the capability of MDT data to represent the real traffic scenario. The real radio coverage map of several cells in the 1800 MHz frequency band is reported with different colours. MDT data are completely anonymous because neither customer nor UE identity data are monitored or gathered. In particular, MDT data are only geolocated radio measurements reported by UE. The MDT data on TIM's RAN are gathered by using a Nokia system called Geosynthesis.

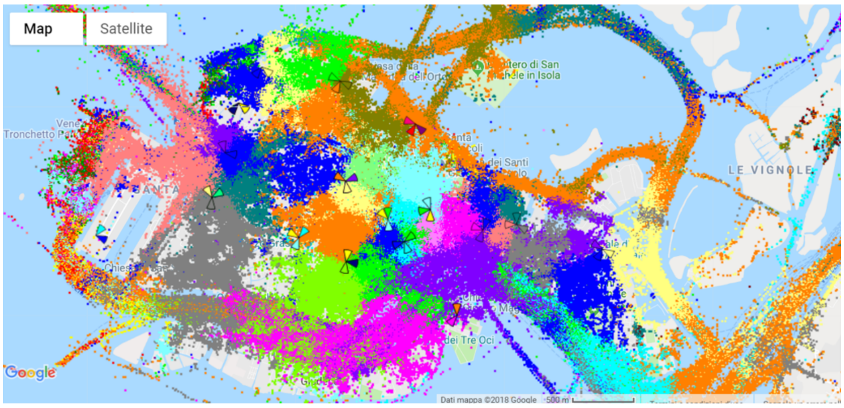


(**a**)

Figure 5. *Cont.*



(**b**)



(**c**)

Figure 5. MDT georeferenced data from TIM live network on the area of Mestre and Venice cities. (**a**) MDT sample density (MDT sample/m^2) in the area of Mestre and Venice Cities (Italy) averaged over 24 h. On the map, each single visible point is a square pixel with 1 m side. (**b**) MDT sample density on Venice city. (**c**) MDT Best Serving Cell radio coverage of Venice Cities (Italy). Each colour identifies a single cell in the 1800 MHz frequency band. On the map, each single visible point is a square pixel with 1 m side.

4. Impacts of the Proposed FED-XAI Approach on V2X Applications in B5G/6G Networks

The practical use value (or benefit) of FL for the considered V2X use cases, involving various local models instantiated at vehicles, roadside units and edge cloud infrastructure is that the QoE predictions issued by the resulting global aggregated model will be characterized by a higher generalization capability in space and in time than local models. In other words, when a QoE prediction request is issued by a (new) vehicle entering the area of interest, the QoE prediction value to be returned in response will be more accurate and of higher confidence than the one generated by a local model trained by using only the data acquired in the vehicle. On top of the increased generalization capability, a second benefit of the FED-XAI approach is the increase in trust in AI for 6G-enabled services. This has an immediate business impact on 6G business entities. This increase in trust is beneficial for all system entities, from end users to operators and service providers, edge-computing providers and other vertical market players (e.g., automotive, industrial automation, etc.), as it better instils collaboration, starting from a business level. For instance, the exemplary V2X applications described in Section 3.1 are typical cases where a collaboration (and related business agreement) is needed between MNOs, possibly in partnership with edge-computing service providers and car OEMs. Both car OEMs and MNOs can benefit from explanations about such predictions and any consequent decision making: MNOs can provide a more explainable set of 6G functionalities (e.g., FL agents enabling QoE predictions) and expose them to their customers (including car OEMs but also application developers and system integrators); OEMs can also benefit from more information on network predictions, exploitable to improve automated driving features offered to their end customers (i.e., the actual drivers).

Figure 6 shows the same V2X service scenario, characterized by a fundamentally different view, depending on the perspective considered: a car dashboard (from the point of view of the user) shows the set of enabled V2X functionalities and their respective space and time availability. Instead, the view of a network operator (providing the needed communication and computation infrastructure) embeds more complexity, including the management and operation of the network infrastructure and the FED-XAI functionalities needed to provide QoE predictions for the offered V2X services. The boundary between these two worlds is typically governed by a set of Service Level Requirements (SLRs), defining the terms and conditions of the agreement between these two stakeholders (see 5GAA reports for the V2X cases [2]). These SLRs are service-specific and can be defined in terms of minimum throughput, maximum delay, but also availability and reliability of the guaranteed KPIs (defined, e.g., in a certain time window). In this perspective, moving towards 6G, accurate and timely predictions (supported by XAI models) are key to providing advanced and very demanding use cases, with a horizon ranging from extremely short to long time windows. Therefore, it is evident how FED-XAI is paramount for improving the understanding and mutual trust among 6G business entities (i.e., MNOs and OEMs here).

Moreover, the industry is moving towards the adoption of MEC Federations, characterized by a multi-MNO environment, where each operator can provide and share with the other federating entities its own edge-computing infrastructure and services to third parties, also in collaboration with other operators. Dually, operators can consume the resources of other operators (e.g., in countries where they do not have a network infrastructure), offering a seamless user experience to their own customers. Again, the FED-XAI approach is particularly beneficial in such challenging scenarios, which are likely to become widely adopted in B5G and 6G systems.

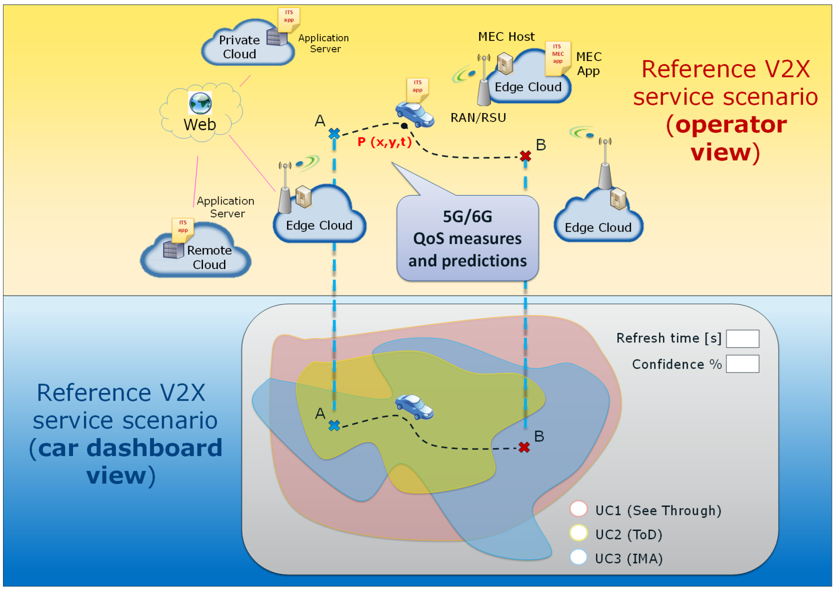


Figure 6. Reference V2X service scenarios, with MNO perspective and customer/car OEM perspective.

Standardization Impact of an Interoperable FED-XAI Implementation

An interoperable implementation of the FED-XAI concept with a focus on an automotive scenario is expected to stimulate discussion within Standards Development Organizations (SDOs) on specifying the involved architectural entities (e.g., FED-XAI CE and FMs), communication interfaces and service protocols including the exchanged data structures. For instance, the International Telecommunication Union (ITU) Focus Group on AI for Autonomous and Assisted Driving (FG-AI4AD) aims to internationally harmonize the definition of a minimal performance threshold for AI "on the road". Additionally, the ETSI Industry Specification Group on Securing AI (ISG SAI) has introduced a new work item on explainability and transparency of AI processing (June 2021). AI platform design issues aiming to provide assurance of explainability and transparency of decisions and allowing independent determination of biases are in scope. ETSI ISG MEC is also of relevance, as one of its specifications (GS MEC 030) focuses on defining the V2X Information Service (VIS) for MEC systems: one of VIS functionalities is to facilitate issuing journey-aware QoS predictions.

5. Conclusions

The aim of this article was to provide a comprehensive vision of AI-pervasive 6G networks that will be extremely high performing, intelligent and trustworthy by design, with a particular applicability to automated vehicle networking. We have introduced the FED-XAI concept, proposing federated learning of XAI models. The FED-XAI concept, applied to advanced 5G systems towards 6G, is expected to improve the user experience of the offered communication services by helping end users trust in-network AI functionality. Benefits of the proposed approach (that could also include standardization) consist in better trustworthiness of operations, e.g., via explainability of QoE predictions, along with security and privacy-preserving management of data from sensors, terminals, users and applications for a range of automotive use cases.

Author Contributions: Conceptualization, A.R., P.D., F.M., D.S., M.C.F., G.N., G.S., A.V., D.M., D.R. and L.G.B.; Investigation, A.R., P.D., F.M., D.S., M.C.F., G.N., G.S., A.V., D.M., D.R. and L.G.B.; Writing—original draft, A.R., P.D., F.M., D.S., M.C.F., G.N., G.S., A.V., D.M., D.R. and L.G.B.; Writing—review and editing, A.R. and P.D. All authors have read and agreed to the published version of the manuscript.

Funding: Part of this work was funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101015956 Hexa-X.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Conflicts of Interest: M.C.F. and L.G.B. are employed by Intel Deutschland GmbH, part of Intel Corporation. D.S. is employed by Intel Corporation Italia SpA, part of Intel Corporation. The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

3GPP	Third Generation Partnership Project
AI	Artificial Intelligence
B5G	Beyond 5G
CE	Computation Engine
DNN	Deep Neural Network
DT	Digital Twin
ETSI	European Telecommunications Standards Institute
FedAvg	Federated Averaging
FED-XAI	Federated learning of explainable Artificial Intelligence
FL	Federated Learning
FM	FED-XAI Manager
FVN	Federated Vehicular Network
GPS	Global Positioning System

ISG	Industry Specification Group
ITS	Intelligent Transportation Systems
MDT	Minimization of Drive Tests
MEC	Multi-access Edge Computing
ML	Machine Learning
MNO	Mobile Network Operators
NFV	Network Functions Virtualisation
NN	Neural Network
OEM	Original Equipment Manufacturers
QoE	Quality-of-Experience
QoS	Quality-of-Service
RAN	Radio Access Network
SGD	Stochastic Gradient Descent
SINR	Signal to Interference plus Noise Ratio
SLR	Service Level Requirements
ToD	Teleoperated Driving
TSK	Takagi-Sugeno-Kang
UE	User Equipment
V2X	Vehicle-to-Everything
XAI	Explainable Artificial Intelligence

References

1. Hexa-X Deliverable D1.2—Expanded 6G Vision, Use Cases and Societal Values—Including Aspects of Sustainability, Security and Spectrum. Available online: https://hexa-x.eu/d1-2-expanded-6g-vision-use-cases-and-societal-values-including-aspects-of-sustainability-security-and-spectrum/(accessed on 3 May 2021).
2. 5GAA Working Item MEC4AUTO. Technical Report Use Cases and Initial Test Specifications Review. Available online: https://5gaa.org/news/working-item-mec4auto/(accessed on 19 July 2021).
3. 5GAA Technical Report. Tele-Operated Driving (ToD): System Requirements Analysis and Architecture. Available online: https://5gaa.org/news/tele-operated-driving-tod-system-requirements-analysis-and-architecture/(accessed on 15 September 2021).
4. Ethics Guidelines for Trustworthy AI, Technical Report. European Commission. High Level Expert Group on AI. 2019. Available online: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai (accessed on 16 August 2022).
5. Barredo Arrieta, A.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [CrossRef]
6. Fernandez, A.; Herrera, F.; Cordon, O.; Jose del Jesus, M.; Marcelloni, F. Evolutionary Fuzzy Systems for Explainable Artificial Intelligence: Why, When, What for, and Where to? *IEEE Comput. Intell. Mag.* **2019**, *14*, 69–81. [CrossRef]
7. Scalas, M.; Giacinto, G. On the Role of Explainable Machine Learning for Secure Smart Vehicles. In Proceedings of the 2020 AEIT International Conference of Electrical and Electronic Technologies for Automotive (AEIT AUTOMOTIVE), Turin, Italy, 18–20 November 2020; pp. 1–6. [CrossRef]
8. 5GAA White Paper: Making 5G Proactive and Predictive for the Automotive Industry. White Paper. Available online: https://5gaa.org/news/5gaa-releases-white-paper-on-making-5g-proactive-and-predictive-for-the-automotive-industry/ (accessed on 8 January 2020).
9. Elbir, A.M.; Soner, B.; Coleri, S. Federated learning in vehicular networks. *arXiv* **2020**, arXiv:2006.01412.
10. Samarakoon, S.; Bennis, M.; Saad, W.; Debbah, M. Federated Learning for Ultra-Reliable Low-Latency V2V Communications. In Proceedings of the 2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 9–13 December 2018; pp. 1–7. [CrossRef]
11. Posner, J.; Tseng, L.; Aloqaily, M.; Jararweh, Y. Federated Learning in Vehicular Networks: Opportunities and Solutions. *IEEE Netw.* **2021**, *35*, 152–159. [CrossRef]
12. Salim, S.; Turnbull, B.; Moustafa, N. A Blockchain-Enabled Explainable Federated Learning for Securing Internet-of-Things-Based Social Media 3.0 Networks. *IEEE Trans. Comput. Soc. Syst.* **2021**, 1–17. [CrossRef]
13. Corcuera Bárcena, J.L.; Ducange, P.; Ercolani, A.; Marcelloni, F.; Renda, A. An Approach to Federated Learning of Explainable Fuzzy Regression Models. In Proceedings of the IEEE WCCI 2022 (World Congress on Computational Intelligence), Padua, Italy, 18–23 July 2022.
14. Takagi, T.; Sugeno, M. Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Syst. Man Cybern.* **1985**, *SMC-15*, 116–132. [CrossRef]
15. Bechini, A.; Corcuera Bárcena, J.L.; Ducange, P.; Marcelloni, F.; Renda, A. Increasing Accuracy and Explainability in Fuzzy Regression Trees: An Experimental Analysis. In Proceedings of the IEEE WCCI 2022 (World Congress on Computational Intelligence), Padua, Italy, 18–23 July 2022.

16. Tong, W.; Hussain, A.; Bo, W.X.; Maharjan, S. Artificial Intelligence for Vehicle-to-Everything: A Survey. *IEEE Access* **2019**, *7*, 10823–10843. [CrossRef]
17. Dong, L.; Sun, D.; Han, G.; Li, X.; Hu, Q.; Shu, L. Velocity-Free Localization of Autonomous Driverless Vehicles in Underground Intelligent Mines. *IEEE Trans. Veh. Technol.* **2020**, *69*, 9292–9303. [CrossRef]
18. Wu, Y.; Liao, S.; Liu, X.; Li, Z.; Lu, R. Deep Reinforcement Learning on Autonomous Driving Policy with Auxiliary Critic Network. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, 1–11. [CrossRef]
19. Peng, Z.; Gao, S.; Li, Z.; Xiao, B.; Qian, Y. Vehicle Safety Improvement through Deep Learning and Mobile Sensing. *IEEE Netw.* **2018**, *32*, 28–33. [CrossRef]
20. Zhan, J.; Ma, Z.; Zhang, L. Data-Driven Modeling and Distributed Predictive Control of Mixed Vehicle Platoons. *IEEE Trans. Intell. Veh.* **2022**, 1. [CrossRef]
21. Renda, A.; Ducange, P.; Gallo, G.; Marcelloni, F. XAI Models for Quality of Experience Prediction in Wireless Networks. In Proceedings of the 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Luxembourg, 11–14 July 2021; pp. 1–6. [CrossRef]
22. Nardini, G.; Sabella, D.; Stea, G.; Thakkar, P.; Virdis, A. Simu5G—An OMNeT++ Library for End-to-End Performance Evaluation of 5G Networks. *IEEE Access* **2020**, *8*, 181176–181191. [CrossRef]
23. Nardini, G.; Stea, G.; Virdis, A.; Sabella, D.; Thakkar, P. Using Simu5G as a Realtime Network Emulator to Test MEC Apps in an End-to-End 5G Testbed. In Proceedings of the 2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications, London, UK, 31 August–3 September 2020; pp. 1–7. [CrossRef]
24. Corcuera Bárcena, J.L.; Ducange, P.; Marcelloni, F.; Nardini, G.; Noferi, A.; Renda, A.; Stea, G.; Virdis, A. Towards Trustworthy AI for QoE prediction in B5G/6G Networks. In Proceedings of the First International Workshop on Artificial Intelligence in beyond 5G and 6G Wireless Networks (AI6G 2022), Padua, Italy, 18–23 July 2022.
25. Micheli, D.; Muratore, G.; Vannelli, A.; Scaloni, A.; Sgheiz, M.; Cirella, P. Rain Effect on 4G LTE In-Car Electromagnetic Propagation Analyzed Through MDT Radio Data Measurement Reported by Mobile Phones. *IEEE Trans. Antennas Propag.* **2021**, *69*, 8641–8651. [CrossRef]