IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received March 14, 2020, accepted March 31, 2020, date of publication April 8, 2020, date of current version April 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2986437

Cloud-Based Implementation of an Automatic Coverage Estimation Methodology for Self-Organising Network

DANIEL FERNANDES[1], **DIOGO CLEMENTE**[1], **GABRIELA SOARES**[2], **PEDRO SEBASTIÃO**[1], **(Member, IEEE)**, **FRANCISCO CERCAS**[1], **(Senior Member, IEEE)**, **RUI DINIS**[3], **(Senior Member, IEEE)**, **AND LÚCIO S. FERREIRA**[2,4,5], **(Senior Member, IEEE)**

[1]Instituto Universitário de Lisboa (ISCTE-IUL)/IT—Instituto de Telecomunicações, Av. das Forças Armadas, 1649-026 Lisbon, Portugal
[2]Multivision—Consultoria, Rua Soeiro Pereira Gomes, Lote N°1, 3°C, 1600-196 Lisbon, Portugal
[3]FCT—Universidade Nova de Lisboa, Monte da Caparica, 2829-516 Caparica, Portugal
[4]ISTEC, A. das Linhas de Torres 179, 1750-142 Lisboa, Portugal
[5]INESC—ID/COPELABS Lusófona University, Campo Grande, 376, 1749-024 Lisbon, Portugal

Corresponding author: Daniel Fernandes (dfsfs@iscte-iul.pt)

This work was supported in part by the OptiNET-5G project and co-funded by the Centro2020, Portugal2020, and European Union under Project 023304, in part by the Instituto de Telecomunicações, FCT/MCTES through national funds and co-funded by the EU funds under Project UIDB/EEA/50008/2020.

ABSTRACT One of the main concerns of telecommunications operators is related to network coverage. A weak coverage can lead to a performance decrease, not only in the user experience, when using the operators' services, such as multimedia streaming, but also in the overall Quality of Service. This paper presents a novel cloud-based framework of a semi-empirical propagation model that estimates the coverage in a precise way. The novelty of this model is that it is automatically calibrated by using drive test measurements, terrain morphology, buildings in the area, configurations of the network itself and key performance indicators, automatically extracted from the operator's network. Requirements and use cases are presented as motivations for this methodology. The results achieve an accuracy of about 5 dB, allowing operators to obtain accurate neighbour lists, optimise network planning and automate certain actions on the network by enabling the Self-Organising Network concept. The cloud implementation enables a fast and easy integration with other network management and monitoring tools, such as the Metric platform, optimising operators' resource usage recurring to elastic resources on-demand when needed. This implementation was integrated into the Metric platform, which is currently available to be used by several operators.

INDEX TERMS Cloud implementation, coverage estimation, drive tests, measurements, propagation model.

I. INTRODUCTION

Nowadays, there is an increasing demand of mobile users, which also increases the use of telecommunication services, making network coverage estimation a concern for operators. According with [1], in 2018 75% of world population was covered by 4G technology and with a predicted increase to 90% in 2025. In 2025 it is also expected that about 65% of world population will be covered by 5G technology. A correct estimation allows operators, not only the guarantee of a better network coverage delivery to its users, but also to perform an efficient optimisation of their resources. This estimation can be done by using several tools, from various vendors [2], [3], that are commercially available. These tools have important features that are very useful to telecommunication operators. However, most of these tools, like [4], require local installation on the operator's machine, and the network planning configuration tasks can be difficult and time-consuming.

To simplify the planning process of a network, this paper presents a novel cloud-based framework of a semi-empirical propagation that using Drive Test (DT) measurements, terrain morphology, buildings information and configurations of the network itself to estimate the coverage in a precise way. This optimised propagation model uses a cloud-based implementation, which allows its integration into a tool for

The associate editor coordinating the review of this manuscript and approving it for publication was Danping He.

monitoring and management of telecommunications networks, called Metric.

Several studies [5], [6] combine propagation models with DT measurements in order to define a realistic propagation model, each time the telecommunications operator wants. We observe that these studies use propagation models that are only applicable to specific macro or micro cells scenarios. In the literature we can also find adaptive systems capable of performing automatic selection of the propagation model according to the scenario (micro and macro cells) [7]. We observe that these different propagation models can only be used under specific conditions, also specifying antenna distance or height. The proposed propagation model intends to overcome these situations and to be generally applicable, regardless of the scenario or antenna properties, for example. In addition, the main novelty of this propagation model is that it is automatically calibrated with drive tests as well as network Key Performance Indicators (KPIs), providing a realistic estimation of path loss without the need for this calibration to be triggered by the user.

This work extends the ideas presented in [8]–[10], where the cell reach value and the handover distance were considered in the calibration of the model. These authors, namely in [9] and [10], include the terrain morphology but they omit the possibility of a Line of Sight (LoS) between the Mobile Terminal (MT) and the Base Station (BS). Despite the cloud implementation, using Amazon Web Services (AWS), has already been used in [10], this proposed work presents further enhancements that generalise and increase its precision, mostly related with fine-tuning of our propagation model. Several aspects, such as the reuse of previous DT measurements and the application of LoS between the Mobile Terminal (MT) and the Base Station (BS) were also considered. One of the novelties of this work is the inclusion of an evaluation metric that allows to estimate the performance of the proposed propagation model. A very important feature of our propagation model is that it fully automates network coverage estimation, while maintaining a simple user intervention, and all aspects of it are carefully detailed. This model can be used by telecommunication operators not only for coverage estimation based in new DT measurements, but also for cell and KPI configurations.

Proposed semi-empirical propagation model, named Automatically Calibrated SPM (ACSPM) presents several innovations. This model integrates various types of KPIs, network configurations, different propagation models, terrain morphology data (as accurate as possible) and the LoS between the Mobile Terminal (MT) and Base Station (BS). The grid is dynamically created. In this paper, the implementation with cloud services, the integration in the Metric platform and the various application for this work, like the network coverage estimation, the crossed sectors identification, the 2G, 3G and 4G planning and the traffic prediction, are presented.

The cloud implementation of the ACSPM enables the Metric platform to have several users to access resources simultaneously, such as the coverage of an antenna. It is also possible to execute multiple tasks at the same time, allowing the user to perform other actions while, e.g. the platform is computing a cell coverage. After the model is integrated into the Metric platform it is possible for telecommunications operators, not only to visualise the realistic area that a given cell covers, but also to optimise the entire network, namely in the establishment of neighbouring cells list.

The main contribution of this paper is the creation of a novel generalised and new semi-empirical propagation model that can be applied simultaneously in micro and macro cell scenarios. This propagation model introduces the innovation of being automatically calibrated with DTs as well as network KPIs, providing a realistic estimation of path loss each time data related to the antenna (tilt, azimuth, KPIs, DTs) is added or changed. The automation of the calibration process follows the Self-Organising Network (SON) paradigm, makes it possible to reduce the human effort, which results in a financial impact on the management of these networks. The overall accuracy achieved by using this ACSPM model, through the use of new measurements and the constant update and storage of results in the cloud, allows telecommunications operators to efficiently plan and optimise their 2G, 3G and 4G and 5G networks and to identify possible problems within their network configurations. The implementation of this methodology is based on cloud-services, efficiently providing elastic, on-demand and pay-per-use computation and storage resources. The resulting work pattern proves to be effective in the integration of various inputs and in the provision of realistic estimations of received signal levels around antennas, essential for network planning and optimisation.

The paper has the following structure: In Section II motivation for the use of cloud-based software as a service solutions, is presented, highlighting the tool Metric where this methodology is implemented and integrated. Key aspects related to the estimation of coverage in cellular networks are presented in Section III. In Section IV a unified coverage estimation methodology and use cases are presented, being the coverage estimation detailed in Section V. In Section VI, the implementation using AWS and its integration in Metric platform is presented. To test and evaluate the proposed model, Section VII presents a reference scenario and performance evaluation results are shown in Section VIII. Section IX highlights several applications where the proposed ACSPM is currently used and the conclusion of this paper is presented in Section X.

II. PLANNING AND OPTIMISATION TOOLS, PARADIGMS AND CLOUD-SERVICES

This section presents an overview of the tools, paradigms and services related or useful for the management of telecommunication networks.

There is currently a growing demand for services accessed through mobile devices, which represents a problem for telecommunications operators due to the excessive usage of network resources. This demand, which is in constant growing, makes mobile networks more complex and dense, resulting also in a more complex planning and optimisation requirements.

As each operator seeks to provide the best service for its customers, there is a major concern for proper planning and optimisation of mobile networks. Several vendors like Nokia [2] and Huawei [3], associated with the equipment they sell, provide planning and optimisation software. Each vendor has associated an Operations Support System (OSS), which congregates configuration parameters as well as KPIs. Still, these software products are proprietary, vendor-specific. Telecommunications operators must recur to these expensive tools from various vendors, not being able to have a unified vision of their network. Vendor-agnostic tools, like Atoll [4], installed on a local machine, enable the planning and optimisation of their network, independently of the vendor, as long as these follow the standards of each technology - Global System for Mobile Communications (GSM), Universal Mobile Telecommunication System (UMTS), or Long Term Evolution (LTE). These can be also associated to the OSS, although, this process requires additional configurations.

Besides the vendors' heterogeneity, telecommunications operators have heterogeneous networks as well. Within the same network, technologies such as 2G, 3G and 4G coexist and interact, each one with different specificities. The networks' densification, the increasing number of users, and the quality assurance demand for telecommunication operators force networks to continuously generate a huge amount of information per cell and per user. One way to monitor the network's performance is through the several available KPIs which results from a mathematical manipulation of several counters (component information). To increase the network's quality, it is possible to optimise the various cell's configuration parameters to reach the optimal configuration and the consequent optimal performance, required by operators.

Once there is a considerable amount of information per cell and per user through time, the generated data can be analysed through a big data approach. The concept of big data considers the volume and the variety of data, the value generated through analysis, the velocity at which the data is generated and processed. An example of big data application is to acquire dynamic data from different sources, in different formats [11]. Big data methods have been used in recent studies to implement Self-Organising Network [12]–[14]. SON aims to reduce operational costs, investment costs and offers a better Quality of Service (QoS) to users [15]. The description of some use cases related to SON is present in [16] and some of these implementations are presented below:

- Self-Planning - When a node is added to the network, the site location and hardware settings are selected;
- Self-Deployment - This implementation receives the data from Self-Planning and performs the installation and validation of the node;
- Self-Optimisation - Using data obtained from network users, such as MT, the network settings are automatically adjusted;
- Self-Healing - A set of preventive actions keeps the network operational and prevents disruptive network problems from arising.

These four use cases can associate the use of SON networks as networks with the "plug & play" feature. This feature can be useful for telecommunication operators, since it can simplify the implementation and maintenance of the network. Whenever an incident is detected, the network can minimise its effect until a permanent resolution decision is made.

With the advance of SON networks, the volume of data useful for the correct operation of the network has increased, and this data may no longer require human intervention in its handling. Ideally, the information should be collected and automatically made available to the SON algorithms [17]. In this paper, in order for this automation to be easily implemented, all the adjacent computing processes were migrated to a cloud-based implementation. This migration presents the advantages of a cloud-based implementation being very scalable and ready for big data manipulation. This implementation provide users with storage and processing services.

Cloud services are a solution based on a flexible, scalable and abstract infrastructure for the user. One of the main advantages of using cloud-based services is that the user can explore all the available resources and is only charged on the applications used in a per-second price basis, which corresponds to the effective running time of processing nodes. Another advantage of these services is that they are always available and accessible through any device connected to Internet. There are many public cloud providers such as Amazon Web Services (AWS) [18], Google Cloud [19], Microsoft Azure [20].

However, in order to be able to migrate the management of the telecommunications network to the cloud, it is necessary to make some changes to the core of the network. The functions performed by the Radio Access Network (RAN), namely the management of base stations, are now performed in a centralised component with a cloud-based architecture, called C-RAN or RAN-as-a-Service (RANaaS). This change allows for Network Functions Virtualisation (NFV) and scalability that was previously non-existent [21]. With C-RAN, the management of telecommunications networks can now move to a cloud environment and thus the concept of SON can be applied.

A Software-as-a-Service (SaaS) is a web-based application that does not require any local installation on any computer or server. This avoids machine limitations and constraints, and other operating system related restrictions. A SON solution for telecommunications networks, Metric [22], was developed by Multivision [23]. It is a web-based SaaS application that can be accessed by any browser on any device, enabling the monitoring and maintenance of a network, allowing SON implementation and quick access to information on the current status of the network. Being one of the main objectives the aggregation of data in a single platform, Metric platform is based on a set of modules. These modules allow the upload of data, regardless of the hardware vendors or network data source. A subsequent manipulation of the data allows them to be visualised on a map, in the form of a table or even in the form of a report. This analysis allows critical situations to be immediately identified in the network, namely crossed feeders, cell in overshooting, missing neighbours and consistency checks. Finally, this platform allows the scheduling of network validation tasks that, depending on the input data, can trigger correction/improvement actions in the network.

III. ESTIMATION OF COVERAGE IN CELLULAR NETWORKS

The estimation of coverage is of key importance in cellular networks planning. These wireless networks must rely in realistic predictions of coverage of their antennas to efficiently provide service within specific geographical areas, taking into account service requirements and offered traffic load predictions. Several methodologies and procedures are available to estimate coverage in cellular networks and are presented and detailed in this section.

A. WIRELESS COVERAGE

Wireless coverage allows MTs to access services and features provided by BSs cells. The aggregation and control of these BS is the responsibility of the RAN. The coverage area of a cell depends on several parameters. Among these parameters are those that defines the transmitted signal such as the position of the antenna, its radiation diagram, the azimuth, the tilt, and the transmission power. The signal received by the MT is the result of the transmitted signal by BS and the path loss attenuation. The location vector **p**, identifying a position with its latitude, longitude and height, the received signal level from antenna **a** operating at frequency f is given by [24]:

$$P_{rx[\mathrm{dBm}]}(f, \mathbf{a}, \mathbf{p}) = P_{tx[\mathrm{dBm}]}(\mathbf{a}) + G_{tx}(\varphi_p - \varphi_a, \theta_p - \theta_a)_{[\mathrm{dB}]} - L(f, \mathbf{a}, \mathbf{p})_{[\mathrm{dB}]}, \quad (1)$$

where,

- P_{tx} [dBm]: Transmitted power;
- G_{tx} [dB]: Transmitted antenna gain, considering the antenna's vertical and horizontal diagrams, for the vertical θ and horizontal φ direction between antenna location vector **a** (antenna latitude, longitude and height, h_a), and position **p**, using the deviation of the antenna azimuth φ_a and tilt θ_a;
- L [dB]: Path loss attenuation between **a** and **p** (x, y, z) positions.

In order to simply the process, it is considered that the gain of the receiving antenna is 0 dB.

For an antenna located at vector location **a** its coverage area C_a is given by the set of points of the service area A that satisfy the condition:

$$C_a = \{\mathbf{p} \in A : P_{rx\,[\mathrm{dBm}]}(f, \mathbf{a}, \mathbf{p}) \geq P_{rx_{min}\,[\mathrm{dBm}]}\}, \quad (2)$$

where $P_{rx_{min}}$ is the system's sensitivity, i.e., the minimum received signal level required to establish a connection with a terminal.

In order for telecommunication operators to provide their services to customers, a preliminary study is carried out on the locations where the antennas should be placed to obtain a large coverage area. These antennas can be divided according to their function, having macro-cells that are responsible for the coverage, the micro-cells responsible for the capacity and the femto-cells that provide the capacity to specific zones.

A weak cellular planning can lead to gaps in coverage and the proper provision of services to users is impossible. It is therefore essential to correctly estimate coverage in order to allow a correct configuration of parameters such as identification of neighbouring cells and identification of overshooting or crossed sectors problems.

B. PROPAGATION MODELS

A propagation model attempts to represent how an electromagnetic signal propagates, and has two types of application [25]:

1) Plan, design, test and validate a wireless system;
2) Optimisation of wireless systems. Operators use these models, which represent reality accurately as possible, to simulate the impact in the network when certain parameters are modified.

According to [26], [27], propagation models can be classified according to the type of information used. They can be classified as:

- *Empirical*: Empirical models are created through field measurements, while not considering the terrain information. They represent a particular propagation environment. This has the advantage that these models are simple and efficient to use. However, it has the disadvantage that it cannot be used in another environment without being updated.
- *Semi-Empirical*: While empirical models are specific to a particular propagation environment, semi-empirical models use terrain information such as the elevation and height of buildings, which makes these models much more specific to a particular region.
- *Deterministic*: Deterministic models combine the laws of electromagnetic wave propagation with terrain information. These models present more realistic results but have strong computational requirements.

In general, telecommunication operators use Empirical models for the design and comparison of systems and the Semi-Empirical and Deterministic models for network planning and deployment. Some of the most well-known propagation models are empirical models such as Okumura-Hata [28], Mishra [29] and Mishra [29] applied to macro and micro-cells, respectively. There are other models that extend these like Standard Propagation Model (SPM) [4].

In order to make propagation models more realistic, it is possible to calibrate them using real data. These calibrations can be performed through Artificial Neural Networks (ANNs) and they are strongly discussed in [30]–[32]. Despite the use of ANNs, the limitations of the propagation models still exist. For example, on an Okumura-Hata implementation (using ANNs or not), can only be applied at distances greater than 1 km, for a given frequency and

TABLE 1. Conditions under which the models are applicable.

Parameter	Range	
	Walfish-Ikegami	SPM
Frequency, f [MHz]	$800 - 2000$	$150 - 2000$
Distance, d_{BS} [km]	$0.02 - 5$	$1 - 20$
BS height, h_{BS} [m]	$4 - 50$	$30 - 200$
MS height, h_{MS} [m]	$1 - 3$	$1 - 10$

TABLE 2. Typical values for the SPM coefficient parameters.

	Minimum	Typical	Maximum
K_2	20	44.9	70
K_3	-20	5.83	20
K_4	0	0.5	0.8
K_5	-10	-6.55	0
K_6	-1	0	0
K_7	-10	0	0

for a given propagation environment. For heterogeneous networks, where several technologies working simultaneously at different frequencies, the calibrated models cannot be used. In order to reduce the limitation presented above, a calibrated propagation model was presented in [33], whose main disadvantage is that it requires the type of environment and its specifications in the calibration process.

In order to generate a propagation model that can be used in heterogeneous networks regardless of the type of propagation environment, this research uses the Walfish-Ikegami and SPM propagation models, used in micro cell and macro cell scenarios, respectively.

One propagation model typically used in micro cell scenarios is the Walfish-Ikegami model [29]. Table 1 presents the conditions under which it is applicable.

For this research, in the distance where the model is applied, considering LoS, its path loss is given by:

$$L_{Walfish_{[\text{dB}]}} = 42.6 + 26\log(d_{BS_{[\text{km}]}}) + 20\log(f_{[\text{MHz}]}), \quad (3)$$

where d_{BS} is the distance between BS and the MT and f is the frequency.

A macro cell model is the SPM [4]. The SPM is an extension of the Okumura-Hata propagation model and its extension COST 231-Hata. The validity intervals, Table 1, must be observed.

The SPM [4] is an extension of the Okumura-Hata propagation model and its extension COST 231-Hata, being a macro cell model. The validity intervals, Table 1, must be observed. The path loss calculation of the SPM model is given by:

$$\begin{aligned} L_{SPM_{[\text{dB}]}} = {} & K_1 + K_2\log(d_{BS_{[\text{m}]}}) + K_3\log(h_{BS_{[\text{m}]}}) \\ & + K_4 L_{dif} + K_5\log(d_{BS_{[\text{m}]}})\log(h_{BS_{[\text{m}]}}) \\ & + K_6(h_{MT_{[\text{m}]}}) + K_7\log(h_{MT_{[\text{m}]}}) + K_{clutter} f(clutter), \end{aligned} \quad (4)$$

where,

- K_1 [dB]: Constant offset;
- K_2: Coefficient for $\log(d_{BS})$;
- d_{BS} [m]: Distance between BS and the MT;
- K_3: Coefficient for $\log(h_{BS})$;
- h_{BS} [m]: Height of BS;
- K_4: Coefficient for diffraction calculation;
- L_{dif} [dB]: Loss due to diffraction over an obstructed path;
- K_5: Coefficient for $\log(d_{BS})\log(h_{BS})$;
- K_6: Coefficient for h_{MT};
- h_{MT} [m]: Height of MT;
- K_7: Coefficient for $\log(h_{Mt})$;
- $K_{clutter}$: Coefficient for $K_{clutter}$;
- $f(clutter)$: Average of weighted losses due to clutter;

The coefficients parameters explained above are limited to the presented in Table 2. Although the frequency of the antenna and the attenuation coefficient is not present in 8, this is considered in the K_1 factor.

Although this model is based from the Okumura-Hata model, the SPM allows values of elevation, diffraction loss, among other parameters in the calculation of path loss, making the model more accurate and realistic.

C. COVERAGE MONITORING

In order to enable telecommunications operators to collect data with network quality information, some field measurements, called Drive Tests (DT), are made. As a rule, these DTs are collected using a vehicle with radio equipment that travels along a predefined route, collecting data from a certain area and each geolocated measuring point. The analysis of the DTs allows operators to understand network failures, such as areas with weak coverage, making it possible for operators to perform corrective measures (like the placement of a new base station). The reception signal suffers some effects like shadowing, slow and fast fading [25]. When working with discrete areas, each one of the sub-areas considered, the signal suffers from fast fading. The use of DTs measurements allows the neutralisation of this effect.

Despite the precise information revealed by the DTs, their execution can be expensive and requires allocation of human resources. The information collected only refers to areas with roads, which limits the true perception of the entire network.

In order to make use of DTs data more efficiently, a standardisation was proposed. This would reduce the costs associated with their execution, allowing them to be made by a MT. This standardisation is called Minimisation of Drive Tests (MDT) and in addition to the objectives mentioned above for DTs, these are triggered whenever there is the implementation of a new base station, the construction of new roads or buildings and if there are complaints from customers [34]. MDT is available in the vast majority of devices, and it is operators' decision to activate this functionality. For the user, the only drawback is the additional power consumption. However, and if the measurements are geolocated, it is possible to create a realistic view of the signal at various points in a cell. By using this data in the proposed model, signal estimation over an entire cell area will become more precise and accurate.

D. AVAILABLE NETWORK PERFORMANCE INDICATORS RELATED TO COVERAGE

In the OSS are available, for the operators, performance indicators, KPIs. These indicators, such as Cell Reach can help in estimating coverage. When an MT connects to a BS, the BS estimates the propagation delay of the communication and sends this value to MT allowing for adjustment in the transmission time ensuring that all communications from different MTs in different places and at different times arrive at BS at the same time. This setting in the transmissions allows the minimisation of interference in the uplink signals of MTs.

Depending on the technology used, the propagation delay presents different denominations. For 2G is referred to the timing advance indicator and for 3G and 4G the propagation delay indicator. These indicators are discrete values defined by each of the vendors. After the information is collected, it is stored and made available in the OSS in the form of counters, which registers the number of occurrences within a given time period [35], allowing an understanding overview of the geographical distribution of MTs within rings of specific ranges.

The propagation delay can be used to estimate the distance between the BS and MT. Considering that an electromagnetic signal propagates at the speed of light, c, and the propagation delays, t_{prop}, the distance is given by:

$$d_{prop_{[\mathrm{m}]}} = c \times t_{prop_{[\mathrm{s}]}} \tag{5}$$

When network planning is performed, telecommunications operators combine some configuration parameters in order to obtain a certain cell size or a planned cell size. This size can be considered as the maximum range of the cell where communications can occur.

IV. NOVEL CELLS COVERAGE ESTIMATION WORK PATTERN

Several methodologies currently used by operators related to cell coverage estimation were presented in the former section. The present work aims to propose a unified methodology to integrate them all and reach a realistic estimation of coverage for any antenna, based on a varied set of information: terrain information, antenna configurations, propagation models, DT measurements, and KPIs. To present it, first a motivation is presented, highlighting the identified needs as requirements for a new methodology. Then, a small use case that exemplifies the way the proposed methodology can be used. Finally, the overall architectural framework is presented, its components being detailed in later sections.

A. MOTIVATION AND REQUIREMENTS

In this section, a motivation and requirements for a unified coverage estimation methodology are presented, based on current practices of operators.

Within the planning and optimisation activities of a telecommunication operator, a key task is the estimation of the coverage area of the cells of its network. For example, given a cell of, e.g., expected 7 km range, the receive power level within the surrounding geographical area of the antenna must be estimated in order to evaluate how it fits with the neighbouring cells. For it, considering the currently used models, the operator will need to apply two distinct propagation models, as the range of micro-cell models (Walfish-Ikegami) only applies up to 5 km, while a macro cell model (SPM) only applies from 1 km on. In this sense, a unified model capable of predicting coverage would be very useful for operators.

To achieve a more realistic estimation of the signal level, the operator recurs to DTs. Still, the criteria to determine the specific areas for the measurement campaign is many times random, which may question from the start the usability of such campaigns. These measurements, once collected, are manually processed by the engineers following arbitrary procedures. They aim at manually fitting certain parameters of the propagation models in order to adapt the model's prediction to the real measurements. This curve fitting is time consuming and erroneous, as statistical indicators of DT measurements (average, standard deviation of measurements) are used, still not taking into account the specific geographic location of them. In fact, propagation conditions of a geographical area where DTs are done may not be representative for another area. Some operators tend to regularly do extensive and expensive DT campaigns to evaluate the quality of the network. Still, they do not have tools to process this large amounts of data nor integrate them in their planning and optimisation tools. In this sense, a coverage prediction model automatically calibrated with available DT measurements would be of key value. A model that would take into consideration the geographical position of each DT, using these geolocated measurements to weight a propagation model to build a complete map of estimated coverage (to estimate coverage in a given position, DTs that would be nearby, in similar propagation conditions, would have more weight than farther DTs). This model would be able to easily integrate MDT measurements, enabling to build a realistic picture of the received signal level around any antenna. This would require an efficient infrastructure for the storage and processing of DT measurements, and computation of propagation grids for each antenna. We think that the use of cloud-services enables an elastic and efficient storage and processing of resources on-demand.

The operator also has available several KPIs. In particular, as formerly discussed, the availability of propagation delay measurements for every connection of each cell is of precious value, enabling to characterise the spatial statistical distribution of users using the cell. In particular, it enables to determine the cell reach. Combined with signal level configurations that trigger handover procedures, these are important indicator that should be integrated in a realistic coverage estimation model of a given cell.

On the other side, one of the challenges for an operator is the variety of infrastructure providers (Nokia, Huawei, Ericsson) and technologies (2G, 3G, 4G, 5G) that compose their network. Each having its specific OSS, with specific KPIs. Manufacturers tend in fact to have different ways of computing KPIs, resulting in incompatible indicators, if we want to have a common view. In this sense, a unified view of KPIs and configuration parameters is essential. It shall be available in a single platform. Ideally, cloud-based, enabling to be available anywhere, and recurring to elastic and on-demand computation and storage resources.

B. USE CASES

Three use cases are presented below to illustrate the advantages of the proposed methodology, illustrated in Fig. 1.

Consider an operator of a large cellular network with thousands of cells of various technologies and vendors. He has available detailed terrain elevation and building data, as well as a large variety of DT measurement campaigns of its cells. He has also available, from each manufacturer's OSS, information of antennas' configuration as well as numerous KPIs, as represented in Fig. 1 as inputs. The proposed methodology is capable of integrating transparently all this data and use it to calibrate automatically a propagation model capable of building with great precision, for each cell, geo-referenced attenuation or P_{rx} grids, as illustrated in Fig. 1. Each time new DT measurements are uploaded, or when configuration or KPI changes occur, the computation of new grids is automatically triggered thanks to a work pattern based on on-demand elastic cloud processing and storage resources. In this way, the operator has always available updated information of its network, without the need of human intervention.

A simple use case is when an operator receives a report specifying that a given geographic area has coverage problems. The available grids of antennas from the area of interest can be easily combined in a heatmap to identify with precision the zones with poor signal.

Another use case is related to a typical situation where the increase of traffic in a site requires the swap from a single omni-directional antenna to a tri-sectorized site. For it, each antenna must be configured (P_{tx}, azimuth and tilt). Several DT measurements are available for the original omni-directional antenna, as well as terrain data. The proposed methodology is capable of creating a realistic P_{rx} grid for each antenna, and build heatmaps of coverage for various antenna configurations. This large amount of data can be easily processed thanks to cloud on-demand processing resources. Based on these, the operator can manually choose, or use an automatic methodology to select the best configuration (taking, e.g., into account neighbouring cells).

Finally, consider an operator that has activated MDT, where mobile terminals report to the OSS geo-located extensive measurements of P_{rx} levels from neighbouring antennas. Although representing a very large amount of data, the proposed cloud-based methodology can automatically calibrate a unified propagation model (for distances from tenths of meters to tenths of kilometers) with great precision, recurring to elastic cloud-based processing resources. This results in realistic heatmaps of received signal levels for each antenna, enabling to easily identify regions with interference as well as poorly covered ones.

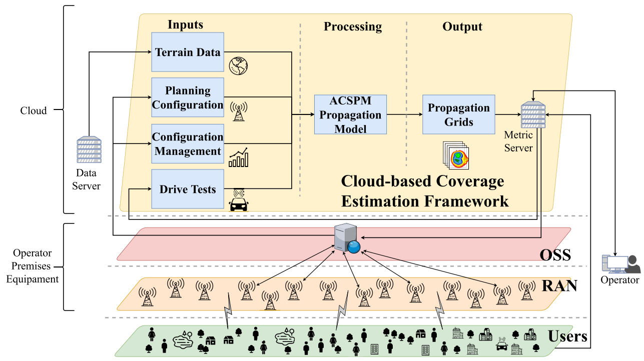

FIGURE 1. A new cell coverage estimation architectural framework composed by four layers for the cloud-based management of radio resources.

All these use cases illustrate the proposed propagation model and supporting methodology, detailed in the following sections.

C. OVERVIEW

A novel cloud-based cell coverage estimation architectural framework is proposed, as presented in Fig. 1. It represents a vision of a cellular network composed by four layers, which clearly identifies what is done and implemented in the cloud and also what are the resources used at the telecommunications operator premises.

The bottom Users layer represents the cellular network users equipped MTs communication devices. These devices use services provided by the RAN layer, generating traffic that is exchanged between these two layers. The RAN layer is composed by all the BSs. BSs are responsible for covering a certain service area (called cell) with wireless connectivity in order for users to establish connection to services.

One of the most important layers for our research is OSS. This layer congregates monitoring counters, performance data of each of the elements, and allows to manage the configuration of these elements through various parameters available. From the OSS, the proposed framework Extracts, Transforms and Loads (ETL) specific data that is used in the coverage estimation model.

The Automatic Coverage Estimation Framework is divided in three main modules: Inputs, Processing and Output. Each of these modules contains one or more submodules.

The Inputs module uses as input data from the OSS, the Configuration and Performance Management, terrain morphology and DTs values that comes directly from the Metric server. The DT data is loaded directly by the user on the Metric server.

Once all of the input data is obtained, the processing of the all information is performed by the novel propagation processing module, which is the most relevant element in the proposed framework, namely the generation of the Propagation Model. In the Output module, for a given cell the received signal power grid is stored on the Metric server and becomes available to the user.

In the following section, the various components of the processing module represented in the architectural framework are detailed.

V. MODULES OF THE AUTOMATIC COVERAGE ESTIMATION FRAMEWORK

In this section, a semi-empirical propagation model that accurately represents reality is detailed. It details and largely extends initial ideas drafted over [8]–[10].

A. INPUTS MODULE

This subsection describes in detail each of the Input submodules of the architectural framework depicted in Fig. 1.

The ***Drive Tests submodule*** is an important input in the presented architecture. DTs in this context are used to calibrate the propagation model to a particular geographical area allowing the propagation model to accurately represent reality. DTs data comes directly from the Metric server and is made available to an antenna at a given time for certain antenna characteristics over a given time period. The DT for a given position provides the receiving power, being represented as P_{rxDT} [dBm].

There are configurable network parameters, which are received through the OSS and serve to calibrate the model. These parameters are aggregated in the ***Configuration Management module***. Some of these parameters are related to the antenna such as its model, the height (h_{BS}), the mechanical and electrical tilt ($\theta_{mec_{BS}}$ and $\theta_{ele_{BS}}$), the azimuth (φ_{BS}), the position and the transmitting power (P_{tx}). Other parameters are related to the technology, frequency, the MT sensitivity ($P_{rx_{min}}$) and the geohash length, Geo_{hash}. The geohash is a encoder that transforms coordinates in a unique hash, where each length of geohash is associated with a certain cell size [36]. Due to the projection systems considered, the geohash has different cell size values depending on the location on the globe. After choosing the length of the geohash and depending on the antenna location, pixel size is defined which is used in the grid construction. For example, if an antenna is located near the Equator line, the cell size is smaller in terms of height when compared to an antenna located in northern Europe.

This information is essential for describing the characteristics of the BS. The handling of this data needs special attention since it is manually entered by the user. This may cause problems if it is not entered correctly.

The ***Planning Management submodule*** corresponds to counters collected by the OSS which monitors the operation of various aspects. An example of this data type is the cell reach which allows obtaining the Prx_{ho} value.

The ***Terrain Data submodule*** is divided into two steps. Initially, information on the terrain morphology of a given work area is requested through the Application Programming Interface (API) provided by *OpenTopography* [37]–[39]. This API returns elevation information for the considered area in raster format 30m by 30m pixels. This information is then integrated into the architecture displayed according to the size of the pixel, defined by the user in the grid generation. For each of the pixels created, the elevation value is searched in the API result. The next step is to add information about buildings and large structures that are requested through the Overpass API [40]. This API returns information from *OpenStreetMap* [41], where data is available under the Open Database License (ODbL), and for a given workspace, it returns the height information of the existing buildings. For each of the pixels of the generated grid, the average of the heights of the buildings in that pixel is calculated. These two steps allow to obtain data similar to Light Detection and Ranging (LiDAR) data and thus realistically portrays the propagation environment.

Table 3 shows some of the parameters used in the propagation model.

TABLE 3. Parameters used in the propagation model.

	Parameter	Description
Antenna	h_{BS}	Height of BS [m].
	$G(p)$	Antenna gain for pixel p [dB].
	$\theta_{mec_{BS}}$	Antenna mechanical tilt [º].
	$\theta_{ele_{BS}}$	Antenna electrical tilt [º].
	φ_{BS}	Antenna azimuth [º].
	P_{tx}	Transmitting power [dBm].
Network	P_{rxDT}	Receiving power for a DT [dBm].
	$P_{rx_{min}}$	Minimum reception power to be considered [dBm].
	Prx_{ho}	Handover power [dBm].
	d_{ho}	Handover distance [m].
	$L(d_{ho})$	Loss at handover distance [dB].
	Geo_{hash}	Length of the geohash, when combined with antenna positions define the pixel size.
	d_{wal}	Minimum distance for application of the Walfish-Ikegami propagation model [m].
	$L(d_{wal})$	Loss at d_{wal} [dB].
Model	K_1	SPM offset.
	K_2	SPM for $\log(d_{BS})$ [dB].
	L_{dif}	Diffraction loss due an obstructed path [dB].
	h_{MT}	Height of MT [m].
	d_{BS}	Distance between BS and the MT [m].

B. PROCESSING MODULE

The core module of our framework is the novel ***Propagation Model*** implemented in the cloud where, for a given cell, based on available inputs, it automatically estimates the cell's path loss within an area of interest. It results from the combination of various aspects. In Fig. 2 it is possible to verify the basic idea of the proposed propagation model, ACSPM. The ACSPM is calibrated with the Walfish-Ikegami model, power level at the handover distance. This calibration allows adjusting SPM to a distance of less than 1 km and adjust to the reality of the antenna, using not only the handover information, but also the DT information. In Fig. 2, d_{max} corresponds to the maximum distance returned by the

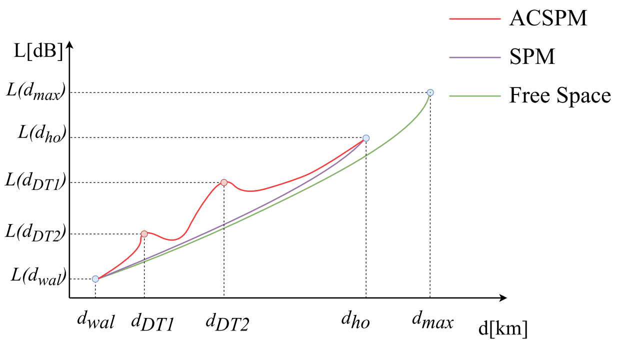

FIGURE 2. Illustration of the combination of various aspects for the construction of the ACSPM.

free space attenuation expression for a given power considered. The remaining variables are defined in Table 3.

The needed inputs to obtain the propagation loss grid are listed in Table 3. Many of these are obtained along the steps detailed in Section V-A. Taking as input the parameters discussed in Section V-A and presented in the Table 3, the calibration of the propagation model and computation of propagation loss grids follows the methodology illustrated in Fig. 3, which is detailed furthermore.

1) DIMENSIONING THE GRID

The proposed methodology aims to estimate propagation loss around a given cell. For that estimation, a rectangular area of interest around the cell area must be identified. Within this area, a grid of pixels is defined. For each pixel the path loss from the antenna shall be estimated.

The maximum communication range, d_{max}, is computed for the antenna. Using the free space propagation model, it corresponds to the distance where, for the antenna's P_{tx} and G_{tx}, the received power reaches the MT sensitivity, $P_{rx_{min}}$. This range will define the initial size of the grid, centred in the antenna position. The grid dimensions are then reduced following a binary search iterative procedure for each of the borders of the grid. In each step, once there is no terrain elevation information, the SPM model, with no diffraction is used to evaluate P_{rx} along the border, determining if, in the next step of the binary search, one should search nearer or farther from the current border. If for a given position the received power is equal to a minimum power $P_{rx_{min}}$ it becomes one of the grid limits. This procedure is applied to each border of the grid.

Once the dimension of the grid of the cell is defined, it is filled with geolocated pixels. The size of the pixels is specified by the Geo_{hash}. Each of the geolocated pixel in the grid will be used to store various information useful to the model and specific for its position. This information is obtained throughout the process illustrated in Fig. 3 and is detailed in the next sections. For each pixel, the following information shall be obtained:

- Terrain height information, h_{MT}, of the pixel and the distance to the BS, d_{BS};
- Existence of LoS, expressed as diffraction losses, L_{dif};
- Gain introduced by the antenna in that pixel, $G(p)$;
- Power information received from DTs, P_{rxDT};
- Specific SPM calibration parameters, K_1 and K_2;
- Estimated path loss value for that pixel.

2) INCLUSION OF TERRAIN MORPHOLOGY INFORMATION

The terrain morphology and the building heights are added to each pixel. In order to validate if line of sight exists between the pixel and the BS, the Bresenham's line algorithm [42] is initially used to find all pixels, in a straight line, between the pixel and the pixel where the BS is located, depicted in Fig. 3.

3) COMPUTATION OF DIFFRACTION LOSS FOR EACH PIXEL

The Knife-Edge Diffraction model [43], [44] is used for each pixel of the line between the antenna and the pixel, to estimate the diffraction losses in that path. For each pixel in the path, a parameter ν is calculated according to the following equation:

$$\nu = h\sqrt{\frac{2}{\lambda}\left(\frac{1}{d_1}+\frac{1}{d_2}\right)} \tag{6}$$

where h is the obstruction height, i.e., height from the obstacle top to the Tx-Rx axis (straight line between transmitter and receiver antennas), λ is the wavelength and d_1 and d_2 is the distance between each side of the path and the obstacle. After calculating the parameter ν for each pixel of the line, the pixel with the highest value of ν is the pixel with the highest diffraction loss. So if the value of $\nu < -0.78$, $L_{dif} = 0$ dB

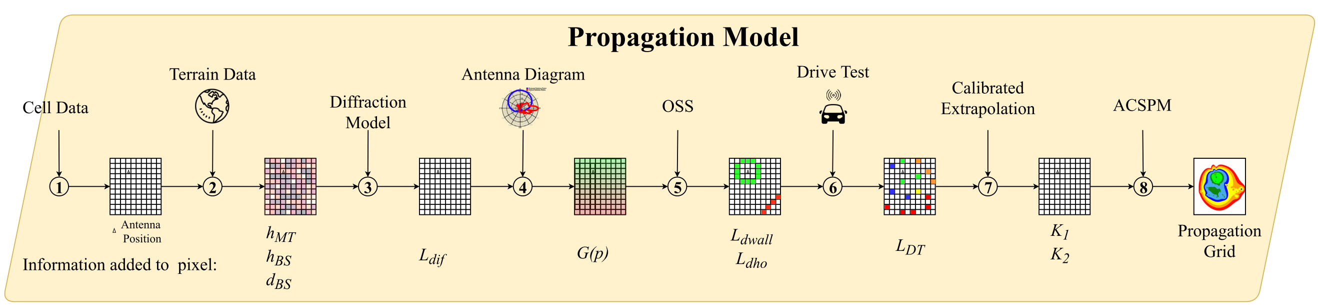

FIGURE 3. Steps for the computation of a propagation grid.

else the L_{dif} is given by [43]:

$$L_{dif} = 6.9 + 20\log\left(\sqrt{(\nu-0.1)^2+1}+\nu-0.1\right) \quad (7)$$

For each pixel, the L_{dif} value is stored in the grid, depicted in Fig. 3, and this value is considered into the ACSPM algorithm.

4) INCLUSION OF BS ANTENNA GAIN FOR EACH PIXEL

According to Fig. 3, other information that is stored in each pixel is the antenna gain perceived at the position of each pixel, $G_{tx}(\varphi_p - \varphi_{BS}, \theta_p - \theta_{mec_{BS}})$ according to its radiation pattern and its azimuth. This information is relevant to make the estimation of path loss independent of the characteristics of the antennas, when DT measurements are available and used to estimate the path loss for that pixel. In fact, by removing to P_{rxDT} the P_{tx} and G_{tx}, a path loss, $L(dt)$, independent of the antenna configuration is obtained from (1).

5) ESTIMATION OF PATH LOSS IN PIXELS OF CELL EDGES

By knowing the antenna's operating frequency, from OSS information it is possible to estimate, using (3), the value of path loss, $L(d_{wal})$, for distance, d_{wal} from the antenna, according to Fig. 3.

In a scenario where the cells are placed side by side, in order to cover a large area, the cell reach may be associated with the handover power, (Prx_{ho}). Prx_{ho} is a parameter defined by the operator in the planning phase of a network, where is considered that handover shall occur and it is related to the distance between cells. It is a minimum received power that the network shall guarantee in the service area. A handover to a new cell is triggered when signals are below the Prx_{ho}. The cell reach information identifies distance ranges within which communicating MTs are monitored/recorded by the OSS. The last ring represents the farthest area where MTs are connected to the BS, where $Prx = Prx_{ho}$.

Combining this two aspects, it is possible to calculate at a distance, d_{ho}, the path loss, $L(d_{ho})$, using (1) and the Prx_{ho} value. For each pixel that is at d_{ho} distance from the antenna, the value of $L(d_{ho})$ is considered.

6) ESTIMATION OF PATH LOSS IN PIXELS WITH DTs

The next step, according Fig. 3, is to register P_{rx} from DT measurements in the corresponding pixels. If, for a given pixel, there is more than one DT, then the value considered is the mean value of DTs in that pixel. The DT values considered are obtained from the Metric platform and are used to calibrate the propagation model.

Once the DTs are registered in the corresponding pixels, it is possible to compute and store for each of these pixels the corresponding perceived path loss, through (1), removing the antenna gain for that pixel as well as the P_{tx} of the antenna.

7) CALIBRATION OF PROPAGATION PARAMETERS OF EACH PIXEL

Initially, the K_2 parameter is estimated for all pixels, using the information, $L(d_{wal})$ and $L(d_{ho})$, calculated in V-B.5. Then, once path loss is registered in pixels with DTs or in cell edges, values of K_1 SPM parameter are estimated as follows:

- **Pixels with estimated path loss**: For each of these pixels, (4) is used to extrapolate values of K_1 parameters, taking for the remaining parameters of the equation the values presented in Table 2. If a pixel already has a value for K_1 (e.g., previous DT), then the average is taken between the two.
- **Pixels without estimated path loss**: Equation (8) is used to calculate the value of $K_1(p)$, for a pixel p without estimated path loss:

$$K_1(p) = \frac{\sum_{m=1}^{N} K_1(m)\alpha(p,m)}{\sum_{m=1}^{N} \alpha(p,m)}, \quad (8)$$

weighted by the N pixels with estimated path loss, $K_1(m)$. The parameter $\alpha(p, m)$ uses the distance between pixels p and m, given by

$$\alpha(p,m) = \frac{1}{[d(p,m)]^2}. \quad (9)$$

If the antenna has been characterised before, a file containing a list of all DTs previously converted to K_1 values is available. These values are distributed over the grid pixels the same way DTs are distributed.

8) ESTIMATION OF PATH LOSS FOR EVERY PIXEL

After calculating a value of K_2 and a value of K_1 for each pixel, it is possible to obtain the path loss value for each pixel using (4) and the remaining parameters presented on Table 2.

C. OUTPUT MODULE

After the processing task, some output files are generated with diversified information that can have many uses.

- *Path Loss File*: This JSON file is created directly from the antenna grid information without further processing. This file maps the path loss value to each of the grid pixel coordinates. The created text file is later represented in the Metric platform.
- *Receive Power File*: From the information present in the antenna grid, it is possible to convert the path loss values of each pixel into received power values, using (1) and the antenna gain values for that same pixel. This information is stored in this JSON file and can later be represented in the Metric platform.
- *K_1 File*: This JSON file stores the information regarding the used DTs. For each DT, the geographical location and the calculated K_1 values are stored. In this way, information related to the DT is stored, regardless of the

characteristics of the scenario under study. This information can later be reused for the same antenna.

D. EVALUATION METRICS

To evaluate the ACSPM, a common metric used in estimation models was applied. This metric allows not only to quantify the accuracy of the model but also to validate the improvement in signal estimation.

One way to visualise the performance of the model is to make a direct comparison between the SPM estimated values and the proposed model values. This analysis can be done visually by representing the coverage estimation by each model and comparing it or, for a more precise comparison, for a location where there are DT values.

In order to evaluate the proposed model estimation against the SPM model, a performance metric can be used to calculate the error between real values, y_i, and estimated ones, $\hat{y}_i$. The Mean Absolute Error (MAE) metric was applied in both models, using (10) [45].

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|, \tag{10}$$

where N refers to the number of considered values.

To calculate the accuracy of the model, for a certain point where DTs exist, DT value is ignored at the time of model calibration and its value is then compared with the value estimated by the model, by using the absolute error. This process is called "blind calibration".

VI. IMPLEMENTATION AND INTEGRATION OF A COVERAGE ESTIMATION WORK PATTERN

After an exhaustive presentation of all the elements and mechanisms of the proposed propagation model, this section discusses the microservices provided by AWS used in the implementation of the architectural framework. The implementation and integration in the Metric platform is also presented and detailed.

A. METRIC SaaS PLATFORM

As previously mentioned, the Metric platform allows bringing together in a single platform all the management of telecommunications networks. The Metric platform collects KPIs from the OSS of various manufacturers and technologies, presenting a unified view of network performance.

Several features of the Metric platform were implemented in the cloud using AWS microservices. An example of these features is the processing of performance indicators and DTs. Another example of using AWS microservices is when the data from multiple operators is received and integrated into the platform to make it available in the ETL process.

Using the AWS microservices allows the Metric platform to automatically respond to user actions, such as processing cell information. The Metric platform shows the characteristics of each of the cells, like the antenna model, azimuth and other information. Whenever the user uploads a file with new cell information, a microservice on AWS is triggered and allows to validate the cells with changed information and make it available through the platform.

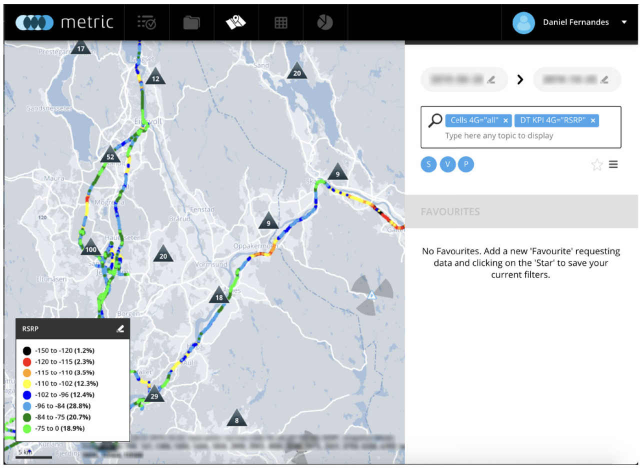

FIGURE 4. Drive Test to a given cell, obtained through the Metric platform.

Multivision company is currently migrating the services available on the Metric platform to an implementation using AWS. The new features developed will provide the platform with new services using AWS microservices. Using AWS cloud microservices allows Multivision to expand scalability and speed of the product developed to any device, in particular in the processing of data through the ETL process.

In Fig. 4 a search result to a given cell, obtained through the Metric platform is presented. The search result includes the DTs in a given time period.

B. AMAZON WEB SERVICES

The AWS provides services for performing computing tasks, such as Amazon Elastic Compute Cloud (Amazon EC2) and AWS Lambda. For database and storage purposes, microservices like Amazon Relational Database Service (Amazon RDS), Amazon DynamoDB, or Amazon Simple Storage Service (Amazon S3) can be used.

Amazon EC2 is a product that was developed with the aim of simplifying web-scale cloud computing. However, its actions cannot be triggered in response to an event, which makes this product unattractive to the solution to be implemented, since it is intended that the presented models run each time there is new cell information in a particular repository. This feature is supported by AWS Lambda which is a service that executes code in response to multiple events and supports multiple programming languages. It runs only when needed, which allows being automatically scalable. It can be used whenever the computing time is less than 15 minutes and the virtual memory required is less than 500 MBytes. In terms of memory, Lambda only provides 3 GBytes.

For database and storage purposes, AWS provides the Amazon RDS, which is a relational database that supports different commercial and open-source products, and does not require users to install, configure, and manage them.

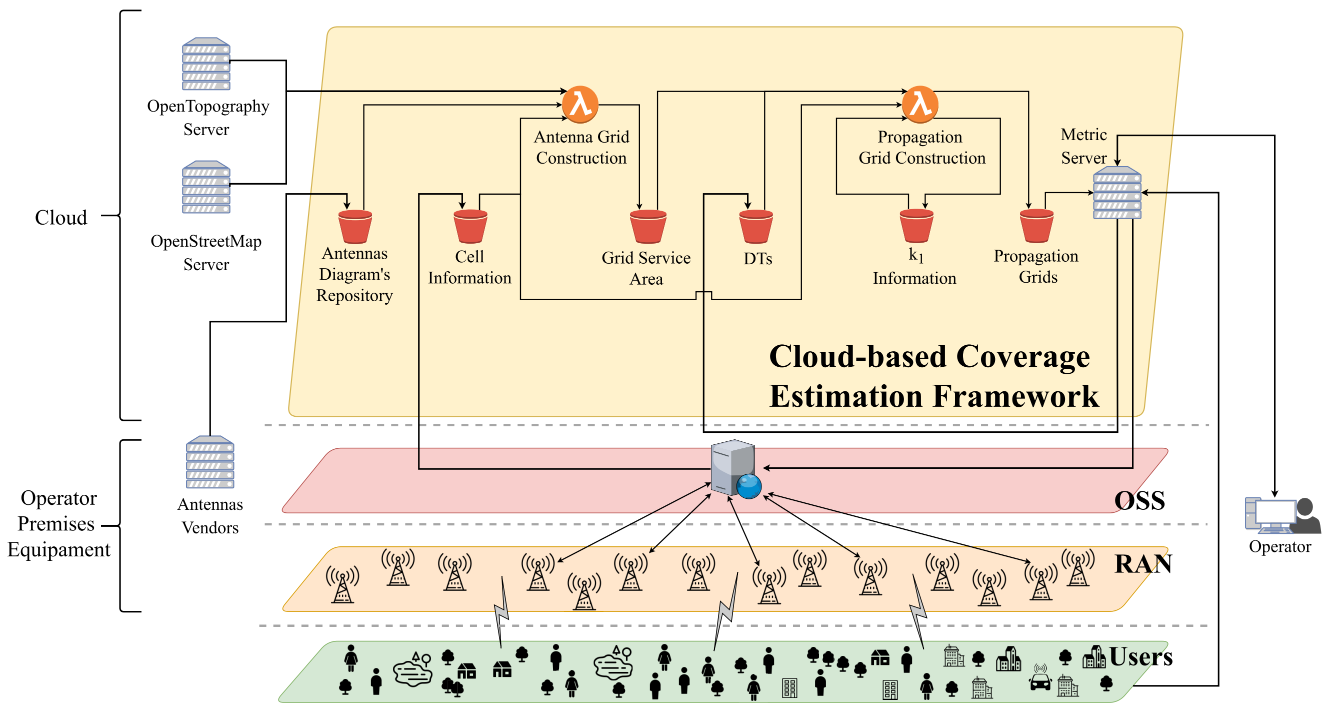

FIGURE 5. Mobile network monitoring and optimisation architecture, with work pattern using AWS services that implements the ACSPM.

In case of non-relational databases, Amazon also offers a user-friendly service, called Amazon DynamoDB. A storage service provided by AWS is Amazon S3, which allows users to store files regardless of their size (each file can have a maximum of 5 TBytes), and is highly scalable. The internal management of Amazon DynamoDB and Amazon S3 is done by AWS without the need for user intervention.

To implement the architectural framework presented in Fig. 1, the microservices AWS Lambda and Amazon S3 were used. Amazon S3 has an internal function that triggers whenever a new cell configuration file is received, starting the whole process.

The system architecture developed and implemented in AWS is depicted in Fig. 5.

C. IMPLEMENTATION OF THE SYSTEM ARCHITECTURE

A new model for estimating coverage, that combines various types of information present in the architecture, is shown in Fig. 1. A proof of concept of this architecture and its implementation using AWS is presented in Fig. 5. This figure also identifies the functions that are performed in the cloud and the resources used at the telecommunications operator premises. This subsection explains the details of each of the components of the proposed system architecture.

1) OVERVIEW

Using AWS, a proof of concept for a coverage estimation model was implemented. This implementation of the model was performed in Java [46] programming language.

The function of each element of the architecture is explained below.

The implementation of this proof of concept allows adding a new dimension in the scope of SON technology to the Metric platform, specifically for planning purposes. The use of this technology allows the decrease of overall operational costs, due to the reduction of human resources allocated to network planning tasks.

2) AMAZON S3

This implementation uses 6 independent Amazon S3 storage location. Each location stores distinct information, which may be used in the various stages of the process, or final information to be made available to the user. This location is also called a *bucket*.

- **Antennas Diagram's Repository** - For each antenna model, frequency and tilt, an ".msi" file containing the vertical and horizontal gain of that antenna are stored. This information is provided by the suppliers of the various antennas.
- **Cell Information** - For each cell of the network there is a JSON file that contains various information necessary for the process. Examples of this information are the technology, the location and model of the antenna, the electric and mechanical tilt, the P_{tx}[dBm] among others. The presence of a file in this Amazon S3 *bucket* triggers the process.
- **Grid Service Area** - Stores a JSON file with a precomputed grid for a cell. This grid contains for each pixel ground elevation and building information, antenna gain, antenna distance and the diffraction loss value.
- **DTs** - Cell DTs, whenever processed by Metric, are sent in form of a JSON file to this Amazon S3 *bucket*, and

they can be used whenever necessary, allowing calibration of the propagation model.

- K_1 **Information** - For each of the studied cells, a JSON file that compiles all the DTs performed for that cell is stored. Also the information of their coordinates and the value of the corresponding K_1 factor are stored. This file have the same usage as the DT file.
- **Propagation Grids** - In this Amazon S3 *bucket*, the propagation grids of each of the studied cell are stored in a JSON file. This *bucket* directly serves as input to Metric platform, making it possible to represent the propagation of a cell.

3) AWS LAMBDA FUNCTIONS

For processing purposes, AWS Lambda features the execution of code. This independent code executions are called AWS Lambda *functions*. For this implementation, 2 AWS Lambda *functions* are used and makes use of the data stored in the different Amazon S3 *buckets* as inputs. After the processing, the resulting information is stored in another *bucket* for later use.

- **Antenna Grid Construction** - From the information file present in *Cell Information bucket*, a AWS Lambda *function* creates a grid with the maximum propagation distance of the cell. Each pixel of the grid has user-defined dimensions, and for each one, the antenna gain considering the *Antennas Diagram's Repository bucket* files, antenna azimuth and tilts are calculated and stored. Elevation information requested from an external server is added to the grid, as explained in the module terrain data detailed in V-A. Once the elevation for each pixel is obtained, it is possible to calculate the existence or nonexistence of a LoS and the corresponding diffraction loss between a given pixel and the antenna. This information is also stored in the pixel. When all these computations are completed, the grid is stored in a JSON file in *Grid Service Area bucket* for later use.
- **Propagation Grid Construction** - For each JSON files in *Cell Information bucket* the corresponding JSON file is searched in *Grid Service Area bucket*. When the file is found, all the information is loaded as well as the existing DTs for that cell. If that cell has already been analysed before, processed information from old DT values can be found in K_1 *Information bucket*. This *function* uses all these information to fine tune the propagation model in order to generate a JSON file with the path loss. This file is stored in *Propagation Grids bucket*.

D. INTEGRATION WITH METRIC

In order to proceed with the integration in the Metric platform, AWS services were used, as explained above.

When a user wants to obtain the propagation of a certain cell through the Metric platform, the user must input "Propagation" in the search box (represented in red in Fig. 6). This action will show the propagation grid for the given cell. If the propagation grid does not exist, the process shown in Fig. 5 will be triggered, and after process completion, the propagation grid will be shown to the user. An example of this situation is shown in Fig. 6, where the propagation of a cell is presented to the user. This representation is the result of an initial implementation.

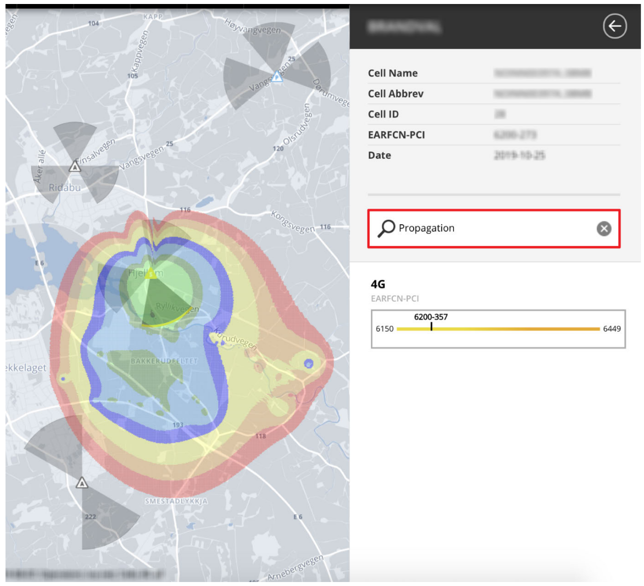

FIGURE 6. Cell coverage estimation module integrated in the Metric **platform.**

This integration in the Metric platform provides it with new features and allows users to explore, visualise and detect problems in their networks. These features can be categorised into three sets:

- **Planning** - With the propagation analysis, it is possible to perform the planning of Broadcast Control Channel (BCCH), Scrambling Codes (SC) and Physical Cell ID (PCI).
- **Evaluation and Optimisation** - Each time a new cell is implemented, it is possible to validate the coverage of this cell and the interference with the existing cells. To implement a new cell, it is necessary to initially generate the propagation grid of that cell, place this antenna in a map with the other antennas, and validate the impacts on the network already implemented. It is also possible through the analysis of coverage to optimise the network.
- **Malfunction Detection** - It is possible to detect some failures in the functioning of the network related to planning. These failures can be the identification of cross sectors and overshooting.

VII. REFERENCE SCENARIO

In order to test the presented architecture, a test scenario was chosen. In this scenario, a macro cell with a 4G antenna implemented in a northern European country is chosen. This antenna has a transmission power, P_{tx}, of 46 dBm and an azimuth value of 150°. The chosen antenna is from Kathrein

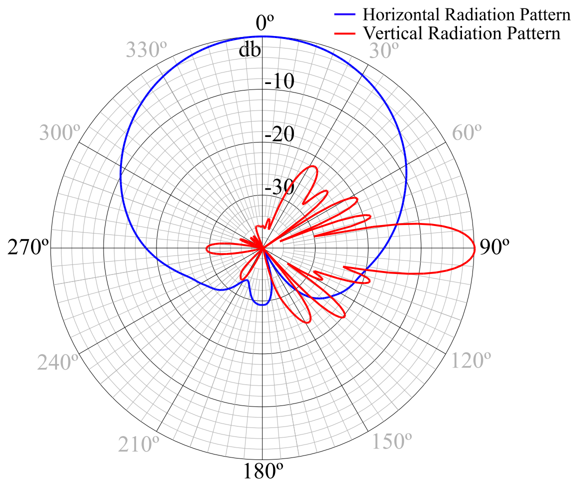

FIGURE 7. Horizontal and vertical radiation pattern of the reference antenna.

brand and the model is *80010665v01* configured with the mechanical and electrical tilt of 0°. The horizontal and vertical radiation patterns are depicted in Fig. 7. The E-UTRA Absolute Radio Frequency Channel Number (EARFCN) of the antenna is 6200, that corresponds to a downlink frequency of 796 MHz.

For this scenario, a Geo_{hash} parameter of 8 was chosen due to the fact that the reference scenario is located in a northern European region, where latitude presents higher values. This parameter allows the pixels of the generated grid to have the dimensions of 38.2 m ×19.1 m (width × height, approximately 2 : 1 pixel ratio) instead of using 1 : 1 pixel ratio (square pixels), which represents a more accurate approximation for geographic representation.

In Fig. 8 it is possible to visualise the scenario under study represented in the Metric platform. It is also represented in Fig. 8 the distance of the cell reach, which in this case is about 12 km, and the DTs associated with this antenna. The distance from the cell reach is associated with a certain receiving power that, in this case, was −85 dBm and up to this distance is where more than 95% of the cell activity is performed, that is, after this distance, the activity that involves the cell is practically residual.

For the scenario under study, the DTs performed over a 6 month period were considered, which contains a total of 248 measurement points. During these 6 months, the antenna configurations were not changed.

The grid generated, also represented in Fig. 8, covers an area of approximately 625 km^2 and the terrain elevation is depicted in Fig. 9. The antenna and its azimuth is represented as well.

VIII. PERFORMANCE RESULTS

This section presents the results obtained with the propagation model proposed in Fig. 3 and implemented as detailed in Fig. 5.

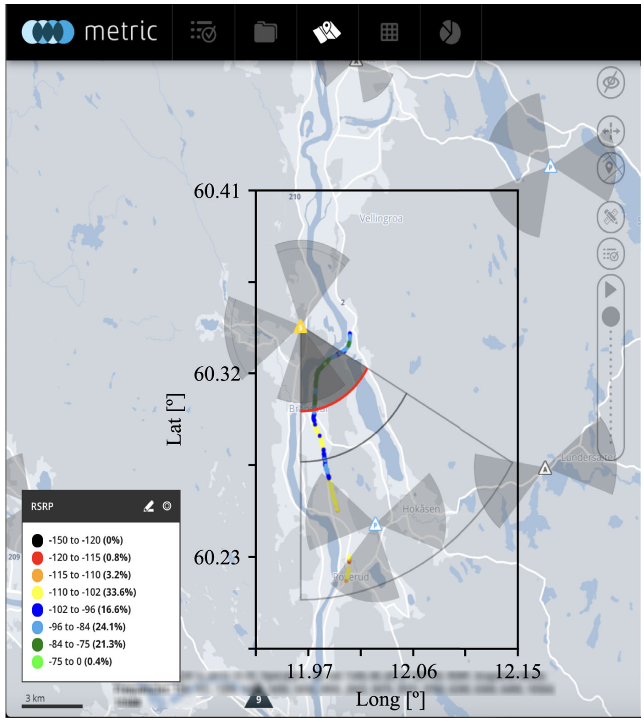

FIGURE 8. Propagation of a cell represented in the Metric **platform.**

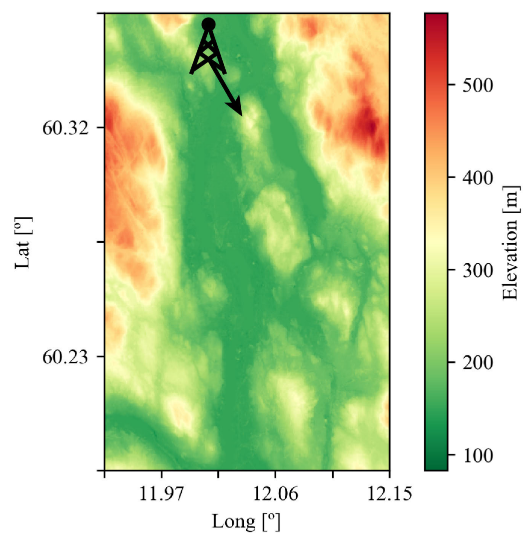

FIGURE 9. Terrain elevation of the reference scenario, with indication of the antenna's azimuth.

A. CELL COVERAGE

Following the model depicted in Fig. 3 and once the grid has been calculated, the attenuations due to obstacles and the gain for each of the pixels are calculated. Later, the DTs are distributed among the various pixels of the grid. The 248 DTs were distributed over 247 pixels, where one pixel contains 2 DTs.

Through the analysis of Fig. 8, it is perceptible that as the distance to the antenna increases, the value of the received power decreases.

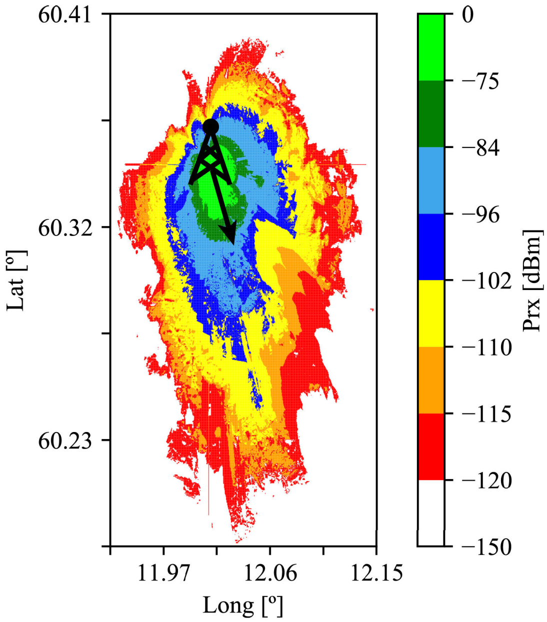

FIGURE 10. Estimated received signal level using the proposed propagation model.

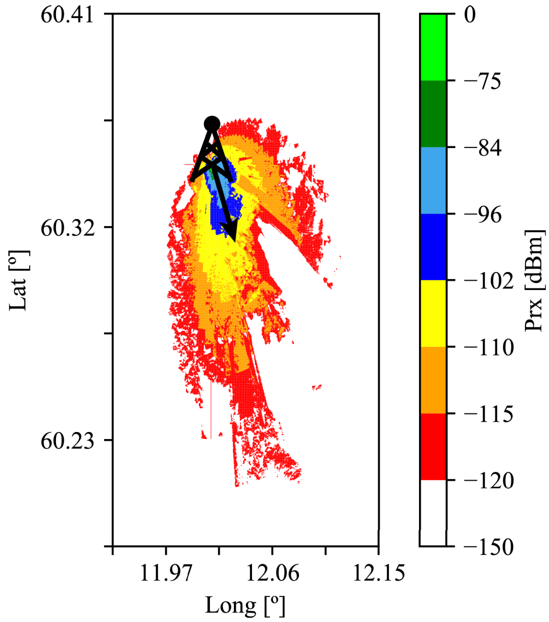

FIGURE 11. Estimated received signal level using the SPM model.

With the tuning of the propagation model, either through cell reach or through the DTs, the result is depicted in Fig. 10.

It is possible to verify that there are areas strongly influenced by DTs, which is the case of the lower zone where this influence can be seen. In this zone, the region between −102 dBm and −110 dBm (represented in yellow color), there is an area with a higher power signal received between −96 dBm and −102 dBm (represented in dark blue). This influence is due, not only by the values obtained with DTs, but also by the morphology of the terrain, which clearly shows that, in this area, there is an terrain elevation.

Although the SPM model is used for distances bigger than 1 km, in order to establish a comparison between the proposed model and a classical model, the SPM model was applied, with the parameters presented on Table 2. For the same scenario, the result of the coverage estimation is presented in Fig. 11.

For each pixel built with data from DTs, the received power value is compared with the estimated power value for the proposed model for the SPM model as explained in V-D.

Since the proposed model is calibrated with DT values, the MAE between the estimated values and the DT values is zero for that pixel. However, there may be some small oscillations for pixels with more than one DT, since the mean value for the received power is considered.

The comparison of the proposed ACSPM with the SPM can also be calculated. To simplify the process, for each propagation model, the MAE between the estimated values and the DT values is calculated and shown in Table 4.

The results of ''blind calibration'', present in V-D, is also depicted in Table 4.

As already mentioned, the absolute error between the values estimated through the proposed model and the DT values is zero. When looking at a SPM model, the situation is different. The absolute error between the value estimated by the SPM and the DTs is quite variable and the MAE value is about 14 dB. The maximum error can reach 30 dB.

TABLE 4. Absolute error between estimated values and DT values for the 247 DT pixels.

	Error [dB]		
	ACSPM	SPM	"Blind Calibration"
Maximum	0.0	30.0	11.1
MAE	0.0	13.6	5.2
Minimum	0.0	1.9	0.0

For the accuracy of the proposed model, is visible that the MAE value, 5 dB is about 8 dB lower when compared with the SPM model. The maximum value is also lower, almost one-third the value of SPM model.

B. IMPACT OF DT ON THE PROPOSED MODEL

Since the proposed model is calibrated using DTs values, it is important to validate the impact on the number of DTs used in the model calibration. Not only the amount of DTs is important but also their distribution around the antenna as the propagation conditions are variable even for the same distance due to the different angles.

Fig. 12 depicts the MAE values, between the values estimated through the proposed model and the DT values, for the set used for evaluation of the methodology. A second order polynomial regression is also shown to visualise MAE trend when DTs measurements increase. As expected, MAE values decrease when more DTs are considered. The increase of available DT measurements, namely with MDT, make the proposed model more robust to outliers, automatically improving the quality of the achieved coverage predictions.

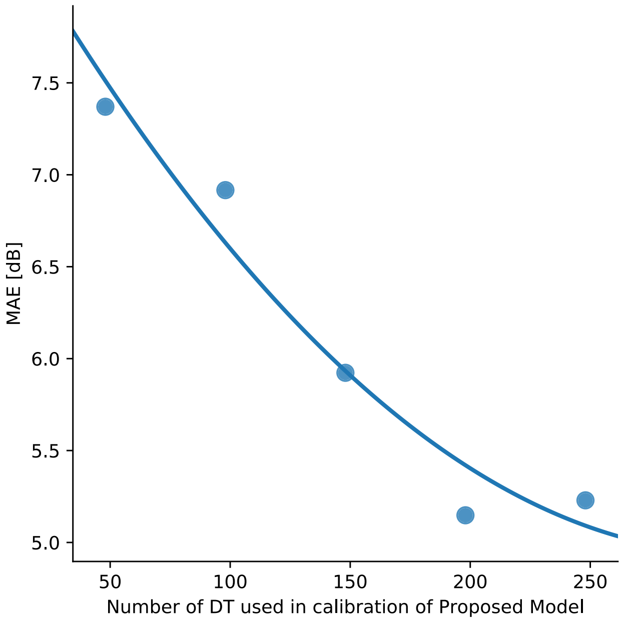

FIGURE 12. Polynomial regression for MAE values.

Once the various network measures are integrated into this model, if mobile phones had the option of monitoring the active network quality by reporting such data, network knowledge would be more realistic as well as its estimation.

C. PLATFORM PERFORMANCE

By splitting the computational effort into two distinct Lambda *functions*, the waiting time for the end-user when making the request is drastically reduced. Initially, the Lambda *function* that needs more computational power is calculated for further reuse. Then, after user request, that information is used to generate the final propagation grid.

The computation of these tasks in a cloud environment allows quick access to the files, fast processing of the pre-defined tasks through computational parallelization. The result of this process can be further accessed simultaneously by several users.

Considering a network with 3 304 cells, the implementation in AWS of the proposed propagation model was tested. Table 5 refers to the computation times using AWS implementation and "No AWS" implementation, of each of the Lambda *functions* as well as the total elapsed time.

The results achieved by the AWS implementation are due to the parallelization feature that AWS services provide in their Lambda *function* execution. The case where parallelization is not considered, being the calculations sequential, was considered to be "No AWS" implementation.

In the case of Lambda *function*, **Antenna Grid Construction**, each cell needs on average 9 minutes and 8 seconds to be computed and in the case of Lambda *function*, **Propagation Grid Construction**, only 62 seconds are needed, being the time that user, in the worst case scenario, can expect to obtain propagation results from a cell.

TABLE 5. Computing time of the ACSPM, for 3 304 cells, on the AWS and not on the AWS. The format of the values is hh:mm:ss.

	Computation Time		
	Antenna Grid Construction	Propagation Grid Construction	Total
AWS	05:01:49	00:37:56	05:39:45
No AWS	502:32:26	56:34:07	559:06:33

In the scenario where the user wants to get the propagation information for 3 304 cells, the use of AWS implementation allows a 89.7% time reduction when compared with not using AWS implementation.

IX. MECHANISMS OF METRIC IMPROVED BY THE PROPOSED METHODOLOGY

The proposed work pattern for estimation of signal level around any cell simplifies the process of planning and optimisation of telecommunication networks, making the process simpler, faster and more accurate. The grids generated have been integrated in several mechanisms that have been developed, implemented and integrated in Metric platform:

- **Network (Cells Deployment) Coverage:** One of the applications where the information from the ACSPM is used is when it is intended to have a coverage overview in a given region. For each of the antennas in the region, a propagation grid is generated, and then all this information is combined. An example of the coverage of a given area is depicted in Fig. 13. This mechanism has been implemented in Metric platform [47]. Through the analysis of this area, it is possible to identify, for example, areas with weak coverage, as discussed in Section IV-B.
- **Cells Neighbourhood List:** The list of neighbouring cells is essential in the operation of the network, helping in the handover of mobile users, as well as in the planning of the 2G, 3G and 4G network resources. Through the superposition of grids, within the coverage area of a cell are calculated the amount of pixels covered by each neighbouring cell. This results in a list of neighbouring cells sorted by estimates of shared coverage, a mechanism that has been implemented in Metric platform, as detailed in [47].
- **Crossed Sectors identification:** When an anomaly occurs in the installation of the transmission cables supplying a certain cell, the signal emitted by that cell is not as planned. These situations can be easily identified by comparing the estimated signal with the DT data taken from the study area. The generation of a path loss grid for a given cell is transparent to the antenna characteristics, so the proposed model is used to estimate the signal with the planned characteristics. This mechanism has been implemented in Metric platform.
- **Overshooting solution:** One problem that also exists in telecommunication networks is the overshooting cells. This type of problems means that the cell in question has a signal outside its normal area of coverage.

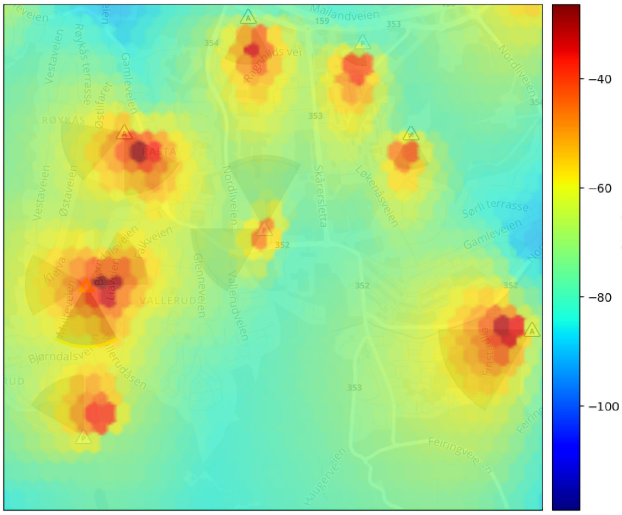

FIGURE 13. Signal received in a certain area.

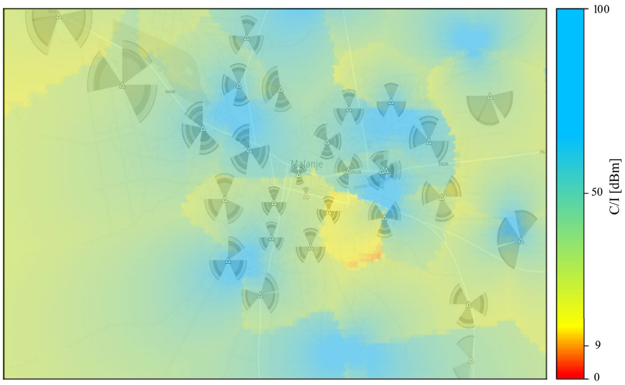

FIGURE 14. Carrier to Interference ratio in a certain area.

This problem drastically reduces the quality of the network. The analysis between the planned area of propagation and the DT quickly detects this problem. This mechanism has been implemented in Metric platform.

- **Planning:** With the generation of the received power grids, it is possible not only to elaborate a coverage map for an area but also a map with the interference of the various antennas, represented in Fig. 14. This information allows the planning of BCCHs and is detailed in [48]. This mechanism has been implemented in Metric platform. On the other side, through the power grids per cell, the lists of neighbors are then created, which are then used for SC and PCI planning (presented in [49]).
- **Energy Efficient Management of Resources**: Based on traffic predictions per cell, the cellular coverage area can be energy-efficiently optimised by a novel mechanism, as detailed in [24], [50], [51]. This proposed mechanism starts with grids of P_{rx} for each of the cells. This mechanism has been implemented in Metric platform.

X. CONCLUSION

This paper presents a cloud-based framework of a novel semi-empirical propagation model that portrays, as accurately as possible, the propagation of an antenna. This model uses samples and antenna configuration data to automatically calibrate the model, making it more accurate. The automation of the calibration process makes it possible to reduce the human effort, which results in a financial impact on the management of these networks. This implementation, allows greater flexibility in the use of this model. This flexibility is related, not only to the ease of integration of the model in the Metric platform, for management and monitoring of telecommunications networks, but also to the features given by cloud services, namely in terms of processing and memory usage, and the availability of elastic resources on-demand, when needed.

The ACSPM was implemented and tested for a telecommunications network located in the northern European region, specifically, in a scenario where an antenna and its DTs were considered. The results obtained present about 13 dB gain when compared with SPM, and a MAE of about 5 dB in terms of model accuracy, which is about 8 dB lower than that obtained by SPM.

This work also shows what happens to the accuracy of the model when the number of DTs used for calibration changes. In fact, our model shows that, as the number of DTs used for model calibration increases, the MAE decreases, which makes our model more accurate, realistic and more robust to outlier samples.

This paper also highlights the importance of a precise propagation model for the correct implementation and optimisation of a telecommunications network. This has been integrated and implemented with other algorithms previously available in the tool, and is now in use allowing to build a neighbours list for a given cell, to plan resources of a 2G, 3G and 4G network and change configurations (overshooting or crossed sectors), as well as for more refined energy efficient mechanisms of radio resources.

The implementation and model described in this paper shows that the integration of MDT, to obtain more data describing the propagation environment, can be done in a fully automatic way, provided that operators make the data available, as it increases the responsiveness of the proposed model to the propagation environment. For future work, it would be interesting to explore a scenario with available MDT data for the calibration of the model, since it allows an even more realistic signal estimation.

REFERENCES

[1] *Ericsson Mobility Report 2019*, Ericsson, Stockholm, Sweden, Nov. 2019.
[2] *Nokia*. Accessed: Jul. 15, 2019. [Online]. Available: https://www.nokia.com
[3] *Huawei*. Accessed: Jul. 15, 2019. [Online]. Available: https://consumer.huawei.com
[4] *Atoll 3.3.0 Technical Reference Guide for Radio Networks*, Forsk, Jaipur, Rajasthan, 2015.

[5] Z. O. N. Martinez, C. Rodriguez, and O. M. Arias, "Propagation characteristics of managua city based on standard propagation model (SPM) at 850 MHz for 3G-WCDMA systems," in *Proc. IEEE Central Amer. Panama Conv. (CONCAPAN)*, Nov. 2014, pp. 1–6.
[6] S. I. Popoola, A. A. Atayero, N. Faruk, C. T. Calafate, E. Adetiba, and V. O. Matthews, "Calibrating the standard path loss model for urban environments using field measurements and geospatial data," *Lect. Notes Eng. Comput. Sci.*, vol. 2229, no. July, pp. 513–518, 2017.
[7] L. M. Correia, *Mobile Broadband Multimedia Networks*. Amsterdam, The Netherlands: Elsevier, 2007.
[8] D. Fernandes, L. S. Ferreira, M. Nozari, P. Sebastiao, F. Cercas, and R. Dinis, "Combining drive tests and automatically tuned propagation models in the construction of path loss grids," in *Proc. IEEE 29th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2018, pp. 1–2.
[9] D. Fernandes *et al.*, "Combining measurements and propagation models for estimation of coverage in wireless networks," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Honolulu, HI, USA, 2019, pp. 1–5.
[10] D. Fernandes, G. Soares, D. Clemente, R. Cortesão, P. Sebastião, F. Cercas, R. Dinis, and L. S. Ferreira, "Integration of a cloud-based realistic and automatic coverage estimation methodology in metric SaaS," in *Proc. 22nd Int. Symp. Wireless Pers. Multimedia Commun. (WPMC)*, Lisbon, Portugal, Nov. 2019, pp. 326–329.
[11] A. De Mauro, M. Greco, and M. Grimaldi, "What is big data? A consensual definition and a review of key research topics," *AIP Conf. Proc.*, vol. 1644, pp. 97–104, Feb. 2015.
[12] K. Zheng, Z. Yang, K. Zhang, P. Chatzimisios, K. Yang, and W. Xiang, "Big data-driven optimization for mobile networks toward 5G," *IEEE Netw.*, vol. 30, no. 1, pp. 44–51, Jan./Feb. 2016.
[13] A. Imran and A. Zoha, "Challenges in 5G: How to empower SON with big data for enabling 5G," *IEEE Netw.*, vol. 28, no. 6, pp. 27–33, Nov. 2014.
[14] E. Bastug, M. Bennis, E. Zeydan, M. A. Kader, I. A. Karatepe, A. S. Er, and M. Debbah, "Big data meets telcos: A proactive caching perspective," *J. Commun. Netw.*, vol. 17, no. 6, pp. 549–557, Dec. 2015.
[15] M. Frankel, N. Shacham, and J. E. Mathis, "Self-Organizing Networks," *Future Gener. Comput. Syst.*, vol. 4, no. 2, pp. 95–115, 9 1988.
[16] *NGMN Use Cases related to Self Organising Network, Overall Description*, Berlin, Germany, NGMN, p. 17, 2008.
[17] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, "A survey of machine learning techniques applied to self-organizing cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2392–2431, 4th Quart., 2017.
[18] *AWS*. Accessed: Jul. 15, 2019. [Online]. Available: https://aws.amazon.com
[19] *Google Cloud*. Accessed: Jul. 15, 2019. [Online]. Available: https://cloud.google.com
[20] *Microsoft Azure*. Accessed: Jul. 15, 2019.[Online]. Available: https://azure.microsoft.com
[21] D. Sabella, P. Rost, Y. Sheng, E. Pateromichelakis, U. Salim, P. Guitton-Ouhamou, M. Di Girolamo, and G. Giuliani, "RAN as a service: Challenges of designing a flexible RAN architecture in a cloud-based heterogeneous mobile network," *Future Netw. Mobile Summit*, no. 2, pp. 1–8, Jun. 2013.
[22] *Metric*. Accessed: Jul. 16, 2019. [Online]. Available: https://metric.pt
[23] *Multivision*. Accessed: Jul. 16, 2019. [Online]. Available: https://multivision.pt
[24] D. Clemente, L. Ferreira, G. Soares, N. Valente, and P. Sebastiao, "Implementation of a cloud-based, traffic aware and energy efficient management of base Stations' activity," in *Proc. 21st Int. Symp. Wireless Pers. Multimedia Commun. (WPMC)*, Nov. 2018, pp. 600–605.
[25] A. F. Molish, *Wireless Communication*, 2nd ed. Hoboken, NJ, USA: Wiley, 2010.
[26] M. F. Iskander and Z. Yun, "Propagation prediction models for wireless communication systems," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 3, pp. 662–673, Mar. 2002.
[27] F. Letourneux, S. Guivarch, and Y. Lostanlen, "Propagation models for heterogeneous networks," *Proc. 7th Eur. Conf. Antennas Propag. (EuCAP)*, Jan. 2013, pp. 3993–3997.
[28] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE Trans. Veh. Technol.*, vol. 29, no. 3, pp. 317–325, Aug. 1980.
[29] A. R. Mishra, *Advanced Cellular Network Planning and Optimisation*, A. R. Mishra, Ed. Chichester, U.K.: Wiley, 2006.
[30] J. M. Mom, C. O. Mgbe, and G. A. Igwue, "Application of artificial neural network for path loss prediction in urban macrocellular environment," *Amer. J. Eng. Res.*, vol. 3, no. 2, pp. 270–275, 2014.
[31] E. Ostlin, H. Zepernick, and H. Suzuki, "Macrocell path-loss prediction using artificial neural networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2735–2747, Jul. 2010.
[32] I. Popescu, I. Nafornita, G. Gavriloaia, P. Constantinou, and C. Gordan, "Field strength prediction in indoor environment with a neural model," in *Proc. 5th Int. Conf. Telecommun. Modern Satell.*, 2001, pp. 640–643.
[33] M. Ayadi, A. Ben Zineb, and S. Tabbane, "A UHF path loss model using learning machine for heterogeneous networks," *IEEE Trans. Antennas Propag.*, vol. 65, no. 7, pp. 3675–3683, Jul. 2017.
[34] W. A. Hapsari, A. Umesh, M. Iwamura, M. Tomala, B. Gyula, and B. Sebire, "Minimization of drive tests solution in 3GPP," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 28–36, Jun. 2012.
[35] *Telecom Hall*. Accessed: Aug. 1, 2019. [Online]. Available: http://www.telecomhall.com/analyzing-coverage-with-propagation-delay-pd-and-timing-advance-ta-gsm-wcdma-lte.aspx
[36] R. Moussalli, M. Srivatsa, and S. Asaad, "Fast and flexible conversion of geohash codes to and from Latitude/Longitude coordinates," in *Proc. IEEE 23rd Annu. Int. Symp. Field-Program. Custom Comput. Mach.*, May 2015, pp. 179–186.
[37] *OpenTopography—High-Resolution Topography Data and Tools*. Accessed: Aug. 1, 2019. [Online]. Available: https://opentopography.org
[38] J. Takaku, T. Tadono, and K. Tsutsui, "Generation of high resolution global DSM from ALOS PRISM," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vols. XL–4, pp. 243–248, Apr. 2014.
[39] T. Tadono, H. Ishida, F. Oda, S. Naito, K. Minakawa, and H. Iwamoto, "Precise global DEM generation by ALOS PRISM," *Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 4, no. 5, pp. 71–76, Apr. 2014.
[40] *Overpass API*. Accessed: Aug. 1, 2019. [Online]. Available: http://overpass-api.de
[41] *OpenStreetMap*. Accessed: Aug. 1, 2019. [Online]. Available: https://www.openstreetmap.org/
[42] M. Abrash, *Michael Abrash's Graphics Programming Black Book: The Complete Works of Graphics Master, Michael Abrash*, 10th ed. Stockbridge, GA, USA: Coriolis Group Books, 1997.
[43] *Propagation by Diffraction*, document ITU-R P.526-14, 2018.
[44] S. Saunders and A. Aragón, *Antennas and Propagation for Wireless Communication Systems*, 2nd ed., Hoboken, NJ, USA: Wiley, 2007.
[45] C. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Res.*, vol. 30, pp. 79–82, Jun. 2005.
[46] *Java Platform Standard Edition 7 Documentation*. Accessed: Aug. 1, 2019. [Online]. Available: https://docs.oracle.com/javase/7/docs/index.html
[47] P. Pina, A. Godinho, D. Fernandes, D. Clemente, G. Soares, P. Sebastião, and L. S. Ferreira, "Cloud-based implementation of an automatic pixel-based neighbour identification system for cellular networks," in *Proc. 22nd Int. Symp. Wireless Pers. Multimedia Commun (WPMC)*, Lisbon, Portugal, Nov. 2019, pp. 305–310.
[48] A. Godinho, D. Fernandes, D. Clemente, G. Soares, P. Sebastião, P. Pina, and L. S. Ferreira, "Cloud-based cellular network planning system: Proof-of-concept implementation for GSM in AWS," in *Proc. 22nd Int. Symp. Wireless Pers. Multimedia Commun (WPMC)*, Lisbon, Portugal, Nov. 2019, pp. 311–315.
[49] R. Cortesão, D. Fernandes, D. Clemente, G. Soares, P. Sebastião, and L. S. Ferreira, "Cloud-based implementation of a SON automatic planning system: A proof-of-concept for UMTS," in *Proc. 22nd Int. Symp. Wireless Pers. Multimedia Commun (WPMC)*, Lisbon, Portugal, Nov. 2019, pp. 316–321.
[50] D. Clemente, G. Soares, D. Fernandes, R. Cortesao, P. Sebastiao, and L. S. Ferreira, "Traffic forecast in mobile networks: Classification system using machine learning," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Honolulu, HI, USA, 2019, pp. 1–5.
[51] D. Clemente, D. Fernandes, D. Clemente, R. Cortesão, G. Soares, P. Sebastião, and L. S. Ferreira, "Assessment of traffic prediction models for mobile communication networks," in *Proc. 22nd Int. Symp. Wireless Pers. Multimedia Commun. (WPMC)*, Lisbon, Portugal, Nov. 2019, pp. 322–325.

DANIEL FERNANDES received the B.S. and M.S. degrees in telecommunications and computer engineering from ISCTE-IUL, Lisbon, Portugal, in 2012 and 2017, respectively, where he is currently pursuing the Ph.D. degree in information science and technology. He is currently working as a Researcher with Multivision—Consultoria Informática. He has four scientific publications as author in international conferences and seven scientific publications as coauthor.

DIOGO CLEMENTE received the B.S. degree in electrical and computer engineering from the Branch of Renewable Systems and Power Systems, Escola Superior de Tecnologia de Setúbal-ESTSetúbal/IPS, and the M.S. degree in telecommunications and computer engineering from the University Institute of Lisbon (ISCTE-IUL), Lisbon, Portugal, where he is currently pursuing the Ph.D. degree in information science and technology. He is currently working as a Researcher with Multivision—Consultoria Informática. He has eight scientific publications in international conferences.

GABRIELA SOARES received the B.S. degree in computational mathematics from the Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil, the M.S. degree in mathematical and computational modeling from the Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), Belo Horizonte, and the Ph.D. degree in computer science from the Universidade de São Paulo (USP), São Paulo, Brazil. She is currently working as a Data Scientist Researcher with Multivision. She participated as a Researcher in three Research and Development projects funded by FAPESP, CAPES, and CNPQ focused on optimization, evolutionary computing, complex networks, and data science. She has 17 scientific publications in journal articles and international conferences.

PEDRO SEBASTIÃO (Member, IEEE) received the Ph.D. degree in electrical and computer engineering from IST. He is currently a Professor with ISCTE-IUL's Information Science and Technology Department. He is also the Board Director of AUDAX-ISCTE-Entrepreneurship and Innovation Center, ISCTE, and responsible for the LABS LISBOA incubator and researcher, Institute of Telecommunications. He developed his professional activity at the National Defense Industries, initially with the Office of Studies and later as the Board Director of the Quality Department, Production of New Products and Technologies. He was also responsible for systems of communications technology at the Nokia-Siemens business area. He has organized or co-organized more than 50 national and international scientific conferences. He planned and developed several postgraduate courses in technologies and management, entrepreneurship and innovation, and transfer of technology and innovation. He has supported several projects involving technology transfer and creation of start-ups and spin offs of value to society and market. He has oriented several master's dissertations and doctoral theses. He has been an expert and evaluator of more than 100 national and international civil and defense Research and Development projects. It has several scientific, engineering, and pedagogical awards. He is the author or coauthor of more than 200 scientific articles. He has been responsible for several national and international Research and Development projects. His main research interests are in monitoring, control and communications of drones, unmanned vehicles, planning tools, stochastic process (modeling and efficient simulations), the Internet of Things, and efficient communication systems.

FRANCISCO CERCAS (Senior Member, IEEE) received the Ph.D. degree from the Instituto Superior Técnico, Lisbon University, Portugal, in 1996, and the ISCTE-University Institute of Lisbon. He was Researcher with INESC, from 1984 to 1985, and CAPS, from 1986 to 1993. He was also a Visiting Researcher with the University of Plymouth, U.K., from 1987 to 1992. Since 1994, he has been a Researcher with the Instituto de Telecomunicações. He is currently a Full Professor with more than 36 years of professional experience, including Research and Development at the industry and 35 years of university teaching with the Instituto Superior Técnico. He also supervised many Ph.D. and M.Sc. students. He has participated in more than a dozen european projects in the telecommunications area, namely as the portuguese delegate in European Cooperation in Science and Technology (4 COST) actions. He has authored and/or coauthored of a new class of codes, Tomlinson, Cercas, Hughes (TCH), one patent, and about 200 publications including book chapters, journal articles and conference papers. His research interests include satellite and mobile communications, coding theory, spread spectrum communications and related topics.

RUI DINIS (Senior Member, IEEE) received the Ph.D. degree from the Instituto Superior Técnico (IST), Technical University of Lisbon, Portugal, in 2001, and the Habilitation degree in telecommunications from the Faculdade de Ciências e Tecnologia (FCT), Universidade Nova de Lisboa (UNL), in 2010. From 2001 to 2008, he was a Professor with IST. He was a Researcher with the Centro de Análise e Processamento de Sinal (CAPS), IST, from 1992 to 2005, and a Researcher with the Instituto de Sistemas e Robótica (ISR), from 2005 to 2008. In 2003, he was an Invited Professor with Carleton University, Ottawa, ON, Canada. Since 2009, he has been a Researcher with the Instituto de Telecomunicações (IT). He is currently an Associate Professor with FCT-UNL. He has been actively involved in several national and international research projects in the broadband wireless communications area. His research interests include transmission, estimation, and detection techniques. He is a VTS Distinguished Lecturer. He is or was an Editor of the IEEE Transactions on Wireless Communications, the IEEE Transactions on Communications, the IEEE Transactions on Vehicular Technology, the IEEE *Open Journal on Communications*, and the *Elsevier Physical Communication*. He was also a Guest Editor of the *Elsevier Physical Communication* (Special Issue on Broadband Single-Carrier Transmission Techniques).

LÚCIO S. FERREIRA (Senior Member, IEEE) received the Licenciado and Ph.D. degrees in electrical and computer engineering from the IST/Technical University of Lisbon, Portugal. He worked as a Researcher with Deutsche Telekom Innovation Laboratories, in 1998, Instituto de Telecomunicações, from 1999 to 2012, and INOV-INESC, from 2012 to 2015. He has worked as an Assistant Professor with the Universidade da Beira Interior, Universidade Lusiada de Lisboa, and ISTEC, in the areas of computer science and telecommunications. He has been with INESC-ID, since 2016. He is currently a Project Manager with Multivision—Consultoria Informatica. He participated, as a Project Manager and a Researcher, in 17 Research and Development projects funded by the European Commission. At the National Level, he has worked as a Consultant for mobile providers and ANACOM national regulator. He has more than 50 scientific publications in books, journal articles, and international conferences. He is an editor and coauthor of 63 technical reports for the European Commission. He reviewed 29 journal articles and conference papers, and participated in eight organisational commitees and 15 technical committees of international conferences. He has supervised 12 M.Sc. and three Ph.D. students. He is a Board Member of the IEEE ComSoc Portugal chapter.