Automated estimation of link quality for Lora

A remote sensing approach

Demetri, Silvia; Zúñiga, Marco; Picco, Gian Pietro; Kuipers, Fernando; Bruzzone, Lorenzo; Telkamp, Thomas

Citation (APA)
Demetri, S., Zúñiga, M., Picco, G. P., Kuipers, F., Bruzzone, L., & Telkamp, T. (2019). Automated estimation of link quality for Lora: A remote sensing approach. In R. Eskicioglu (Ed.), *IPSN 2019 - Proceedings of the 2019 Information Processing in Sensor Networks: Proceedings of the 2019 Information Processing in Sensor Networks* (pp. 145-156). (IPSN 2019 - Proceedings of the 2019 Information Processing in Sensor Networks). Association for Computing Machinery (ACM). https://doi.org/10.1145/3302506.3310396

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Automated Estimation of Link Quality for LoRa:
A Remote Sensing Approach

Silvia Demetri
University of Trento, Italy
silvia.demetri@unitn.it

Marco Zúñiga
TU Delft, The Netherlands
m.a.zunigazamalloa@tudelft.nl

Gian Pietro Picco
University of Trento, Italy
gianpietro.picco@unitn.it

Fernando Kuipers
TU Delft, The Netherlands
f.a.kuipers@tudelft.nl

Lorenzo Bruzzone
University of Trento, Italy
lorenzo.bruzzone@unitn.it

Thomas Telkamp
Lacuna Space
thomas@telkamp.eu

ABSTRACT

Many research and industrial communities are betting on LoRa to provide *reliable, long-range communication* for the Internet of Things. This new radio technology, however, provides widely heterogeneous coverage; a LoRa link may span hundreds of meters or tens of kilometers, depending on the surrounding environment. This high variability is not captured by popular channel models for LoRa, and on-site measurements—a common alternative—are impractical due to the large geographical areas involved.

We propose a novel, *automated* approach to estimate the coverage of LoRa gateways *prior to deployment and without on-site measurements*. We achieve this goal by combining free, readily-available multispectral images from remote sensing with the *right* channel model. Our processing toolchain automatically classifies the type of environment (e.g., buildings, trees, or open fields) traversed by a link, with high accuracy (>90%) and spatial resolution ($10\times10m^2$). We use this information to explain the attenuation observed in experiments. As signal attenuation is not well captured by popular channel models, we focus on the Okumura-Hata empirical model, hitherto largely unexplored for LoRa, and show that *i)* it yields estimates very close to our observations, and *ii)* we can use our toolchain to *automatically* select and configure its parameters. A validation on 8,000+ samples from a real dataset shows that our automated approach predicts the expected signal power within a ~10dBm error, against the 20–40dBm of popular channel models.

KEYWORDS

LoRa, link quality, LPWAN, remote sensing, multispectral images

1 INTRODUCTION

The success of the Internet of Things (IoT) will depend fundamentally on the ability to provide *reliable communication* to the billions of devices that will monitor our cities and rural areas. Many RF technologies are competing to achieve this goal; among these, LoRa and, more generally, Low-Power Wide Area Networks (LPWAN) are at the forefront of commercial interest. The reason is that LoRa promises communication coverage spanning tens of kilometers; a potentially high number of IoT devices can be covered by a single gateway and with a simple star topology, simplifying deployment, operation, and management of the communication infrastructure.

Problem and motivation. These potential advantages, however, must be checked against reality: it is well-known that the performance of a wireless link is intrinsically dependent on the environment it traverses—and LoRa is no exception. Open, urban and rural areas attenuate and distort the wireless signal in very different manners. Unfortunately, although there is plenty of anecdotal evidence about LoRa links being much shorter than claimed, channel models and empirical evidence are still largely lacking in the literature, with the exception of a few preliminary works [8, 25, 28, 45].

One reason is that accruing the necessary empirical knowledge, even for a well-defined and limited deployment, is significantly more demanding with LoRa. In short-range technologies like the IEEE 802.14.5 radios popular in wireless sensor networks and IoT, the vagaries of wireless communication are assessed by deploying a few nodes over a relatively small area (typically ~100-500 m^2) to obtain key radio propagation parameters required by channel models. These can later be used to estimate the coverage and reliability of communication in the same (or similar) environment, which are typically uniform across the deployment.

In contrast, performing the same on-site measurement campaigns to assess LoRa long-range entails deploying nodes over areas ~10–100 km^2, i.e., orders of magnitude larger, with a corresponding increase in logistical complexity and personnel effort. More importantly, these results do not necessarily transfer directly to similar environments. Indeed, unlike short-range radio links, LoRa long-range links *i)* traverse *multiple* types of environments, *ii)* whose spatial extension and sequence are highly variable, and typically *a priori unknown*, in the large area under consideration.

Considering the thousands of LoRa gateways deployed worldwide and the millions of km^2 they cover, here we ask ourselves:

How can we model LoRa links and estimate their quality
in a simple and low-cost manner?

The crux of the matter is that the *actual* performance of LoRa links, and therefore the ability to *model and estimate them beforehand*, strictly depends on *local and specific* features of the environment. These features are however difficult to obtain with on-site measurements, due to the *large scale of the target area*. A key asset of our approach is the use of remote sensing techniques to obtain environment-specific information in a low-cost manner.

Key asset: Remote Sensing. Remote Sensing (RS) systems acquire data (mainly images) over wide areas by exploiting the propagation and reflection properties of electromagnetic waves. The target scene is illuminated by a source of electromagnetic radiation, and sensors mounted on satellites, airplanes or UAVs, measure the radiation

reflected by the objects in the scene. The acquired data enables the *automatic* extraction of detailed information over a large-scale area.

Remote sensing is a very active area, with several systems in operation [11, 29, 41], roughly distinct in two classes. Passive systems (e.g., multispectral and hyperspectral scanners) mainly exploit the sun as the source of radiation and measure the spectral response of the objects across several spectral channels. The multidimensional images acquired represent the "spectral signature" of the investigated objects, enabling the extraction of information about the composition of materials. Active systems (e.g., Light detection and ranging, LiDAR, Synthetic Aperture Radar, SAR), generate a coherent radiation themselves and are able to capture also the vertical structure of the scene (i.e., heights) exploiting the delay, amplitude and phase of the reflected signal.

Remote sensing systems have been widely applied to large-scale monitoring of the Earth surface, and their data exploited for a plethora of services. These notably include the generation of *land-cover and land-use maps* [24, 29], which classify each pixel of the image with, e.g., the presence of man-made artifacts like buildings and roads, or natural features like trees, vegetation, and water. It is precisely this capability and related body of work that we leverage to classify the types of environment traversed by LoRa links.

Contributions and relevance. The key contribution of this paper is a novel, *automated* approach to estimate the quality of LoRa links *prior to deployment and without on-site measurements*. We achieve this goal by performing a structured analysis of the LoRa link.

First, we analyze empirically the maximum range of LoRa in free space with a weather balloon (§3), and use that evaluation as a baseline to gain insights about the negative effects of the environment at ground level (§4).

Second, motivated by the insights gained in our initial experiments, we propose a dedicated processing toolchain (based on Sentinel-2 multispectral images [1]), to automatically analyze the types of environment traversed by LoRa links. Our system *i)* exploits automatic classification techniques to discern seven land-cover classes typical in LoRa deployments (§5) with high accuracy (>90%) and spatial resolution ($10{\times}10\text{m}^2$), and *ii)* allows us to establish correlations between the type of land-cover and the Packet Reception Rate (PRR) of links (§6).

Third, we analyze the models proposed in the literature for LoRa [7], identify their shortcomings and propose a framework that builds upon the Okumura-Hata empirical model, hitherto largely unexplored for LoRa (§7). Our approach *i)* uses our remote sensing toolchain to automatically select and configure the model parameters, and *ii)* yields estimates that are very close to our observations.

Fourth, we validate the combined use of the Okumura-Hata model and our toolchain on a large real-world dataset from "The Things Network" (TTN) (§8). Our results show that we can estimate the signal power in complex urban environments within a 10 dBm error, compared to the 20−40 dBm error of existing models. This confirms that our approach is a significant, albeit initial, step towards the goal of predicting the quality of LoRa links over large areas and in a completely automated fashion.

2 THE LORA LINK IN THEORY

Obtaining a *long* communication range with *high power* radios (e.g. satellite), or a *short* communication range with *low power* radios (e.g. ZigBee) does not defy intuition. Obtaining a *long* communication range with a *low power* radio is less intuitive. Low-power long-range technologies, including LoRa, build upon the key concept of *link budget* to achieve this seemingly contradicting goal.

Wireless signals attenuate while travelling through free space. The link budget, i.e., the difference between the output power at the transmitter and the sensitivity at the receiver, determines the maximum amount of attenuation that a signal can tolerate. Once the signal strength falls below the sensitivity of the receiver, the information cannot be decoded. A simple way to increase the range of a signal is to increase the output power, but requirements from technical standards and IoT scenarios prohibit that; devices should use a very low power, to increase their lifetime. Thus, the only other option to increase the link budget (and range) is to improve the receiver sensitivity.

LoRa can achieve a high sensitivity: −140 dBm. To put this value in context, consider a WiFi receiver with a sensitivity of −90 dBm: the 50 dB difference means that LoRa can recover signals that are 10^5 times weaker. This property can increase significantly the range of a link. Taking into account that the maximum output power of WiFi and LoRa is 20 dBm and 14 dBm (in Europe), the resulting link budget is 110 dB and 154 dB, respectively. Using the well-known free-space path loss (FSPL) model for 2.4 GHz (WiFi) and 868 MHz (LoRa) signals, the aforementioned link budgets map to ranges of approximately 3 km and 1400 km.

The FSPL ranges are upper bounds. In practice, all radio links achieve much shorter ranges. For example, in *real open scenarios*, WiFi can reach distances of ∼90 m, and LoRa of ≥30 km [37]. This is approximately 3% of their free-space ranges (FSPL), still the long-range advantage of LoRa in *open spaces* is overwhelming. The problem arises when LoRa operates in urban or rural environments where the signal needs to travel through multiple obstacles. In those scenarios, the coverage of short-range signals, such as WiFi, could vary by a factor of 5; but the range of LoRa links can vary by a factor of 100 or more. This high variability is unique to low-power long-range links, and creates complex and highly heterogeneous coverage patterns that are hard to estimate.

3 THE BASELINE: OPEN-SPACE SCENARIOS

LoRa claims long ranges, but how long can they be in the wild? And under what conditions can they be achieved? A proper link analysis requires evaluating an important baseline: identifying the longest range by performing Line-of-Sight (LOS) experiments. Researchers are devising ingenious ways to analyze LoRa links in big open spaces, e.g., installing a gateway close to a coastline and carrying a node on a boat [37]. We use a high-altitude weather balloon. The latter was launched on March 15th, 2017 on a clear afternoon and was on-air for 3 hours covering a distance of 164.4 km. The balloon carried a GPS receiver and a LoRa device registered to TTN. The LoRa node in the balloon transmitted packets periodically with
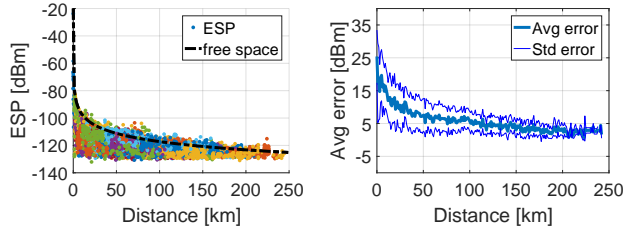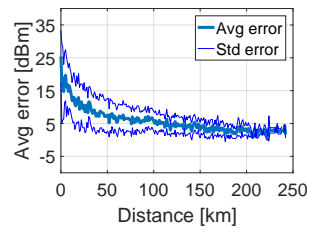
Figure 1: ESP vs. distance for all gateways.

Figure 2: Error between Friis model and measurements.

(a) $G_{id} = 7$ **(5 m)**   **(b)** $G_{id} = 79$ **(38 m)**   **(c)** $G_{id} = 116$ **(11 m)**
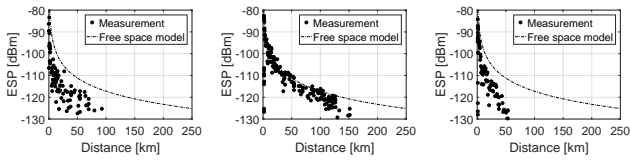
Figure 3: ESP vs. distance for selected gateways.

spreading factor[1] $SF = 7$, bandwidth $BW = 125$ kHz, and coding rate $CR = 4/5$.

For open spaces, the most accurate link model is given by the Friis equation:

$$P^{rx} = P^{tx} + G^{tx} + G^{rx} - FSPL \qquad (1)$$

where P^{rx} is the expected received power, also dubbed the Received Signal Strength Indicator (*RSSI*); P^{tx} is the transmission power; G^{tx} and G^{rx} are the transmitting and receiving antenna gains, all in dB(m) units. The last term is the free-space path loss:

$$FSPL = 20log(d) + 20log(f) - 27.55 \qquad (2)$$

where d is the distance in meters and f is the frequency in MHz. In our experiments, $P^{tx} = 14$ dBm and $f = 868$ MHz, and we assume a worst-case scenario with no antenna gains, $G^{tx} = G^{rx} = 0$.

Given that LoRa enables the reception of signals whose power is up to 20 dB under the noise floor, LoRa defines a metric called the *Expected Signal Power (ESP)* to obtain the energy of the signal. The ESP can be derived from the *RSSI* and *SNR* measurements (available from TTN) as per the following equation:

$$ESP_{[dBm]} = RSSI_{[dBm]} + SNR_{[dB]} - 10 \cdot log_{10}\left(1 + 10^{0.1\,SNR_{[dB]}}\right) \quad (3)$$

Figure 1 reports the ESP received by all stations and the Friis model. We collected 8578 packets from 141 TTN gateways; each color represents the samples received at a particular gateway. Figure 2 reports the average error ± the standard deviation between the Friis model and the measurements.

Our results show some interesting insights. First, in open spaces it is not uncommon for LoRa to attain very wide coverages. Several gateways provide a range between 50 and 200 km. Second, at shorter distances (<50 km) the error is quite high. This occurs because several samples fall significantly below the Friis model,

---

[1]$SF = 12$ provides the longest range, but also increases the on-air time of packets by a factor of 32. Thus, $SF = 7$ is a lower bound on the maximum range and provides more packet samples.

Table 1: Selected gateways with their integer ID (G_{id}), TTN ID (TTN_{id}), number of receptions (N_{rx}), connection duration (T_c), and connection distances from-to [km] (D_c).

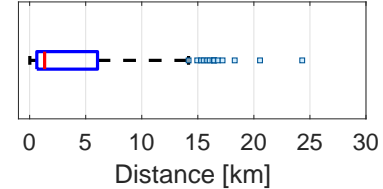| G_{id}	TTN_{id}	N_{rx}	T_c	D_c
7	eui-0000024b080602ed	98	1h 30'	0.01 - 93
79	eui-aa555a00080605b7	173	2h 45'	2 - 151
116	eui-fffeb827eb5d8d35	79	1h 13'	1 - 53

Figure 4: Variation of communication ranges for LoRa links.

which indicates the presence of paths with high attenuation. The attenuation is caused by the presence of obstacles, as we discuss later in detail. At longer distances, the error decreases. Some gateways consistently obtain long ranges, which indicates paths with low attenuation.

To gain a better understanding about the cause of these differences in coverage, we focus on 3 gateways, $G_{id} \in \{7, 79, 116\}$. Table 1 summarizes some relevant characteristics, and Figure 3 shows the ESP vs. distance. We can observe that $G_{id} = 79$ fits very well the Friis model, apart from some samples in the first kilometers. In contrast, the trend observed for $G_{id} \in \{7, 116\}$ presents stronger decay rates. The reasons for this different performance are the height of the gateway and the surrounding environment. The best gateway, $G_{id} = 79$, is placed very high (38 m), and its vicinity consist of open agricultural fields. In contrast, $G_{id} = 7$ is placed at only 5 m height and in an area surrounded by residential buildings, and $G_{id} = 116$ is placed at 11 m and surrounded by trees.

This baseline evaluation provides two important insights. First, the need to classify the attenuation of the surrounding environment (§5 and §6). Second, the need to use a model that considers height as an integral parameter of the propagation model (§7).

4 ASSESSING THE IMPACT OF OBSTACLES

In practice, LoRa will rarely be deployed in *open* spaces. In rural and urban environments, LoRa signals travel through several obstacles, e.g., foliage and buildings. *In those cluttered scenarios, a large link budget does not necessarily translate into a long coverage.* Depending on the obstacles that are encountered on a given direction, a large link budget can be consumed rather fast, leading to short ranges. For example, the link budget difference between WiFi and LoRa is 44 dB. This extra budget can stretch a link quite far in an open environment, but a brick wall can attenuate the signal by 10 dB; a few walls can make the LoRa link disappear within 100 m.

To capture the joint effect of large link budgets and obstacles on the variability of ranges, we collected 65821 samples from TTN Mapper [2]. The traces were collected over one year (June 2016 to June 2017) from 5 TTN gateways placed at different heights in various cities. The maximum communication range was 76 km, but
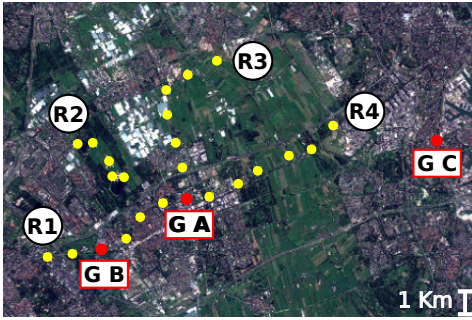
Figure 5: Measurement sites (yellow dots) along four routes (R1–R4) around TTN gateways (GA–GC).
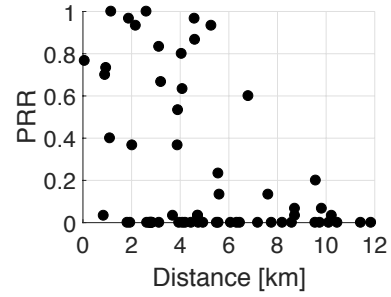

Figure 6: Link reliability (PRR) vs. distance for three gateways. LoRa provides only a single, long transitional region.

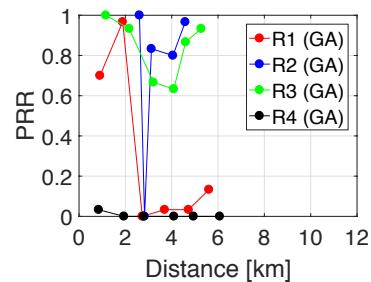Figure 7: PRR vs. distance for gateway GA.

98% of samples were within 20 km, with a median distance of 1.3 km and 25th and 75th percentiles at 0.6 and 6 km, respectively (Figure 4). Note that in urban environments, the range of LoRa links can have a remarkable variability; two orders of magnitude or more. This variance is not present on prior wireless technologies.

To analyze the effect of cluttered environments in more detail, we perform controlled experiments over a relatively large area. Based on these experiments, we provide a qualitative description of the impact of the environment on the reliability of links.

4.1 Experimental Setup

The user device is based on a Dragino LoRa shield v1.3 embedding a RF96 radio chip. The TX power is set to 14 dBm; the other LoRa parameters (SF, BW and CR) are set as in §3.

The experiments took place across urban and rural areas of Delft, The Netherlands. Measurements were taken every ~1 km along 4 routes departing in different directions from a TTN gateway (GA) chosen as reference (Figure 5). This gateway is placed inside a tall building, 62 m above ground. The shape of the routes is a compromise between the desire to follow a straight line and practical limitations, e.g., fences preventing access to farming fields. The tests were performed on June 7, 2017 (R1, R3) and June 12, 2017 (R2, R4). Overall, we selected 23 measurement sites, 6 for each route (except for R2), where the user device transmitted 30 packets, one every 40 s. These packets were received by two other gateways in range, GB and GC in Figure 5. For each received packet, the dataset contains the GPS position and the information provided by TTN: application and device IDs, packet payload, time of reception, frequency, modulation scheme, data rate, coding rate, RSSI, SNR, ID and GPS coordinates of the receiving gateway.

4.2 LoRa: A Long Transitional Region

Overall reliability. Figure 6 depicts the link reliability, i.e., the packet reception rate (PRR), vs. distance. We can observe that, for any distance under 6 km, we get a mix of link quality, from good to poor. For short-range technologies, like those commonly used in wireless sensor networks, the research community established three clear communication regions [46]: connected (links with reliability close to 1), transitional (links with reliability between 0 and 1), and disconnected (no links). Even though short-range radios have anisotropic coverages, the above described regions still

appear because, probabilistically, the propagation properties are the same in *all directions*; an environment does not change much within a couple hundred meters. Longer links, on the other hand, can see widely different environments depending on the direction they travel; a big open park or a series of houses yield very different ranges. On a kilometer scale, there is no clear connected (i.e., reliable) region. We then formulate the following insight:

The large budget of LoRa links does not yield homogeneous coverage, but a long "transitional region" with complex coverage.

This wide variability leads to highly irregular coverage patterns, like those later shown in Figure 19. We now look deeper into this phenomenon and qualitatively relate it to the type of land cover.

Reliability per gateway and route. We focus our analysis on gateway GA, as the data for GB and GC provide similar insights.

Figure 7 shows the PRR vs. distance for each route. For the gateway of interest (GA), R1 shows a short range (<3 km) because it unfolds only through buildings. R2 and R3 mostly traverse farming fields and exhibit better reception, with the exception of a single site in R2 (at ~3 km) with $PRR = 0$ further investigated in §6. R4 is an interesting case. Even though it also unfolds across farming fields, almost none of the packets are received at GA. We found out that GA is placed indoors and near the NW side of the building (i.e., pointing towards R1 in Figure 5). Therefore, the bulk of the building sits exactly in between the gateway and the measurement sites along R4, yielding a significant shielding. In our evaluation (§8) these disagreements will show as estimation errors.
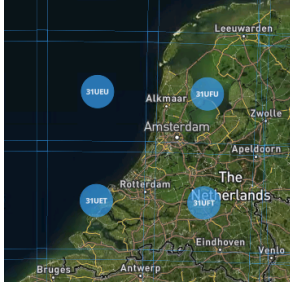
Figure 8: Sentinel-2 multispectral images for the area of interest (tiles 31UET, 31UFT, and 31UFU).

Table 2: Land-cover classes.

NLOS	BUILDING	buildings
	GREENHOUSE	greenhouse structures
	TREES	trees
LOS	FIELD	farming field or grassland
	SOIL	bare soil
	ROAD	streets, roads, and highways
	WATER	lakes and rivers

These preliminary tests highlight the value of remote sensing to acquire information about how cluttered is the landscape; we describe next the toolchain we develop towards this task.

5 EXTRACTING LAND COVER MAPS FROM MULTISPECTRAL IMAGES

We describe how we automatically extract land-cover classes from multispectral images acquired via the Sentinel-2 satellite constellation, a last-generation remote sensing system of the European Space Agency (ESA). Images are acquired at global scale with high spatio-temporal resolution: *i)* every 2–3 days at mid-latitudes *ii)* across 13 spectral bands in the visible, near-infrared, and short-wave infrared range, and *iii)* with a spatial resolution of 10–60 m, depending on the spectral band. These data can be downloaded from ESA archives free of charge.

We extract accurate land-cover maps by exploiting the spectral response of different land-cover classes and applying supervised classification techniques based on machine learning. We consider kernel based approaches, specifically pixel-based Support Vector Machines (SVM) [12, 23, 24, 31], due to their good properties, including high generalization capabilities, high classification accuracy, and relatively simple design through few control parameters.

Dataset. We perform classification on 3 multispectral images acquired on May 26, 2017. Each image (hereafter *tile*) covers a ground area of $\sim$100×100 km^2; together, they cover the full area (Figure 8) where the experimental data discussed in §4–§8 were collected.

Land-cover classes. We define the land-cover classes (Table 2) based on three criteria: *i)* presence in the target area *ii)* usefulness in characterizing LoRa links, and *iii)* possibility to discriminate them in the multispectral images [33, 36]. We further divide them in two groups, depending on whether they introduce significant attenuation (NLOS) or instead do not affect signal propagation (LOS).

***Automatic* classification.** Our system automatically maps each 10×10m^2 *pixel* in the images to the land-cover class that best matches
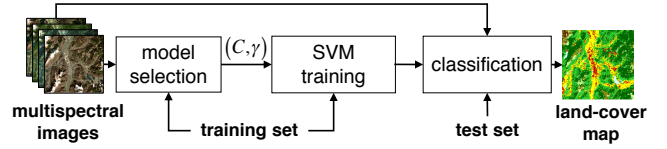

Figure 9: Automatic classification system.

a predefined criterion based on spectral features, i.e., the spectral values of the raw pixels, the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI) computed for each pixel.

As the problem involves multiple non-linearly-separable classes, it is solved via non-linear kernel methods. Specifically, we use SVMs with Radial Basis Function (RBF) kernel [13], along with the one-against-all (OAA) strategy, a state-of-art multiclass approach [9].

The toolchain (Figure 9) requires a *training set*, i.e., a small set of reference labeled pixels manually associated to a land-cover class via photo interpretation. This dataset is used for model selection and the training of the SVM in the learning phase of the classifier. Once the SVM is trained, it is applied to all considered tiles for *automatically* producing the land-cover map. We classify each tile independently after collecting tile-specific training (200 samples) and test (100 samples) sets for each land-cover class. The test samples are used for validating classification accuracy.

The *model selection* phase is crucial in determining application-specific optimal SVM tuning parameters (C, γ) to use in the other two phases. C is the regularization parameter determining the penalty for mis-classified samples [10, 43], while γ is the width of the RBF kernel [13]. The selection aims at *i)* accurately discriminating the classes, and *ii)* minimizing the expected generalization error. We perform on each tile a grid-search model selection based on 5-fold cross-validation [6], testing $C \in [100, 1000]$ in increments of 20 and $\gamma \in [0.1, 2]$ in increments of 0.1 [16]. The training set is randomly divided into 5 folds; the SVM is trained on 4 folds and validated on the 5th. This process is repeated 5 times for each (C, γ) pair, by swapping the training and validation folds, and the correspondent average classification accuracy is computed. The (C, γ) with the best cross-validated estimate of classification accuracy is selected and used in SVM training and classification, performed with standard tools and techniques and not discussed further.

Classification accuracy We evaluate the classification of our test set in terms of *i) Overall Accuracy (OA)*, the percentage of test pixels correctly classified *ii) Producer's Accuracy (PA)*, the percentage of correctly classified pixels for the given class *iii) User's Accuracy (UA)*, the percentage of correctly classified pixels computed w.r.t. the overall number of pixels associated to the given class. While OA focuses on the entire dataset, PA and UA are per-class metrics relating to the error of omission and commission, respectively.

Table 3 shows high ($\geq$ 92%) overall accuracy for all three tiles. This is visually confirmed by Figure 10, which compares a true color composition of the RGB bands from a portion of tile 31UET vs. the corresponding land-cover map. On the other hand, a few mis-classified pixels for BUILDING and ROAD can be noticed, e.g., along the highway. Indeed, the worst performance is for BUILDING

Table 3: Classification accuracy metrics, per tile and per class, estimated on the test samples.

	31UET			31UFT			31UFU		
	OA	PA	UA	OA	PA	UA	OA	PA	UA
BUILDING		92.5	96.3		70	77.7		86	92.4
GREENHOUSE		100	100		100	100		100	99
TREES		100	99.5		100	97		97	99
FIELD	98.3	99.5	100	92	95	96	95	99	94.2
SOIL		99.5	98.5		98	88.3		98	85.2
ROAD		97	94.2		80	82.5		86	98
WATER		100	100		100	100		100	99

and ROAD on tile 31UFT (Table 3); by comparing their PA and UA we can deduce that these two classes are often confused, likely due to relatively similar spectral signatures of the related materials. A similar yet less marked trend occurs in tile 31UFU. Confusing BUILDING and ROAD pixels may seem concerning, as these classes affect communication differently. However, this (still acceptable) error occurs sparsely on 10×10 m^2 pixels (Figure 10b), a very small scale w.r.t. the long (km) range of LoRa. Therefore, the error does not significantly affect the analysis in the next sections.

6 EXPLOITING LAND-COVER KNOWLEDGE

Armed with the land-cover classification enabled by the remote sensing toolchain, we now revisit the empirical observations in §4; instead of intuitive and qualitative arguments about landscape characteristics, we provide *quantitative* and detailed *evidence*. For each of the measurement sites (Figure 5), we use our automated classification (§5) to retrieve the sequence of land-cover classes present on each link between the transmitting device and receiving gateway, and compute the occurrence (in percentage) of each class.

Measurement routes vs. land cover. Figure 11 shows the results for all 7 classes considered. Each plot focuses on a route, with one subplot for each measurement site, ordered in increasing distance w.r.t. the gateway. We omit R4 as it exhibits very poor reception, as discussed earlier.
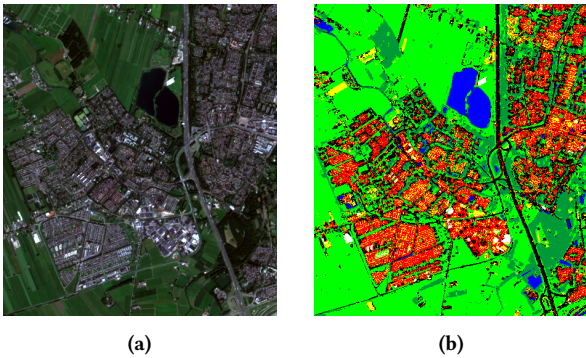

■Water ■Field ■Soil ■Road □Greenhouse ■Building ■Trees

(a) (b)

Figure 10: True color composition of red, green, and blue spectral bands (a) and classification map (b) of a 4×5 km^2 subset of the area of interest.

Reliability vs. land cover. In Figure 7, R1 has the shortest range. Figure 11a shows that this route has a heavy presence of the high-attenuation BUILDING class, 32% to 61% depending on the site. R3, on the other hand, has the longest and most reliable link. Along this route, FIELD and ROAD prevail, which can be considered open spaces. FIELD covers 38–62% of each link, and ROAD covers 10–26% (Figure 11c). R2 has an intermediate quality, with a comparable presence of high- (BUILDING and TREES) and low-attenuation (FIELD and ROAD) classes (Figure 11b).

These quantitative data support our previous considerations about reliability observed in Figure 7. The trends of PRR vs. distance diverge significantly on R1 and R3 at 3 km. This is reflected in the predominance of different land-cover classes along the links, i.e., BUILDING on R1 and FIELD on R3 (Figure 11). Similar considerations can be drawn by comparing the PRR for distances ≥ 3 km of R1 and R2 and observing that the two routes are dominated by BUILDING and FIELD, respectively.

As these two classes (BUILDING and FIELD) are predominant in our dataset, we investigate further the correlation between their presence and the link performance. Figure 12 helps visualizing the impact of the land cover on the signal attenuation. We differentiate the data obtained in §4.1 between measurement points whose links are dominated by BUILDING and FIELD classes. We observe that the ESP, and consequently[2] the PRR, have a noticeably better performance for FIELD than for BUILDING.

Impact of obstacle height. Figure 7 shows that one measurement site along R2 (2.8 km) exhibits no packet reception from GA, despite being FIELD-dominated as the the rest of the route. We observe that the composition of land-cover classes in the *vicinity* of the transmitter presents peculiar characteristics in this site. Figure 13a shows that FIELD is predominant on the overall path (top), while TREES is present for 41% along the first kilometer (center) and 100% in the first 50 m (bottom). Trees are exactly in front of the TX device (held at 1.5 m height), and completely obstruct the line of sight towards GA. It is interesting to compare against the next measurement site along R2 (3.1 km), where PRR = 0.8 (Figure 7). Figure 13b shows that in this case FIELD remains predominant at all distances, including next to the TX device.

We therefore observe that, in addition to the height of transmitting and receiving devices (§3), the height of obstacles belonging to NLOS classes plays a role in determining link quality. As mentioned in §1, passive remote sensing does not capture obstacle height, which could instead be derived automatically with, e.g., LiDAR, whose data is however less pervasive (and much more expensive) than multispectral images. In §8, we propose a simple geometric approach to attribute more weight to the obstacles in the vicinity of the end-device. Alternately, other sources of information could be exploited, e.g., cadastral maps for BUILDING or forestry surveys for TREES. However, even in absence of these, the detailed horizontal structure extracted from high-resolution land-cover maps already enables accurate estimates, as shown in §8.

---

[2]The PRR is a function of the ESP (or RSSI) and it depends on the response of the radio receiver. We do not delve into those details in this paper as they are not relevant for our discussion.
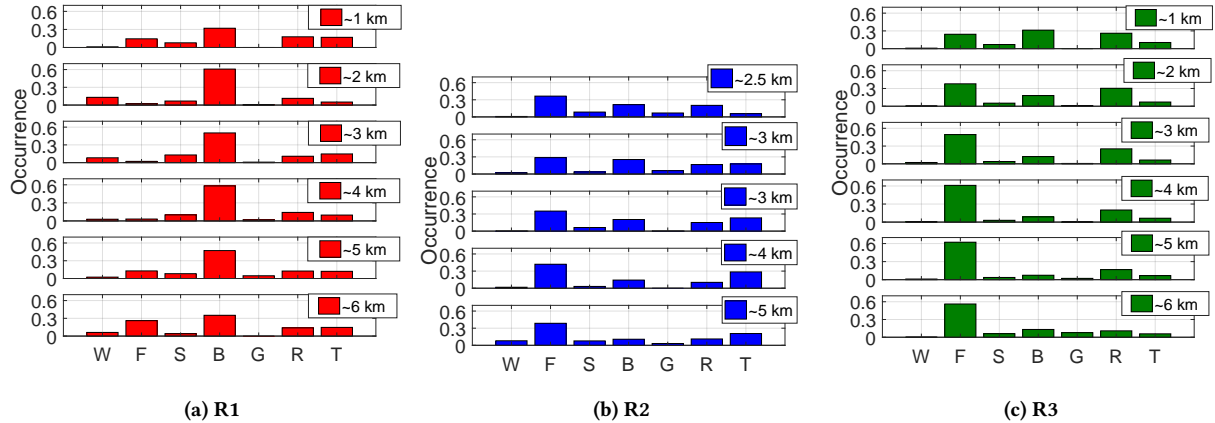
(a) R1   (b) R2   (c) R3

Figure 11: Occurrence of the land-cover classes in Table 2 along the links between the user device and gateway GA along routes R1–R3. Each subplot concerns a measurement site at the distance shown from GA.
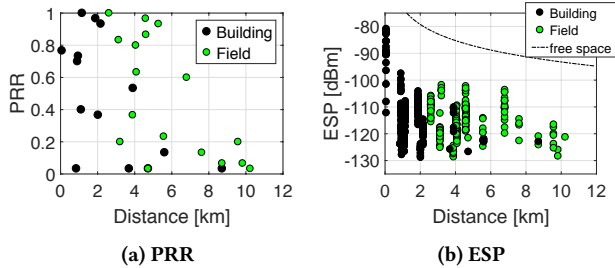

(a) PRR   (b) ESP

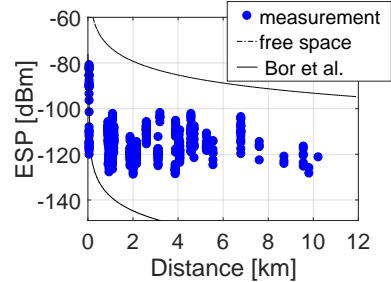Figure 12: Reliability and power vs. distance for BUILDING and FIELD dominated links.

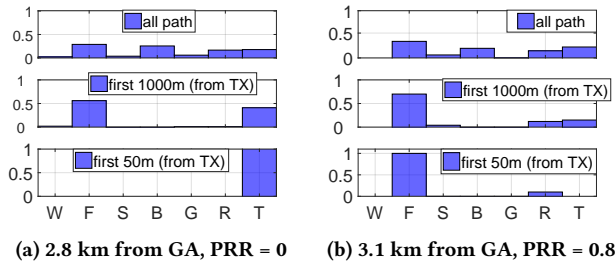(a) 2.8 km from GA, PRR = 0   (b) 3.1 km from GA, PRR = 0.8

Figure 13: Occurrence of land-cover class at different vicinities from the TX device, for two peculiar sites on R2.

7 MODELING THE IMPACT OF LAND COVER

It is now quantitatively evident that a different combination of land cover characteristics results in a very different packet reception, due to differences in signal attenuation. Can we *predict* these trends with the land cover knowledge distilled by the automated toolchain relying on satellite images? We provide a positive answer in this section, corroborated by the validation on real-world data in §8.

Figure 14: ESP vs. distance for measurements (dots), free-space model (dashed line) and Bor's model (solid line).

7.1 A Model that Needs Measurements

In §3, we discussed the free-space model (Eq. 1). That model only considers the attenuation in free space, FSPL (Eq. 2). Bor et al. [7] builds on top of a more realistic model called log-normal shadowing model [39] to analyze LoRa links:

$$PL[dB] = PL(d_0) + 10 \cdot n \cdot log_{10}\left(\frac{d}{d_0}\right) + X_\sigma \qquad (4)$$

$$P^{rx} = P^{tx} + G^{tx} + G^{rx} - PL[dB] \qquad (5)$$

where d is the distance from the transmitter, $PL(d_0)$ is the path loss at a known reference distance d_0, n is the path loss exponent of the environment and σ is the standard deviation of a zero-mean Gaussian random variable X. This model captures the attenuation (path loss) of the environment, but it requires empirical data. Based on an indoor building environment, Bor et. al. perform measurements and estimate these parameters as $PL(d_0) = 127.41$ dB, $n = 2.08$, and $\sigma = 3.57$ at a reference distance $d_0 = 40$ m.

Figure 14 compares the expectation (mean) of the Bor model, the free-space model, and all the ESP measurements for the data we gathered in §4.1. The free space equation (Eq. 1), accurately models the behavior of LoRa links in free space, but it *over*estimates the received power by ~20 dBm, on average, because it does not consider the effect of obstacles. The Bor model, on the other hand,

Table 4: Least-square estimate of path loss exponent n, standard deviation σ of the gaussian random variable X_σ and average error between measurements and fitted model.

subset	n	σ	avg err [dB]
GA NLOS	3.34	3.63	5.21
GB NLOS	3.89	6.64	7.22
GA LOS	3.11	3.39	5.79
GC LOS	2.84	3.39	5.14

(a) Free-space and Bor models. (b) Okumura-Hata model.

Figure 15: ESP vs. distance and fitting curves for each gateway and land-cover macro-class.

severely *under*estimates signal strength. This occurs because the this model is suited solely for the particular environment where it was trained, as stated by the authors, and that environment is particularly harsh. Considering that their $PL(d_0)$ is 127.41 dBm at 40 m, and the LoRa budget is 140 dB, the link has already lost most of its budget at that short distance.

The performance of the log-normal shadowing model, i.e., the basis of the Bor model, can improve significantly if we train it with the data we gathered. To obtain a more stable and accurate path loss reference $PL(d_0)$ in Eq. (4), we use a distance d_0=1 m [19, 39]. With 200 samples we obtain a $PL(d_0)$ = 23.9 dB with a standard deviation of 1.1 dB. To obtain the n and σ propagation parameters, we use a least mean square approximation based on Eq. (4) and (5). The parameters we use are P^{tx} = 14 dBm for the transmission power, and $G^{tx} = G^{rx} = 2$ dBi for the antenna gains[3].

Table 4 reports the results we obtain for the path loss parameters along with the average error in dB between measurements and the fitted model. Hereafter, we group the land-cover classes according to their effect on communication links, i.e., into the LOS and NLOS macro-classes of Table 2. The curve fitting is done separately for each gateway/class combination. We omit the combinations GB/LOS and GC/NLOS due to the very few measurements available. We observe that n is much larger in NLOS than in LOS ([3.34, 3.89] vs. [2.84, 3.11]), coherently with the stronger attenuation induced by the former class. Moreover, for the same class, n increases as the gateway height decreases, e.g., n = 3.34 for GA/NLOS (62 m) vs. n = 3.89 for GB/NLOS (6 m). The value of σ and the average error are generally ~3 dB and ~5 dB, respectively, except for GB/NLOS. This is due to the high variability of the ESP measured in the site next to GB (Figure 5), caused by dynamic components (e.g., cars) across this very short link (67 m).

The results can be visually evaluated in Figure 15a, where each fitting curve is represented with the color of the corresponding measurement. The chart shows that trends are captured well. Interestingly, the fitting curve for GB/NLOS is the one best approximated by Bor's model, as the land cover in the former is similar to the one where the latter model was derived.

7.2 A General Model

The previous section reasserted the impact of land-cover on communication, but the model requires training data. Instead, our goal is to derive *a priori* accurate estimates of the expected received power, based on knowledge of the land cover and an appropriate general model. Further, our results also highlighted the interplay of

---

[3]GA mounts a typical half-wave dipole antenna; we assume that also GB and GC mount an antenna with similar characteristics and gain.
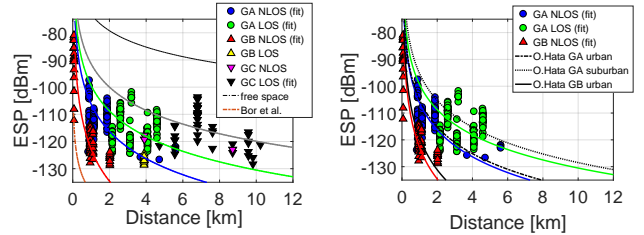
landscape with the gateway height, which must therefore be taken into account by model estimates.

We tackle both concerns by using the Okumura-Hata model [21, 34, 39], widely applied in the context of cellular communications. The key advantage of this model is that it provides equations that can be selected based on the properties of the surrounding environment. The model only needs as input the type of environment, which is provided by our automated tool. The Okumura-Hata model relates the height of transmitter and receiver h_m and h_b (m), their distance d (km), and transmission frequency f (MHz) based on the following equation determined empirically:

$$L_U[dB] = 69.55 + 26.16 \, log_{10}(f) - 13.82 \, log_{10}(h_b)$$
$$- a(h_m) + (44.9 - 6.55 \, log_{10}(h_b)) \, log_{10}(d) \quad (6)$$

This basic model is then adapted with a correction factor $a(h_m)$ depending on the environment where the communication link is placed. Considering that we have two land-cover macro-classes (Table 2), we use the two equations (urban and suburban) that best match from the Okumura-Hata set:

- *Urban model.* If our automated toolchain determines that the link path is dominated by the NLOS class, we use the correction factor below for Eq. (6):

$$a(h_m)[dB] = (1.1 \, log_{10}(f) - 0.7)h_m$$
$$- (1.56 \, log_{10}(f) - 0.8) \quad (7)$$

- *Suburban model.* If our automated toolchain determines that the link path is dominated by the LOS class, we use the following equation:

$$L_{SU}[dB] = L_U - 2 \left(log_{10} \left(\frac{f}{28} \right) \right)^2 - 5.4 \quad (8)$$

In our case, the transmitter height is h_m = 1.5 m, while the receiver height is h_b = 62 m for GA and h_b = 6 m for GB. The height of GC is unfortunately not available from TTN. Figure 15b compares the estimates obtained in this way with the fitting described in §7.1. The curves are very close, with a difference ~3 dB; the average difference is 1.13 dB for GA/NLOS (urban), 3.33 dB for GA/LOS (suburban) and 3.36 dB for GB/NLOS (urban).

This result is significant under two respects. On one hand, it shows that the Okumura-Hata model can accurately capture the characteristics of LoRa communication signals; to the best of our knowledge this result is hitherto unreported in the LoRa literature. We further validate this result to a larger scale in the next section. On the other hand, the Okumura-Hata model requires knowledge
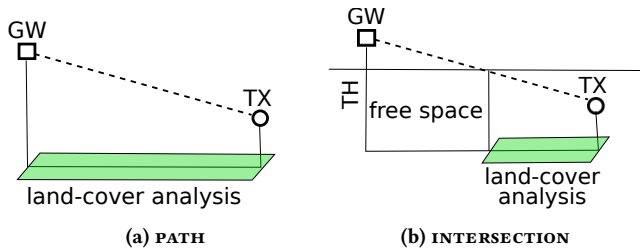
Figure 16: Land-cover analysis (a) on the whole link path (PATH) and (b) on the link portion that intersects objects given a threshold TH on their average height (INTERSECTION).

about the environment surrounding the links being estimated. Nevertheless, our results show that the toolchain described in §5 can be effectively used to *easily and automatically* select the appropriate variant of the Okumura-Hata model based on land cover classes, therefore greatly simplifying its practical application.

8 VALIDATION "IN THE WILD"

In this section, we validate the approach presented in §7.2 for the automatic estimation of ESP in LoRa links. We compare our predictions against real traces from TTN. Our validation *i)* predicts the expected ESP by exploiting the combination of the Okumura-Hata model and remote sensing data, and *ii)* shows quantitatively that LoRa communication range is strongly not isotropic.

8.1 Dataset & Approach

The TTN dataset was collected through the TTN Mapper application. The gateways and end-devices provide GPS coordinates. This data set is different from the one used in §4.1. We consider data collected over one year, from June 2016 to June 2017, for 5 gateways in The Netherlands; the number is limited by difficulties in retrieving information about the gateway height. Our estimation approach targets communication in *outdoor* scenarios, therefore we consider only traces associated to end-devices whose height is ≤2m; devices at higher altitudes are likely transmitting from inside buildings. We consider 8642 traces overall. The number of traces for each gateway is reported in Table 5 together with the gateway ID and height.

As stated in §7.2, our automated tool needs to find the dominant land cover to select between the urban (NLOS) or suburban (LOS) equations. We consider two prediction approaches to estimate the dominant land cover, as shown in Figure 16: *a)* PATH (Figure 16a), where we consider *all* the land cover between the end device and the gateway, and *b)* INTERSECTION (Figure 16b), where we consider only the land cover surrounding the end-device. The INTERSECTION approach is considered to capture in a better manner the effect of obstacles in the vicinity of the end-device, as discussed in §6.

8.2 Estimation Accuracy of ESP

We quantitatively evaluate the prediction accuracy of PATH and INTERSECTION by comparing each per-link estimate of ESP against the corresponding real value. For INTERSECTION, the threshold TH

Table 5: Gateways: integer identifier (G_{id}), TTN identifier (TTN_{id}), height [m] (H_{gw}) and number of samples (#samples) for end-device height ≤2m.

G_{id}	TTN_{id}	H_{gw}	#samples
1	eui-0000024b080e015d	100	3010
7	eui-aa555a000806047a	79	1235
8	eui-aa555a00080605b7	38	3619
11	eui-0000024b080e015b	32	365
13	eui-0003ffff1d09ce86	23	413

Table 6: Estimation error on ESP: average [dBm] (avg) and standard deviation [dBm] (stdd) for PATH, INTERSECTION (TH = 10 m), as well as free-space and Bor models.

G_{id}	PATH		INTERSECTION		free-space		Bor	
	avg	stdd	avg	stdd	avg	stdd	avg	stdd
all	8.73	6.67	8.71	6.62	32.24	10.61	33.53	10.71
1	9.73	8.01	9.64	7.69	25.66	9.93	40.58	10.59
7	7.11	5.73	6.53	5.25	32.18	6.58	33.91	6.63
8	7.90	5.36	8.03	5.56	35.68	8.36	29.92	8.31
11	10.14	5.55	9.67	5.25	42.58	8.43	22.97	8.46
13	12.28	7.22	13.63	7.69	43.22	8.28	21.97	8.17

represents the average height of objects (e.g., buildings) expected in the target area. We set TH = 10 m to represent the three-storey buildings commonly found in the areas targeted by the evaluation.

The statistics of the errors (average and standard deviation) are reported in Table 6 for the whole dataset and for each gateway; Figure 17 offers a visual representation. These results provide some important insights. First, the average error for PATH and INTERSECTION is ≤9 dB; further, it is slightly better for INTERSECTION, as discussed later. This is an acceptable error considering that the attenuation of a single brick wall is around that value. Second, the error of our approach is far lower than the ones obtained for the free-space and Bor models, whose errors are ≥32 dB. This is a significant difference. To put these values in context, consider that in urban environments, a 10 dB difference roughly doubles the range of a link [22]. For example, if the range of a real link is 1 km in an urban environment, our method would provide a range between 0.5 and 2 km, the free-space model would provide a range above 8 km, and the Bor model would provide a range below 125 m. Third, our method does not *consistently* over- or under-estimate the ESP. Figure 17 shows that free-space consistently overestimates the ESP, while Bor consistently underestimates it. With our approach, the *average* ESP estimations of all samples deviate by only 1.1 dB from the real measurements, while the free-space and Bor estimations deviate by more than 30 dBm (Figure 18). This balanced oscillation around the real mean is an important property for channel models, since it avoids the under- or over-provisioning of gateways during deployment.

8.3 Predicting the Gateway Coverage

The per-link estimates of ESP can be exploited to predict the *non-isotropic* radio coverage of LoRa gateways. Based on the map in Figure 20a, which has a 6×6 km² area, we used our automated tool to generate the coverage maps for a gateway positioned in the middle of the map (G_8), as shown in Figure 19. In this figure the same colorscale of [−40, −140] dBm is used for all models. All the
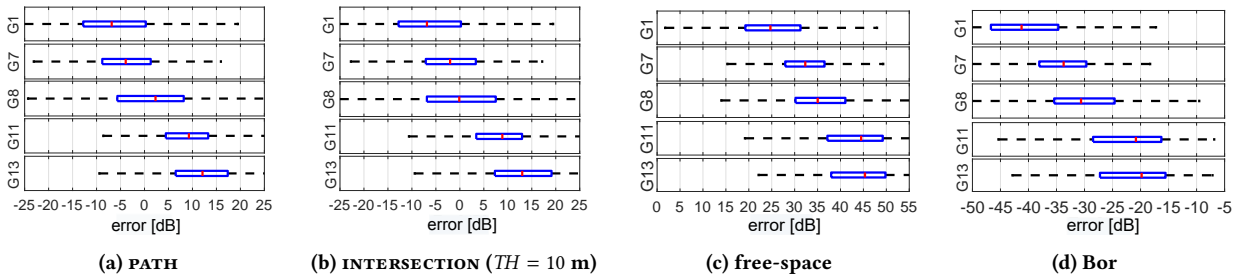
(a) PATH   **(b) INTERSECTION ($TH = 10$ m)**   **(c) free-space**   **(d) Bor**

Figure 17: Per-gateway statistics of the estimation error (median value, 25th and 75th percentile) for PATH, INTERSECTION, as well as free-space and Bor models, for end-devices with height $\leq$ 2m.

predicted values < -140 dBm are saturated at the bottom of the scale (i.e., black); this happens for most of Bor predictions, where we only plot the mean ESP for clarity.

Considering that the samples we took are sparse (red dots in Figure 20a), we cannot obtain a complete ground truth image to compare with our estimations, and we only provide per-sample values for the ground truth in Figure 20b. Nevertheless, our coverage evaluation provides some interesting insights. First, both PATH and INTERSECTION improve the coverage estimation w.r.t. free-space (too optimistic) and Bor (too conservative). Using the free-space model in a real deployment planning would leave several areas without coverage, while using the Bor model would lead to an unnecessarily dense deployment of gateways. Second, the gateway coverage is *strongly non-isotropic*, showing significant variability in space depending on both the direction and the distance. This anisotropic behavior is better captured with the INTERSECTION approach than with the PATH one. By comparing the BUILDING (gray) and FIELD (green) areas in Figure 20a with the coverage provided by the PATH and INTERSECTION approaches, we can notice that the INTERSECTION approach captures in a better manner the "pockets" of BUILDING environments, compared to the PATH approach. This helps explain the better performance of INTERSECTION in Table 6. We hyphotesize that the reason the difference is small in our data is because the samples are very sparse; a denser data set would likely amplify the better performance of INTERSECTION.
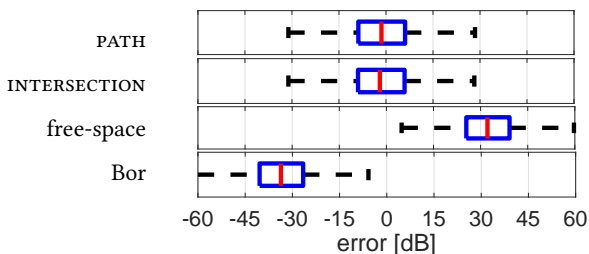

Figure 18: Overall estimation error (median, 25th and 75th percentile) on samples with devices $\leq$ 2m height: PATH, INTERSECTION, free-space and Bor.

9 RELATED WORK

Characterizing LoRa links: Empirical studies. Realistic coverage estimation is crucial to provide connectivity guarantees, satisfactory service, and resource optimization (e.g., number of gateways to deploy). This has led to numerous empirical studies on LoRa coverage ranges, but unfortunately with equally many different findings [7, 15, 25, 28, 32, 37, 45]. Real-world observations show both a significant gap w.r.t. theoretical expectations [26, 27] and significant variability depending on the specific environment at hand. For example, Centenaro et al. [15] estimated the number of gateways required to enable city-wide LoRaWAN coverage in Padova (Italy). They experimentally obtained a coverage estimate of 2 km in a high-building area. Bor et al. [7] observed a range of 2.6 km in rural areas and of 100 m in a built-up environment, while in the central business district of Glasgow (Scotland) the communication range was 1–20 km [45]. In Hyde Park (London), Kartakis et al. [28] achieved 2.4 km with semi line-of-sight conditions, but 450 m in a built-up area. Petajajarvi et al. [37] reported a range of 15–30 km in a urban/maritime environment. Instead, a maximum range of 90 m was measured in a mountain forest with dense vegetation [25].

Besides the environment, the PHY-level configurable settings of LoRa also determine different trade-offs between range, consumption and data rate. For example, a quantitative assessment of the impact of PHY settings on PRR is presented in [14]. The authors observed through experiments indoor, outdoor, and underground *i)* a drastic decrease of link reliability at high temperatures, and *ii)* the benefits of using energy-hungry PHY settings to increase link quality, instead of relying on retransmission schemes.

Characterizing LoRa links: Modeling studies. Georgiou and Raza [18] evaluated the coverage probability by exploiting stochastic geometry and observed that when collisions occur between packets with the same SF, the stronger signal can be successfully received if it is at least 6 dB stronger than any other. In the paper, the expected received signal power was estimated by considering path loss attenuation, assuming a path loss exponent equal to 2.7 in suburban environments and 4 in urban environments. However, these approximations may not be representative enough for real scenarios, in that the attenuation as a function of distance usually varies also within environments belonging to the same category and it is not isotropic in practice, due to the intrinsic non-homogeneity of the propagation environment. Voigt et al. [44] analyzed the impact of inter-network interference due to independent LoRa networks
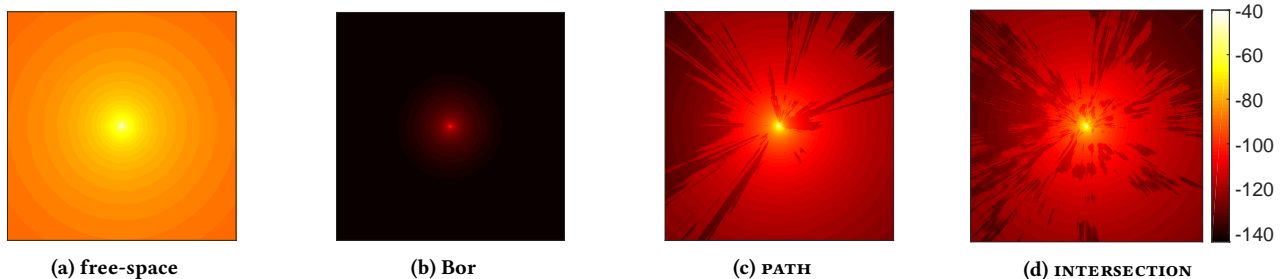
(a) free-space   (b) Bor   (c) PATH   (d) INTERSECTION

Figure 19: ESP prediction on a 6×6 km^2 area centered in G_8.

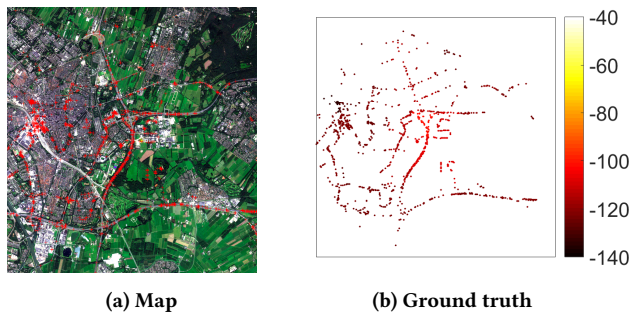(a) Map   (b) Ground truth

Figure 20: Deployment map and ground truth.

operating over the same deployment area. In their simulations, the best improvement was obtained by placing additional gateways to ensure that all devices are in reach of at least one of them, under the assumption of circular (i.e., isotropic) coverage. However, in real scenarios, the desired interference mitigation can be guaranteed only by properly accounting for the non-uniform spatial coverage that gateways can provide.

Although these prior works present valuable insights on link performance, they do not provide a systematic understanding about how LoRa behaves in real conditions. The development of models and tools that consider these conditions are therefore needed to provide realistic coverage estimates.

An approach toward such a model is to use machine learning to model connectivity of (low-power) wireless deployments, e.g. as in [35], where topographic and vegetation features are taken into account. However, the training of their algorithm requires the collection of a significant amount of in-field connectivity measures in the target deployment environment, difficult to obtain for long-range links. In this paper, we have therefore combined techniques from remote sensing with machine learning to replace on-site measurements with an automated approach based on satellite information.

Coverage planning in cellular networks. The idea of using a digital representation of the target area for coverage planning has been exploited in the context of cellular networks, mostly in combination with deterministic radio models based on ray-tracing [17, 20, 30]. These models are very precise but computationally complex, and require highly accurate 3D city models [17, 20] usually

not available for free. Some commercial products adapted these techniques to LoRa. For instance, the provider of 3D city maps Siradel [3] exploits the Volcano ray-tracer for LoRa coverage predictions; however, detailed descriptions of the techniques employed are generally not available. In contrast, our methodology is based on publicly available satellite images and yields good estimates without requiring full 3D maps.

Alternately, cellular coverage planning also employs empirical or physical radio models, e.g., Okumura-Hata or Walfisch-Ikegami [30], whose accuracy is often improved by correction factors accounting for the "clutter type" (e.g., urban, rural, or forest). These factors are determined via extensive measurements and on-purpose planning tools that make use of digital cartography, e.g., terrain models and clutter classes [30, 38]. Vendors and operators typically rely on their own proprietary tools, often patented [40, 42], or on commercial tools [4, 5], some of which now support LoRa. However, again, information about how these tools derive and use digital cartography is not available, including their combination with radio models.

In summary, motivated by the fact that on-site experiments are costly and their result difficult to generalize, we have combined two fields—radio path loss modeling and remote sensing—to develop a novel tool that uses *i)* multi-spectral images, *ii)* a machine-learning based classification of the environment based on those images, and *iii)* the right path loss model per environment type, to estimate LoRa coverage in an automated, low-cost manner.

10   CONCLUSIONS AND OUTLOOK

This paper presents a novel approach to estimate the quality of LoRa links in a fully automated way over the geographical scale typical of this radio technology, based on two key elements. First, a remote sensing toolchain based on freely available multispectral images from satellites, exploits land-cover classification techniques to enable the automated analysis of the landscape characteristics on a per-link basis. Second, the actual estimation of expected received power is achieved via the Okumura-Hata model, hitherto largely neglected by LoRa literature, configured—again, automatically— based on land-cover knowledge. Our validation on 8000+ samples from the TTN network confirms that our approach yields signal power estimates considerably more accurate than popular channel models for LoRa.

This work is a first, albeit crucial, stepping stone towards the goal of accurately predicting the quality of LoRa links; several aspects are therefore not yet addressed. For instance, we focus on modeling

the *outdoor* environment, implicitly assuming that gateways are unencumbered. This is a problem when gateways are instead placed *indoor*, as discussed in §4, as walls induce a significant attenuation near the gateway. However, in principle, this attenuation could be modeled (or measured) and factored as a correction of model. Similarly, and more importantly, we do not consider weather conditions (e.g., rain or fog) or physical parameters (e.g., temperature and humidity) known to affect wireless links. These dynamic parameters are significantly harder to model than the static features we focused on; moreover, empirical studies for LoRa are scarce, let apart reliable models. Therefore, in general the estimates returned by our toolchain should be considered as an *upper bound* w.r.t. the conditions one may expect in reality.

On the other hand, these upper bounds are *much* more accurate and spatially fine-grained than popular channel models for LoRa, as Figure 19 clearly shows—yet they are derived *entirely automatically*. These two aspects, accuracy and automation, hold the potential for significant practical impact, in that LoRa gateways, unlike cellular base stations deployed by telco operators, are often deployed by public bodies (e.g., smart cities) and individual citizens (e.g., as in the TTN network). We argue that our contribution is key in enabling the tools necessary to simplify a truly decentralized and grass-root deployment of the future IoT infrastructure, while at the same time ensuring its reliability and performance.

REFERENCES

[1] [n. d.]. http://scihub.copernicus.eu/.
[2] [n. d.]. https://ttnmapper.org/.
[3] [n. d.]. https://www.siradel.com/portfolio-item/alliance-lora/.
[4] [n. d.]. https://www.forsk.com/atoll-lpwaiot.
[5] [n. d.]. https://www.teoco.com/products/planning-optimization/asset-radio-planning/.
[6] A.I. Belousov, S.A. Verzakov, and J. Von Frese. 2002. A flexible classification approach with optimal generalisation performance: support vector machines. *Chemometr. Intell. Lab. Syst.* 64, 1 (2002).
[7] M. Bor, U. Roedig, T. Voigt, and J. M. Alonso. 2016. Do LoRa low-power wide-area networks scale?. In *Proc. of MSWiM*.
[8] M. Bor, J. E. Vidler, and U. Roedig. 2016. LoRa for the Internet of Things. *Proc. of EWSN*.
[9] L. Bottou, C. Cortes, J. S. Denker, H. Drucker, I. Guyon, L. D. Jackel, Y. LeCun, U. A. Muller, E. Sackinger, P. Simard, et al. 1994. Comparison of classifier methods: a case study in handwritten digit recognition. In *Proc. of IAPR*, Vol. 2. IEEE.
[10] C. J.C. Burges. 1998. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* 2, 2 (1998).
[11] J. B. Campbell and R. H. Wynne. 2011. *Introduction to remote sensing*. Guilford Press.
[12] G. Camps-Valls and L. Bruzzone. 2005. Kernel-based methods for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* 43, 6 (2005).
[13] G. Camps-Valls and L. Bruzzone. 2009. *Kernel methods for remote sensing data analysis*. John Wiley & Sons.
[14] M. Cattani, C. A. Boano, and K. Römer. 2017. An Experimental Evaluation of the Reliability of LoRa Long-Range Low-Power Wireless Communication. *J. Sens. Actuator Netw.* 6, 2 (2017).
[15] M. Centenaro, L. Vangelista, A. Zanella, and M. Zorzi. 2016. Long-range communications in unlicensed bands: The rising stars in the IoT and smart city scenarios. *IEEE Wirel. Commun.* 23, 5 (2016).
[16] B. Demir, F. Bovolo, and L. Bruzzone. 2013. Updating land-cover maps by classification of image time series: A novel change-detection-driven transfer learning

[17] approach. *IEEE Trans. Geosci. Remote Sens.* 51, 1 (2013).
[17] T. Fugen, J. Maurer, T. Kayser, and W. Wiesbeck. 2006. Capability of 3-D Ray Tracing for Defining Parameter Sets for the Specification of Future Mobile Communications Systems. *IEEE Trans. Antennas Propag.* 54, 11 (2006), 3125–3137.
[18] O. Georgiou and U. Raza. 2017. Low power wide area network analysis: Can LoRa scale? *IEEE Wireless Commun. Lett.* 6, 2 (2017).
[19] A. Goldsmith. 2005. *Wireless communications*. Cambridge University Press.
[20] E. Greenberg and E. Klodzh. 2015. Comparison of deterministic, empirical and physical propagation models in urban environments. In *Proc. of IEEE COMCAS*.
[21] M. Hata. 1980. Empirical formula for propagation loss in land mobile radio services. *IEEE Trans. Veh. Technol.* 29, 3 (1980).
[22] K. T. Herring, J. W. Holloway, D. . Staelin, and D.l W. Bliss. 2010. Path-loss characteristics of urban wireless channels. *IEEE Trans. Antennas Propag.* 58, 1 (2010), 171–177.
[23] C. Hsu and C. Lin. 2002. A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Netw.* 13, 2 (2002).
[24] C. Huang, L.S. Davis, and J.R.G. Townshend. 2002. An assessment of support vector machines for land cover classification. *Int. J. Remote Sens.* 23, 4 (2002).
[25] O. Iova, A. L. Murphy, G. P. Picco, L. Ghiro, D. Molteni, F.ederico Ossi, and F. Cagnacci. 2017. LoRa from the City to the Mountains: Exploration of Hardware and Environmental Factors. In *Proc. of MadCom*.
[26] Recommendation ITU-R P ITU-R. 2016. 525-3, Calculation of free-space attenuation. *International Telecommunication Union* (2016).
[27] Recommendation ITU-R P ITU-R. 2017. 837-7, Characteristics of precipitation for propagation modelling. *International Telecommunication Union* (2017).
[28] S. Kartakis, B. D Choudhary, A. D. Gluhak, L. Lambrinos, and J. A. McCann. 2016. Demystifying low-power wide-area communications for city IoT applications. In *Proc. of WiNTECH*.
[29] T. Lillesand, R. W. Kiefer, and J.onathan Chipman. 2014. *Remote sensing and image interpretation*. John Wiley & Sons.
[30] A. R. Mishra. 2007. *Advanced cellular network planning and optimisation: 2G/2.5 G/3G... evolution to 4G*. John Wiley & Sons.
[31] G. Mountrakis, J. Im, and C. Ogole. 2011. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* 66, 3 (2011).
[32] B. Moyer. 2015. Low power, wide area: A survey of longer-range IoT wireless protocols. *Electronic Engineering Journal* (2015).
[33] A. Novelli, M. A Aguilar, A. Nemmaoui, F. J. Aguilar, and E. Tarantino. 2016. Performance evaluation of object based greenhouse detection from Sentinel-2 MSI and Landsat 8 OLI data: A case study from Almería (Spain). *Int. J. App. Earth Obs. Geoinf.* 52 (2016).
[34] Y. Okumura. 1968. Field strength and its variability in VHF and UHF land-mobile service. *Rev. Elec. Comm. Lab.* 16, 9 (1968).
[35] C. Oroza, Z. Zhang, T. Watteyne, and S. D Glaser. 2017. A Machine-Learning Based Connectivity Model for Complex Terrain Large-Scale Low-Power Wireless Deployments. *IEEE Trans. Cogn. Commun. Netw.* (2017).
[36] M. Pesaresi, C. Corbane, A. Julea, A. J. Florczyk, V.asileios Syrris, and P. Soille. 2016. Assessment of the added-value of Sentinel-2 for detecting built-up areas. *Remote Sens.* 8, 4 (2016).
[37] J. Petajajarvi, K. Mikhaylov, A. Roivainen, T. Hanninen, and M. Pettissalo. 2015. On the coverage of LPWANs: range evaluation and channel attenuation model for LoRa technology. In *Proc.of ITST*.
[38] S. I Popoola, A. A. Atayero, N. Faruk, C. T. Calafate, E. Adetiba, and V. O. Matthews. 2017. Calibrating the standard path loss model for urban environments using field measurements and geospatial data. In *Proc. of The World Congress on Engineering*.
[39] T. Rappaport. 1996. *Wireless communications: Principles and Practice*. Vol. 2. Prentice Hall PTR New Jersey.
[40] T. Rappaport and R. Skidmore. 2010. Method and system for using raster images to create a transportable building database for communications network engineering and management. US Patent 7,711,687.
[41] R. A Schowengerdt. 2006. *Remote sensing: models and methods for image processing*. Elsevier.
[42] V. E. Somoza, G. Almeida, P. D McDonald, and P. Hill. 2002. Tools for wireless network planning. US Patent 6,336,035.
[43] V. Vapnik. 1998. *Statistical learning theory*.
[44] T. Voigt, M. Bor, U. Roedig, and J. Alonso. 2017. Mitigating inter-network interference in LoRa networks. In *Proc. of MadCom*.
[45] A. J Wixted, P. Kinnaird, H. Larijani, A. Tait, A. Ahmadinia, and N. Strachan. 2016. Evaluation of LoRa and LoRaWAN for wireless sensor networks. In *Proc. of IEEE SENSORS*.
[46] M. Zuniga and B. Krishnamachari. 2004. Analyzing the transitional region in low power wireless links. In *Proc. of IEEE SECON*.