

# Towards a Low-cost RSSI-based Crop Monitoring

JAN BAUER and NILS ASCHENBRUCK, Distributed Systems Group, Institute of Computer Science, University of Osnabrück

The continuous monitoring of crop growth is crucial for site-specific and sustainable farm management in the context of precision agriculture. With the help of precise *in situ* information, agricultural practices, such as irrigation, fertilization, and plant protection, can be dynamically adapted to the changing needs of individual sites, thereby supporting yield increases and resource optimization. Nowadays, IoT technology with networked sensors deployed in greenhouses and farmlands already contributes to *in situ* information. In addition to existing soil sensors for moisture or nutrient monitoring, there are also (mainly optical) sensors to assess growth developments and vital conditions of crops. This article presents a novel and complementary approach for a low-cost crop sensing that is based on temporal variations of the signal strength of low-power IoT radio communication. To this end, the relationship between crop growth, represented by the leaf area index (LAI), and the attenuation of signal propagation of low-cost radio transceivers is investigated. Real-world experiments in wheat fields show a significant correlation between LAI and received signal strength indicator (RSSI) time series. Moreover, influencing meteorological factors are identified and their effects are analyzed. Including these factors, a multiple linear model is finally developed that enables an RSSI-based LAI estimation with great potential.

CCS Concepts: • **Information systems** → **Sensor networks**; • **Applied computing** → **Agriculture**; • **Computer systems organization** → **Sensor networks**; • **Hardware** → *Sensor devices and platforms*; • **Networks** → Physical links; Network experimentation;

Additional Key Words and Phrases: Internet-of-Things (IoT), precision agriculture, wireless sensor network (WSN), link quality (LQ), radio propagation, received signal strength indicator (RSSI), leaf area index (LAI), vegetation, temperature, humidity

## ACM Reference format:

Jan Bauer and Nils Aschenbruck. 2020. Towards a Low-cost RSSI-based Crop Monitoring. *ACM Trans. Internet Things* 1, 4, Article 21 (June 2020), 26 pages.  
<https://doi.org/10.1145/3393667>

## 1 INTRODUCTION

Information technology may be the most important enabler for a sustainable and efficient precision agriculture. Recent advances have enhanced not only agricultural infrastructure and machinery but also autonomous monitoring and control systems. Under the umbrella of the Internet-of-Things (IoT), smart, distributed, and collaborating sensors greatly support the entire chain from data gathering, communication, processing, and analytics to yield modeling and predictions

Authors' addresses: J. Bauer and N. Aschenbruck, Distributed Systems Group, Institute of Computer Science, University of Osnabrück, Wachsbleiche 27, 49090 Osnabrück, Germany; emails: {bauer, aschenbruck}@uos.de.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2020 Association for Computing Machinery.

2577-6207/2020/06-ART21 \$15.00

<https://doi.org/10.1145/3393667>

[7, 30]. By supporting the understanding of factors influencing crop growth and yields, IoT-based crop monitoring has the potential to provide new insights and to improve productivity.

In the context of Wireless Sensor Networks (WSNs), a lot of research on radio propagation has been conducted, particularly for indoor scenarios but also for outdoor deployments, e.g., References [3, 8, 31]. With the emerging digitization of industrial agriculture, outdoor deployments in farmlands [9, 17, 29] and forest terrains [18, 21, 24] also attracted scientific attention in the past two decades. Here, not only the impact of communication distances on the link quality (LQ), predominantly given by the RSSI, is studied but also the effect of different meteorological factors [19, 31]. In addition, the vegetation itself is reported to have a measurable effect on the attenuation of radio signal propagation in outdoor environments in more recent studies [16, 24]. However, this impact is yet not comprehensively explored for frequency bands in the lower microwave spectrum ( $<5$  GHz) as used for short-range IoT communication. Only very few publications on the subject exist, predominantly in forest environments (e.g., References [21, 24]) and very rare studies for short-range communication in farmland deployments (e.g., References [9, 17]). In contrast, the corresponding impact is well-known for higher frequencies in the microwave spectrum as used for satellite or Public Land Mobile Network (PLMN) communication. Here, there have even been vegetative signal propagation models for decades. [14].

Although the impact of vegetation on IoT radio communication has been observed and partially analyzed in recent literature, even in the context of precision agriculture, there is—to the best of our knowledge—no other work exploiting the relationship between vegetative growth and signal strength degradation for the monitoring of crop parameters, such as biomass or LAI. However, such an approach would have a large practical relevance, not only limited to agricultural applications. It does not require additional sensor hardware and, thus, could be easily integrated into existing agricultural monitoring systems.

In our previous work, we presented an agricultural IoT monitoring system and discussed real-world challenges [4]. This monitoring system uses a specific transmittance-based LAI sensor [6]. Recently, we realized a long-term deployment of our system in experimental wheat fields and presented an adequate approach for processing and filtering *in situ* sensor data to derive reliable LAI time series [5]. In this article, we now focus on a novel and complementary approach for crop monitoring that exploits the effect of vegetation on radio propagation. Based on the extensive data set collected in a wheat field deployment and the assessed LAI time series, we experimentally investigate whether cheap low-power IEEE 802.15.4 transceivers, as typically used in the IoT domain, are appropriate and sensitive enough to perceive the effect of vegetation and whether the 2.4 GHz ISM radio band is suited for such an approach. In summary, the major contributions of this article are:

- a novel approach enabling a new RSSI-based crop sensing technique;
- a holistic analysis of the impact of vegetation (winter wheat) and meteorological factors on LQ as well as their respective relative importance;
- a model for an RSSI-based LAI estimation that demonstrates the practical feasibility and the potential of our approach.

The remainder of this article is organized as follows. In Section 2, relevant background information is introduced. Then, the state of the art is discussed in Section 3. Afterward, in Section 4, the attention is turned on the IoT deployment that is used for the field experiments. A stepwise analysis of the importance of environmental factors having a significant impact on LQ is subsequently presented in Sections 5–7. This leads to our novel approach for RSSI-based crop monitoring and an evaluation of its potential in Section 8. Finally, our findings as well as the limitations of this study are discussed in Section 9 and concluded in Section 10.

## 2 BACKGROUND

### 2.1 Link Quality Metrics

The quality of wireless links is determined by the path loss, i.e., the degradation of the transmitted radio signal during its spatial propagation. This degradation can have different reasons [2]. First, there is a physical signal attenuation caused by the so-called free-space loss and several other physical effects, such as absorption, diffraction, refraction, and shadowing. Amplified by the environment, i.e., obstacles in Fresnel zones, these effects also lead to interfering multipath propagation. Second, interference induced by background noise and concurrent transmissions further degrade the signal. Finally, the location and orientation of antennas have a significant impact. In addition to path loss, the LQ can also be degraded by the hardware, which is more prone to internal noise, especially with low-cost transceivers.

To quantify the quality of links, many LQ metrics, or estimators, have been introduced. A comprehensive survey on existing estimators is given by Baccour et al. [2]. The idea of a crop monitoring based on the degradation of LQs assumes the significance of the impact of crop growth on radio propagation. Thus, rather than the LQ that is important for the final application from a higher layer perspective, the path loss and actually the physical attenuation induced by the vegetation is of major interest for our investigation. In this work, we hence focus on receiver-side, hardware-based LQ estimators, as we believe that they reflect physical effects more directly. Two widespread estimators are used, namely, the received signal strength indicator (RSSI) and the link quality indication (LQI), both provided by the transceiver. Also, the packet delivery ratio (PDR), sometimes denoted as packet reception ratio (PRR) or inversely referred to as packet error ratio (PER), is initially used as a basic software-based estimator. Whereas the PDR is typically obtained by simply counting received packets on application-layer, the hardware-based implementation of both, RSSI and LQI, is more complex.

RSSI estimates the input power of a radio signal, usually given in dBm. For typical IEEE 802.15.4 transceivers as the CC2400 product family, it is calculated over eight symbols (128  $\mu$ s periods) [11]. The chosen sensor platform for our experiments is the widely-used TelosB device that integrates a CC2420 chip that returns RSSI values in the range of  $-100$  dBm to  $0$  dBm with an accuracy of  $\pm 6$  dB. Since the RSSI also includes background noise, the probability of distinguishing received signals decreases if RSSI is lower than  $-90$  dBm, resulting in a sensitivity of roughly  $-95$  dBm. LQI, in contrast, is a dimensionless score that represents a vendor-specific statistical value, which can be considered as chip error rate. For the CC2420 chip, it is calculated over the first eight symbols following the start-of-frame delimiter. Returned values are usually between  $110$  and  $50$ , corresponding to maximum and minimum LQs, respectively [11].

Due to transient (short-term) fluctuations in signal strengths, averaging over multiple values is recommended. However, RSSI was observed to be relatively stable over a short period. In contrast, LQI is highly variable in transitional regions, i.e., for intermediate links with  $10$ – $90\%$  PER. Therefore, a large packet window for averaging with a minimum of  $40$  packets should be used in practice [2, 28]. Hence, LQI requires more transmissions and, thus, is not only less responsive but also less applicable in precision agriculture deployments due to the higher energy demand. Nevertheless, the suitability of both metrics for crop monitoring will be evaluated in this work.

### 2.2 Crop Parameters

For the characterization and monitoring of crop growth and vitality as well as for grain yield modeling and field phenotyping, a large number of various plant parameters are considered across a variety of research disciplines, such as biology, agronomy, ecology, and even climatology and meteorology. Agronomical crop parameters range from crop height and root depth to more complex parameters as the standing or dry biomass, the vegetation water content (VWC), the leaf

chlorophyll content, the fraction of absorbed photosynthetically active radiation (fPAR), and most notably the leaf area index (LAI) [1]. This index is one of the most important bio-physical plant parameters and widely used in vegetation studies. It is a dimensionless quantity that is commonly defined as the amount of one-sided leaf area ( $m^2$ ) per unit ground area ( $m^2$ ). Particularly in the context of precision agriculture, yield-limiting and -reducing processes caused by plant diseases and mismanagement can be indicated by LAI monitoring.

LAI assessment is mainly conducted indirectly, i.e., by actually deriving LAI estimates from a related quantity. This can be done *in situ* by using commercial instruments, digital hemispherical photography (DHP), or more recently by WSNs [6, 23]. Alternatively, LAI information can be derived via remote sensing from satellite or airborne images, enabling large-scale monitoring [20, 32]. A different widespread vegetation index assessed by remote sensing is the normalized difference vegetation index (NDVI) that is derived from characteristic absorptions of specific wavelengths of vegetative reflectance. Both LAI and NDVI are, to a certain extent, related to the biomass. Overall, *in situ* assessment is generally costly and time-consuming but indispensable for the validation of remote sensing products.

In the context of radio signal propagation, whenever its impact is taken into account, vegetation is sometimes represented by fPAR or NDVI, but predominantly by LAI [14, 17, 25]. Hence, LAI is assumed to be an appropriate parameter for the empirical study in this work.

### 2.3 Meteorological Variables

For radio propagation in higher frequencies in the microwave spectrum ( $>10$  GHz), meteorological variables are known to have a significant influence, especially for earth-space telecommunication [15]. Likewise, for low-cost transceivers in the 2.4 GHz band, outdoor experiments reported in literature sometimes revealed similar influences. Both humidity and air temperature ( $T$ ) are fundamental meteorological factors and closely connected. The humidity describes the amount of water vapor in the air. The water content in a particular volume of air is defined as the absolute humidity ( $AH$ ) and usually given in  $g/m^3$ . Because the maximum amount of vapor the air can hold is highly temperature-dependent, the relative humidity ( $RH$ ) is often used that defines the proportion of present vapor to this maximum. Changes in temperature also imply  $RH$  changes, although the amount of water vapor in the air ( $AH$ ) remains unchanged. Hence, this metric can be misleading [19, 31]. For this reason, both  $RH$  and  $AH$  will be carefully reflected. Since only  $RH$  is available,  $AH$  ( $g/m^3$ ) is calculated from  $RH$  (%) and  $T$  ( $^{\circ}C$ ) based on the following general formula, which is also used by Luomala and Hakala [19]:

$$AH(T, RH) = 216.7 \cdot \frac{\frac{RH}{100\%} \cdot A \cdot e^{\frac{m \cdot T}{T_n + T}}}{273.15 + T}, \quad (1)$$

with  $A = 6.112$  hPa,  $T_n = 243.12$   $^{\circ}C$ , and  $m = 17.62$ .

Beside both key factors, also other variables as rainfall, fog, snow, cloud coverage, and even solar radiation and wind are occasionally considered in the literature [8, 22, 24, 29]. However, due to their strong correlation with temperature and humidity, the analysis of their impacts is severely complicated. It requires comprehensive data sets covering a wide range of different weather variables and mutual combinations, which is barely the case in many existing studies. Consistent with related studies, our primary focus is, therefore, on the impact of temperature and humidity. However, to some extent, the effects of other variables will also be analyzed.

## 3 RELATED WORK

Outdoor radio propagation characteristics of low-power transceivers for short-range communications have been considered by several studies in the past. Particularly for the IEEE 802.15.4 WSN



technology, there is a relatively large number of publications that differ in terms of location (urban vs. rural), experiment duration (few hours vs. months), and the number of included environmental factors.

### 3.1 Meteorological Effects

Thelen et al. [29] pioneered the research on low-power radio propagation in agricultural crop fields. Their work was motivated by the potential effects of environmental factors that were missing in simulators and theoretical models and often unforeseen by network developers. In a potato field, a correlation of RSSI with temperature and humidity was observed, and the latter was found to have the main influence. A higher *RH*, e.g., at night or in the presence of rain, was shown to improve LQ. This improvement was attributed to changes in the reflection coefficient of the canopy, thus, environmentally induced. However, Bannister et al. [3] studied the effect of *T* in both an outdoor desert and an indoor lab experiment, neglecting humidity effects. They found a strong linear correlation of RSSI and *T* and showed the possible degradation of network connectivity caused by temperature increase. In a forest deployment, Luo et al. [18] found that both *T* and *RH* correlate with RSSI but concluded that the effect of *T* is more significant than *RH*. Based on seasonal experiments in controlled open-field environments, later studies on meteorological factors, conducted by Wennerström et al. [31] and Luomala and Hakala [19], explicitly differentiate between *RH* and *AH*. These studies provide a more comprehensive insight and show that *T* has the main impact on LQ. Hence, sometimes empiric and platform-specific *T* correction coefficients were determined [3, 8, 19, 27]. The diurnal variations of *T* induce a characteristic diurnal pattern, also referred to as day-night regime, that was observed in many experiments [3, 8, 19, 21, 29, 31]. Although there are some controversial findings, most studies agree that *T* affects low-cost transceiver hardware rather than physically contributing to the path loss. Despite this broad consensus, it is unclear whether transmitters or receivers are more impaired by increasing temperature. Based on controlled lab experiments, contradictory findings were reported [3, 8, 27].

The influence of humidity is interestingly still not fully understood. Markham et al. [22] observed a significant effect of rainfall in a 26-day forest deployment. Similar to Thelen et al. [29], they found that the presence of rain increased the quality of poor links but, however, it simultaneously decreases the quality of good links in connected regions. They argue that neither fading caused by rainfall itself has an impact, nor changes in the reflection coefficient. Instead, their findings are explained by the presence of water on the exterior of the transceiver's enclosure inducing a capacitive loading of the antenna. Gernert et al. [12] used sensors planted in the soil of a potato field. They observed that the water content of the surrounding soil has a negative impact on the RSSI. Moreover, the water on leaves and the humidity affect the LQ during rain.

In contrast, based on controlled outdoor experiments, Boano et al. [8] claim that the effects of rainfall and also of fog are almost negligible in practice. Only heavy rainfall might impact LQ. This conclusion is supported by Wennerström et al. [31] and the worst-case maize field campaign of Li and Gao [17], who also found that there appears to be no correlation between precipitation and either RSSI or PDR. Results are, furthermore, consistent with findings reported by Brinkhoff and Hornbuckle [9] for WLAN communication in a cotton field and also with rain fade theory [15].

Despite several other studies, there are some discrepancies in conclusions and sometimes contradictory findings, in particular regarding the meteorological factors that have the main influences and physical causes. The major challenge is the strong mutual correlation of those factors and their complex interaction with one another. Discrepancies might further be explained by different methodologies, e.g., regarding communication hardware, meteorological data sources, or the duration of conducted experiments, cf. References [2, 31]. Moreover, we believe that in some studies,

also a sparse and possibly insufficient data range of meteorological variables additionally contributes to misinterpretations and incompatibilities.

### 3.2 Attenuation in Vegetation

Whereas the vegetation loss of radio propagation has been extensively investigated for telecommunications [14, 25], the impact of vegetation on low-power, short-range communication is much less explored. Compared to the investigations on the effects of meteorological factors, only a very few studies exist, particularly in the context of precision agriculture, often with the focus on forestry. Based on PDR, Marfievici et al. [21] showed the adverse effect induced by vegetation for varying densities in a long-term forest campaign. Rankine et al. [24] present an extensive study on the temporal relationship between RSSI and plant canopy development of a boreal forest. By multiple linear regression, they concluded that, amongst other variables, both temperature and NDVI have the most influential contribution to RSSI variance.

Especially for radio propagation in arable land, Thelen et al. [29] reported a dramatic decrease in the communication range (from 70 to 10 m) due to the emerging foliage of potato crops, similar to the observations of Kulau et al. [16]. Such a decrease that is much more significant than the effect of varying weather conditions was also found in [9] for cotton and rice crops. The foliage and stems of plants have a great effect on diffraction, absorption, and scattering loss but also plant height and field layout matters in the case of row crops. Li and Gao [17] therefore experimentally evaluated adequate antenna placements for radio communication in maize fields.

In the domain of remote sensing, the prevailing passive microwave technique is GNSS-Reflectometry (GNSS-R), which has been applied to several applications, including the sensing of soil moisture and vegetation growth. Chen et al. [10] demonstrated the feasibility of vegetation sensing (VWC and height) via GPS interferometric reflectometry.

In contrast to many previous studies, our long-term deployment during the wheat growing season has a clear focus on precision agriculture and takes daily and seasonal variations into account. Moreover, our experiment and its analysis are motivated by the general idea of GNSS-R, namely, vegetation sensing by utilizing a change in the radio signal induced by vegetation, and, thus, are aimed at a novel approach for low-cost RSSI-based crop monitoring.

## 4 EXPERIMENTAL SETUP

The overarching goal of our agricultural IoT monitoring system [4] that is used for the LQ investigations in this work is an adequate long-term evaluation environment for a previously developed transmittance-based LAI sensor [5, 6]. Motivated by commercial LAI handheld instruments, this specific low-cost sensor leverages optical measurements to assess the transmittance, i.e., the amount of solar radiation that is intercepted by the canopy, and subsequently derives LAI estimates. To this end, sensors deployed in the field are placed on the ground below the canopy. They are supported by additional reference sensors gathering the unshaded down-welling irradiance above the canopy. By integrating these sensors into a WSN as sensing component of the IoT monitoring system, this approach was shown to achieve promising results in wheat cultivars ( $R^2 = 0.91$  in comparative analyses with a traditional approach) and enables monitoring of LAI time series [5], often referred to as LAI trajectories. However, this transmittance-based LAI monitoring requires specific sensor hardware. Because of the impact of vegetation on LQ reported from previous experiments in related work, we hence ask whether there is an exploitable correlation between crop growth and LQ decrease. Such a correlation could potentially enable a complementary crop sensing technique.

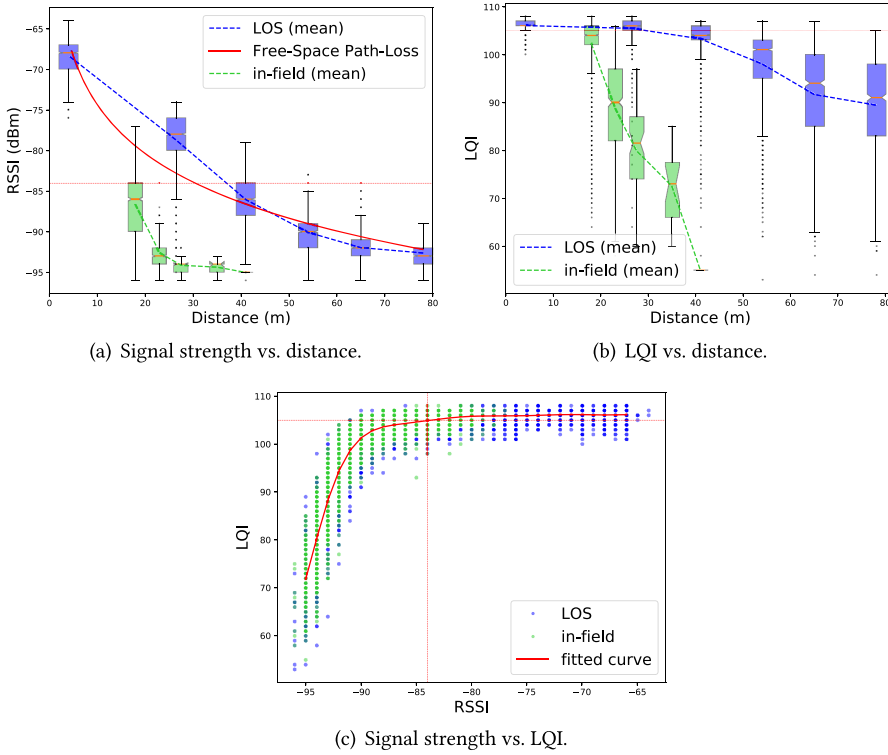
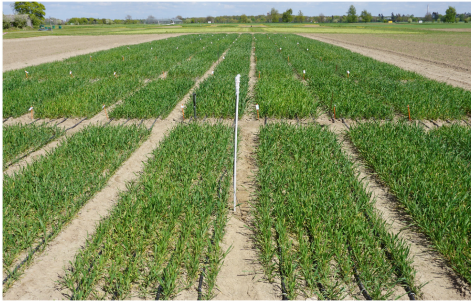


Fig. 1. Motivating results from a representative link of a preliminary worst-case experiment showing a significantly degraded signal propagation through vegetation (mustard field). Thresholds (red lines,  $\text{RSSI} > -84$  dBm and  $\text{LQI} > 105$ ) were also identified for both hardware-based LQ metrics that indicate a mean PDR better than 0.95.

#### 4.1 Pre-deployment Phase

Before the IoT deployment that is used for this study, several preliminary experiments were conducted that were necessary for network planning. Different low-cost commercial off-the-shelf sensor platforms were evaluated, and the widespread TelosB device (8 MHz TI MSP430 MCU, 10 kB RAM) was finally selected as the underlying platform. As already mentioned in Section 2.1, it features the IEEE 802.15.4-compliant CC2420 radio transceiver. For the deployment, we use the low-gain internal antenna that is embedded in the circuit board of the device, radio channel 26 (2,480 MHz), and the maximum transmission power (level 31) [11].

Amongst others, the preliminary experiments comprised a communication range evaluation, which was conducted in December 2014 in a field located near the University of Osnabrück, Germany. White mustard (*Sinapis alba* L.) was grown as a catch crop with a varying plant height of roughly 1 m. As a kind of worst-case scenario, a rainy day (day of year (DOY) 344) was chosen. One device was placed on a fixed tripod above the canopy (2 m) and periodically transmitted probe messages (10 packets/s). Another device was placed on a small bracket at ground level, either in an open area or in the mustard field. This device acted as a potential receiver and LQ logger. Its distance to the sender was gradually increased. Transmissions are monitored for approximately 4 min per distance increment. A significant impact of vegetation was observed for different transceivers that either operate in the 2.4 GHz or the 868 MHz band. Due to the focus on the selected TelosB platform, the results presented here are limited to its CC2420 radio chip.



(a) DOY 105: small vegetation ( $LAI < 0.5$ ) in the tillering stage at the beginning of the experiment.



(b) DOY 179: dense vegetation ( $LAI \approx 2-3$ ) in the ripening stage at the end of the growing season.

Fig. 2. Study site with controlled wheat plots on experimental fields of the Julius Kühn-Institute (JKI), Braunschweig, Germany.

The evaluation clearly showed the effect of vegetation (Figures 1(a) and 1(b)). For line of sight (LOS) transmissions, reasonable distances up to 80 m could be achieved. The logarithmical decrease of the signal strength over distance corresponds to signal propagation theory. It fits the free-space path loss model (red curve in Figure 1(a)). The presence of vegetation in the Fresnel zones results in an attenuation of the radio signal and, thus, significantly decreases the communication range and LQ values, respectively. As long as distances are not too large (i.e., within connected regions), LQI values are very stable (Figure 1(b)). However, our experiences confirm that within transitional regions, LQI values are highly variable. This observation is further illustrated by a scatterplot (Figure 1(c)) that unveils a strong but non-linear relationship between both metrics. To distinguish between connected and transitional regions, hereinafter, the commonly used *LQI threshold* of 105 was adopted. The mean PDR observed for transmissions with LQI values above that threshold was better than 0.95. As shown by red lines in the scatterplot, the corresponding *RSSI threshold* is roughly  $-84$  dBm, similar to values found in References [28, 29].

Whereas RSSI is more sensitive to small changes of signal strengths, also in case of reliable links with good quality, LQI is less responsive to those changes and decreases not until LQ is drastically reduced. Consequently, LQI better reflects the PDR ( $R^2 = 0.70$  vs.  $R^2 = 0.25$  of RSSI-PDR correlation) and, thus, the LQ from an application perspective. In fact, due to its higher responsiveness, RSSI is commonly considered to be more suitable for localization purposes.

Overall, the findings derived from this preliminary experiment essentially reproduced general results from outdoor deployments reported in the community, particularly in agricultural contexts such as References [17, 29]. Moreover, the clear difference of LOS communication and transmission through vegetation additionally motivated our idea of LQ-based crop sensing very early and encouraged us to continue with further experiments.

## 4.2 Study Site

The IoT monitoring system was deployed on an experimental field of the Julius Kühn-Institute (JKI) located at Braunschweig, Germany. During the 2016 growing season, various winter wheat cultivars (*Triticum aestivum* L.) were grown under different water regimes in controlled plots (Figure 2). The main deployment was operating for 74 days (DOY 105–178, i.e., 2016/04/14 to 2016/06/26), covering the relevant growth period from tillering to ripening. The temporal development of crop growth is represented by LAI trajectories. These are sensed by the deployment (primary motivation) and are steadily complemented by traditional LAI measurements carried out manually. Further details on the deployment site can be found in Bauer et al. [5], whereas additional details on the



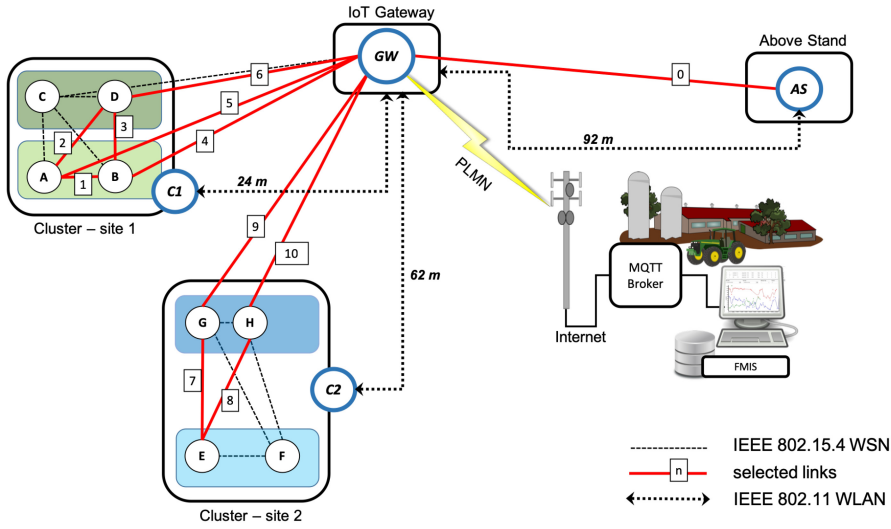


Fig. 3. Architectural overview of the IoT-based agricultural monitoring system and its concrete IEEE 802.15.4 topology in the experimental deployment (not in scale). Representative links with different qualities are emphasized by solid red lines. Further details on selected links are listed in Table 1.



(a) TelosB and the enclosure with a diffuser cap for its PAR sensor.



(b) Sensor node deployed in the wheat field.

Fig. 4. The transmission-based LAI sensor developed in Bauer et al. [6].

agronomic experiment are presented in Schittenhelm et al. [26]. Moreover, there is an official German Weather Service (DWD) station (ID 662) close to the field ( $< 1$  km). Observations of its high-quality meteorological sensors are freely available<sup>1</sup> and used to support internal sensor readings.

### 4.3 IoT Deployment

Our IoT system combines different technologies for continuous crop monitoring. Its design and implementation details were presented in Bauer and Aschenbruck [4]. An overview of the system's architecture is provided in Figure 3. The *in situ* tier of the system comprises low-cost and resource-constrained WSN technology. Based on the TelosB platform, a specially developed LAI sensor [6] was deployed, depicted in Figure 4. Two fields, each with different wheat cultivars, were monitored. According to the conducted preliminary investigations (Section 4.1), particularly

<sup>1</sup>[http://ftp-cdc.dwd.de/pub/CDC/observations\\_germany/climate/hourly](http://ftp-cdc.dwd.de/pub/CDC/observations_germany/climate/hourly).



Table 1. Overview of Links Selected from the Deployment (Figure 3)

ID	Sensors		Distance (m)	Type	CAT	Quality
0	GW	AS	92	above-above	2	medium
1	A	B	0.3	ground-ground	1	high
2	A	D	1.2			
3	B	D	1.2			
4	B	GW	22.5	ground-above	2	medium
5	A	GW	23			
6	D	GW	22			
7	G	E	3.7	ground-ground	3	low
8	H	E	3.9			
9	G	GW	65	ground-above		
10	H	GW	65			

because of the limited in-field communication range, a clustered network topology was designed. A small-scale cluster of four sensors (sensors *A–D* and *E–H*, respectively) with short communication distances (<5 m) were assigned to each field. Moreover, there is an additional cluster outside the field with an unshaded above reference sensor *AS* used for transmittance-based LAI estimation. Using cluster heads *C1* and *C2* that are (wirelessly) interconnected in a WLAN, data collected by sensors is forwarded to a specific IoT gateway *GW*. *In situ* data is further delivered from the gateway via the IoT infrastructure tier to a central server for data processing and persistence.

Despite the topological separation into clusters, broadcast transmissions over a shared communication channel are consciously used, so that all devices can potentially overhear messages from one another. By this means, not only links between sensors in the same cluster but also different clusters are possible, depending on the particular LQ. From those potential links, a representative set (emphasized in Figure 3) was selected for the analysis. Amongst them, there are links of different distances and types, in terms of pairwise transceiver locations, cf. Table 1. According to the findings in the pre-deployment phase and based on link distance and type, links are grouped into three categories (CATs). These do not necessarily correspond to a certain communication region, because links are expected to change over the growing season dynamically.

#### 4.4 Data Retrieval

A custom data retrieval software for WSN devices was implemented in TinyOS, an open-source operating system tailored for resource-constrained devices. Devices were programmed to capture readings of all available environmental sensors, i.e., temperature, humidity, and photosynthetically active radiation (PAR) (for the actual transmittance-based LAI estimation) with a constant two-minute sampling interval. Subsequently, these readings are bundled in a single data frame and transmitted to the assigned cluster heads. Due to the possible overhearing of concurrent transmissions, LQ values of neighboring sensors can potentially be obtained. In this case, those values are attached to the next outstanding data frame and delivered to the central server. During night phases (00:00–03:00), the data collection was paused.

## 5 IMPACT OF VEGETATION ON LINK QUALITY

### 5.1 Observation

An initial overview of LQs of selected links during the relevant part of the wheat growing season is given in Figure 5. For each link, first, daily averages of LQ values were calculated (arithmetic

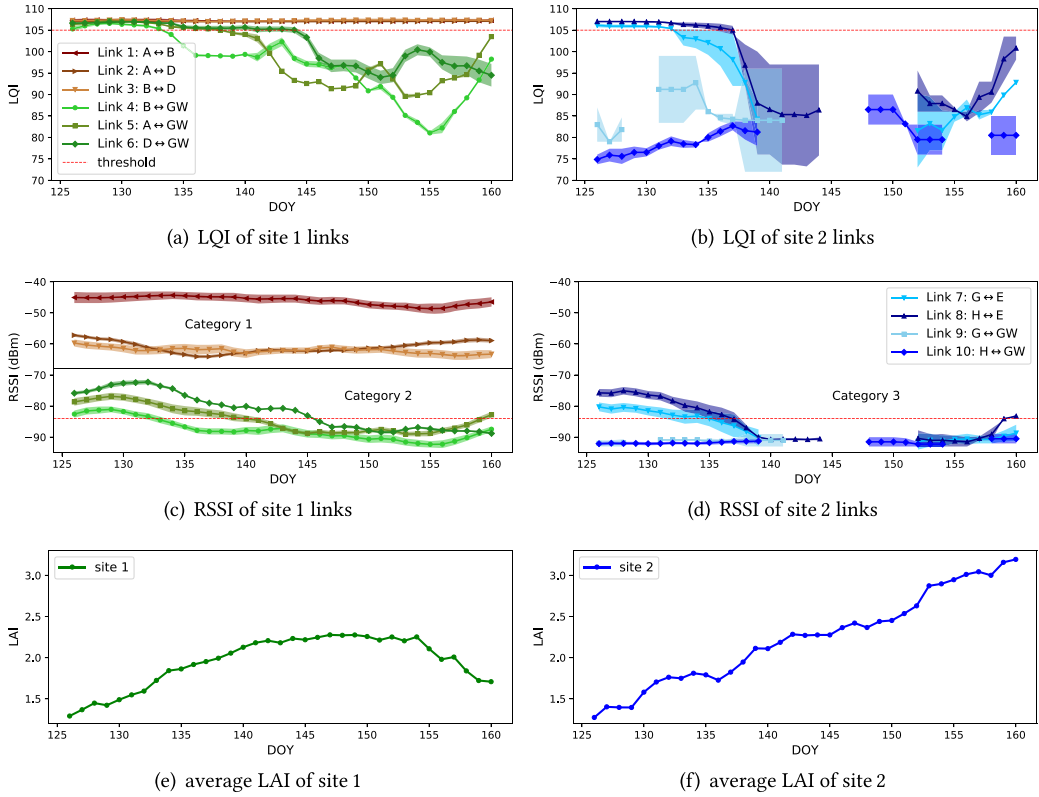


Fig. 5. Overview on LQs and crop development during the experiment providing a first insight into the adverse effect of vegetation on radio propagation.

mean). Then, these averages are smoothened by a central simple moving average (SMA) with a sliding window size of  $\pm 3$  days. This window size is consistently used in this article unless otherwise stated. Figure 5 shows the LQI and RSSI development over time for both sites. Here, the bidirectional mean LQs are presented by lines, whereas the difference between directed LQs is visualized by the corresponding filled areas between both unidirectional curves, indicating possible link asymmetries. However, links turned out to be quite symmetric if LQ is at a reasonable level. Only if LQs decrease towards disconnections, link asymmetries increase, as can be observed for CAT 3 links (shaded blue) in Figures 5(b) and 5(d). It can also be seen that LQI values suffer more from low LQs than RSSI values, since LQI curves are highly varying under these conditions. Moreover, the maximal communication range of in-field ground-ground links is found to be drastically reduced by vegetation (Links 7 and 8 in Figures 5(b) and 5(d)). Because ground-ground links are also impaired by the adverse positioning of transceivers, i.e., restricted Fresnel zones, distances of only a few meters were possible, which is consistent with Li and Gao [17].

Concerning link categories, it turns out that our initial assignment is well suited, because links of different categories are separated from each other. Moreover, at the beginning of the deployment, CAT 2 and 3 links obviously have sufficient LQ (values above thresholds) and, thus, are in connected regions. However, during the experiment, LQs decreased so that CAT 2 links were shifted into transitional regions and CAT 3 links even into disconnected regions (missing values between DOY 140 and 152 in Figures 5(b) and 5(d)).

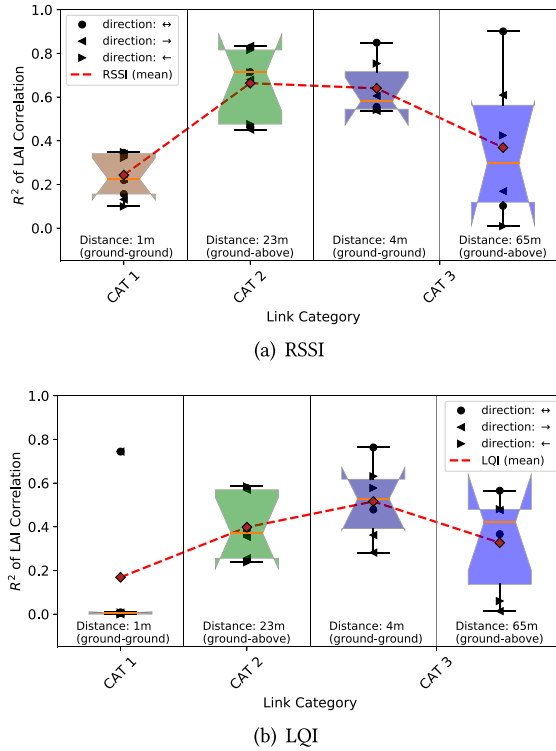


Fig. 6. Correlation of LQ and crop growth (LAI) grouped in three link categories.

Besides, crop growth dynamics, represented by LAI trajectories, are also included (Figures 5(e) and 5(f)). The notable difference in crop development in both sites is intentionally caused by different water treatments in the experimental fields (cf. Section 4.2 and Schittenhelm et al. [26]). Comparing the LQ and LAI curves, it seems that there is the expected causal relation between signal propagation decrease and vegetation increase.

## 5.2 Analysis

To analyze the correlation between LQ and LAI time series, for each link and both LQI and RSSI, the linear correlation (Pearson product-moment) is determined. Resulting coefficients of determination ( $R^2$ ) are visualized in Figure 6 and grouped in CATs and distances, respectively. Both link directions (labeled as  $\leftarrow$  and  $\rightarrow$ ) were separately considered and additionally their bidirectional means (labeled as  $\leftrightarrow$ ). The analysis reveals a moderate linear correlation between RSSI and LAI in the experiment for CAT 2 links (green boxplot in Figure 6(a)). Concerning CAT 3 links (blue boxplots), a similar correlation was found for the 4 m ground-ground links, but the correlation of 65 m ground-above links is much weaker and extremely varying, depending on the particular link and its direction. However, the correlation achieved by CAT 1 links (i.e., over 1 m distances; brown boxplot) is very weak, as well. Hence, the overall LQ (and, thus, also the distance between transceivers) apparently has a significant impact on the capability of LQ-based perception of vegetation. In other words, it seems that the overall LQ should not be too good so that the influence of vegetation is still perceptible; but the influence on low-quality links might be too intensive and bears the risks of link disconnections. In our experiment, CAT 2 links between ground and above sensors (typical for agricultural monitoring systems) were found to be well-suited for this purpose.

For LQI, determined LAI correlations have a similar trend but are generally much weaker than for RSSI (Figure 6(b)). Thus, in accordance with our expectations, LQI is evaluated to be less appropriate for our purpose. As a consequence, from now on, we focus on RSSI as a promising LQ metric. LQI will play only a minor role in further analyses. Moreover, due to an increased asymmetry of CAT 3 links and particularly due to their temporal disconnections, these links are also neglected hereinafter.

Overall, already promising results (median  $R^2 \approx 0.7$  for CAT 2) were achieved for the RSSI-LAI correlation. However, meteorological effects are ignored so far. Hence, RSSI dynamics observed in the experiment could also be incidentally caused by temperature or humidity impacts, for instance, and not necessarily by vegetation. To this end, in the following two sections, the meteorological impact on RSSI measurements will be extensively analyzed and whether it is negligible or instead dominates influences on LQ.

## 6 SHORT-TERM ENVIRONMENTAL EFFECTS

### 6.1 Observation

It is well-known that meteorological factors have an impact on low-cost, low-power radio transceivers. However, as stated in Section 3, the primary factor is controversially discussed. Hence, it is worth to have a closer look at this subject. Therefore, it is crucial to divide the total duration of the experiment into shorter observation periods with similar vegetation and weather conditions to focus on the effects of single meteorological factors.

Figure 7 shows the unidirectional raw RSSI values of Link 5 (both directions) over a short period of five days at the beginning of the experiment. By applying an SMA over a short time window ( $\pm 2$  h) to both unidirectional RSSI series and then calculating their mean, the bidirectional RSSI curve (red) is obtained. RSSI values clearly show the diurnal pattern, i.e., the characteristic behavior over the time of a day (cf. Section 3.1). To identify the main cause for this phenomenon, the time series of three main influencing factors,  $T$ ,  $RH$ , and  $AH$ , were considered in our investigation (Figure 7). On the one hand, there is weather data from an external DWD station (black) and, on the other hand, weather data gathered by both sensors,  $A$  and  $GW$  (red). It can be seen that the ambient temperature measured by the sensors is much more extreme than the actual outdoor temperature (DWD). In peaks, values that are roughly twice as high are reached. That can be explained by the black and weatherproofed plastic enclosures of sensor devices that were directly exposed to sunlight. Furthermore, there is an increasing difference between  $RH$  series of  $A$  and  $GW$ , indicated by the red-filled area behind the mean  $RH$  curve. We think this is caused by the different sensor positions (ground vs. above) due to a more humid microclimate at ground level. The outdoor  $AH$  curve (DWD) is relatively constant over the observation period and does not show a diurnal pattern. In contrast, the corresponding curve derived from sensor readings is completely skewed because of the extreme temperatures in the interior of the enclosure that dominate the  $AH$  conversion (Equation (1) in Section 2.3). Thus, unlike DWD values,  $AH$  readings of sensors do not reliably represent the real outdoor  $AH$ .

### 6.2 Analysis

**6.2.1 Comparative Analysis.** In the first step, a simple comparative analysis is conducted between RSSI and  $T$ ,  $RH$ , and  $AH$  values, respectively. For several links,  $R^2$  of the corresponding correlations is determined, and the three meteorological factors are explicitly distinguished. Furthermore, it is also differentiated between sensor readings of sender and receiver devices. The results clearly show a stronger correlation of RSSI with meteorological factors measured by the sensors than by the more reliable external DWD source (Figure 8). From this finding, we conclude

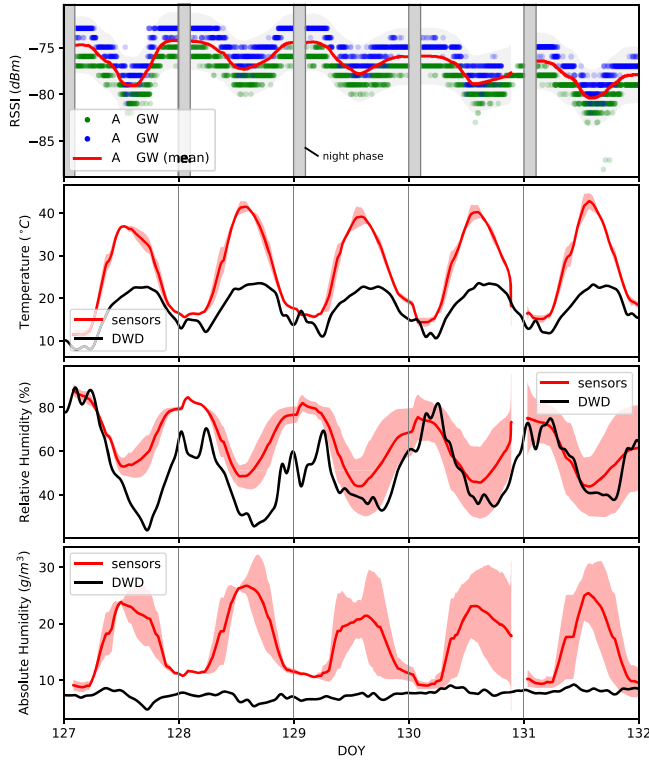


Fig. 7. Short-term RSSI and weather overview showing small-scale variances and a typical diurnal pattern of RSSI readings, that is apparently correlated with the meteorological factors.

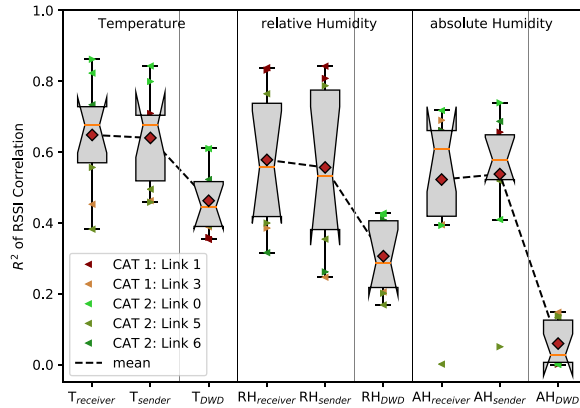


Fig. 8. Correlation of RSSI with  $T$ ,  $RH$ , and  $AH$  for data gathered by deployed sensors (sending or receiving devices) and external DWD weather information.

that the hardware rather than the transmission medium is primarily affected by those factors. Moreover, temperature correlations are slightly higher than humidity correlations. This indicates that  $T$  is the most dominant factor, which would be very likely and consistent with many related studies, e.g., References [8, 19, 24, 31]. However, due to the only slight difference between  $T$  and  $RH$ , further analysis on this subject will be conducted.



Another interesting observation is that RSSI and  $AH_{DWD}$ , which is assumed to be a reliable measure of the real  $AH$ , are actually not correlated (Figure 8). The reason is that the absolute humidity is quite constant over the observation period. Existing  $AH$  correlations with readings from sensors ( $AH_{sender}$  or  $AH_{receiver}$ ) are induced by the misleading diurnal pattern of the dominating  $T$  factor. Therefore,  $AH$  is hereinafter represented by  $AH_{DWD}$  and  $AH$  readings from sensors are discarded.

Finally, no significant difference is observed between  $R^2$  values derived from sender and receiver readings. Assuming  $T$  as the main factor, our experimental setup is unfortunately not suited to differentiate between the impact of temperature on sender or receiver due to similar temperatures of both devices (cf. Figure 7). Otherwise, if humidity, that differs between sender and receiver (cf. Figure 7), would be the main influencing factor, then there would apparently be no difference between impacts on sender and receiver because of similar  $AH$ -RSSI correlations (Figure 8).

**6.2.2 Relative Importance.** The correlations above might predict a possible causation but do not necessarily imply a causal relationship between corresponding factors. To discriminate the impact of the considered meteorological factors, particularly  $T$  and  $RH$ , the relative importance of those variables for the explanation of the observed RSSI curve is investigated in the second step. As long as explanatory variables, or regressors in this context, are uncorrelated, the determination of the relative importance is simple. However, this is not the case for meteorological factors. Therefore, inspired by Rankine et al. [24], multiple linear regression methods for model assessments are used for the quantification of the individual contribution of a certain regressor to a multiple linear regression model.

Leveraging *relaimpo* [13], a powerful package for the statistical software *R*, four different relative importance metrics are considered. The *First* metric simply compares the direct contribution of each regressor and completely ignores other regressors. Therefore, it is not really suited for the decomposition of contributions from individual regressors. The next metric, *Pratt*, multiplies the standardized coefficient with the marginal correlation. Because it can also allocate negative contributions but is unusable then, it is often criticized and not applicable in all situations. A more sophisticated and compute-intensive metric is introduced with *LMG*, one of the most robust metrics. It is based on sequential  $R^2$ s and addresses the dependence of orderings by averaging over orderings. Both the direct and indirect effects of other regressors are taken into account. Next, *PMVD* is a newly proposed metric that also can be seen as a specific (weighted) average over orderings. It was introduced to mitigate the disadvantage of *LMG* that, in case of a high correlation of a certain regressor with one or more strong regressors, possibly also assigns a positive contribution to this regressor. For all of these four metrics, a bootstrapping with a replication size of 1,000 and 95 % confidence intervals are used throughout our investigations.

In this section, we include the three meteorological variables,  $T$ ,  $RH$ , and  $AH$ , as well as the LAI in a multiple regression by defining the following additive four-regressor model:

$$RSSI \sim T + RH + AH + LAI. \quad (2)$$

Based on the findings of the previous section, here,  $T_{receiver}$  and  $RH_{receiver}$  of averaged sensor readings (of all CAT 1 and 2 links) are consciously used for  $T$  and  $RH$ , whereas  $AH$  is represented by  $AH_{DWD}$ . The results of multiple regressions with different decomposition metrics are visualized in Figure 9. First, it can be seen that the short-term RSSI variance can be nearly completely explained by the selected regressors ( $R^2 \approx 0.96$  with very small confidence intervals). As expected for the short observation period, across all metrics,  $AH$  and LAI are found to be almost not influential. It is moreover not surprising that our former results are replicated by the *First* metric. Similar to *First*, *LGM* is unable to decompose the relative importance, probably due to the very strong correlation

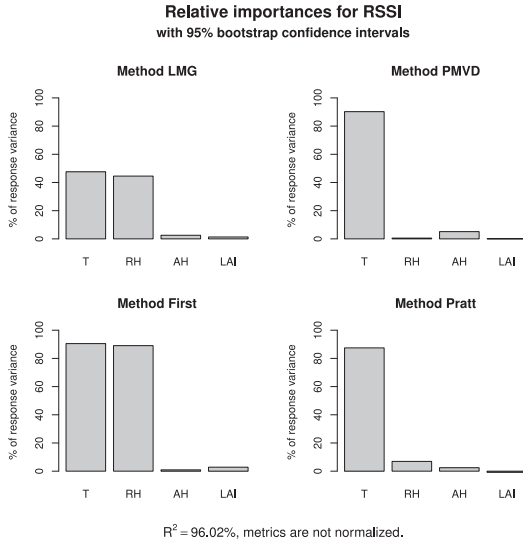


Fig. 9. Relative importance of basic meteorological factors ( $T$ ,  $RH$ , and  $AH$ ) and crop growth ( $LAI$ ) for short-term RSSI variance determined with four different decomposition metrics using the simple RSSI model (Equation (2)).

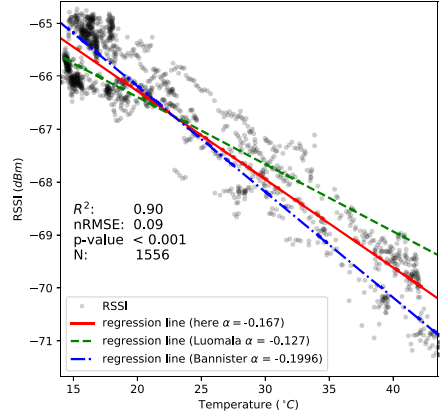
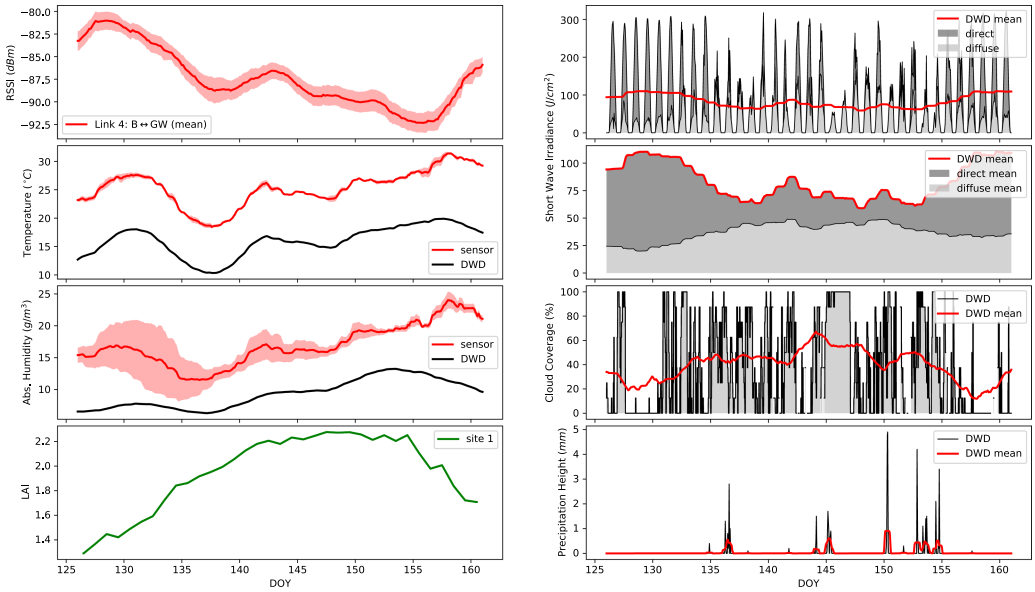


Fig. 10. RSSI vs. temperature shows a strong negative linear correlation ( $R^2 = 0.9$ ), which results in a 1.7 dBm RSSI decrease for a 10 °C increase in temperature.

of  $T$  and  $RH$ . In contrast, both *PMVD* and *Pratt* confirm  $T$  as the only key factor that significantly influences LQ here.

**6.2.3 Temperature Effect.** Finally, a correction factor for the temperature effect is determined. To this end, again, sensor readings of all CAT 1 and 2 links of the previous section (without SMA) are averaged. Then, linear regression is applied to RSSI and  $T$  values (Figure 10). The regression confirms a strong, statistically significant ( $p < 0.001$ ), linear correlation with  $R^2 = 0.90$ . This observation enables a reliable correction of temperature-depended impacts, using the slope ( $\alpha = -0.167$ ) of the fitted regression line (solid, red) as a correction factor. Hence, as a rule of thumb, a 10 °C increase in temperature decreases RSSI with roughly 1.7 dBm. This result is credible and consistent with correction factors for the CC2420 transceiver reported in the literature. For instance, Luomala and Hakala [19] found a factor of  $-0.127$  using a deployment in a university campus area (green dashed line) and Bannister et al. [3] a factor of  $-0.1996$  in a desert deployment (blue dashed line).

**6.2.4 Summary.** In conclusion, the analysis of relevant meteorological factors having a short-term impact has shown that  $T$  is the key factor. Humidity, in contrast, is evaluated to be negligible, at least as long as the absolute amount of water vapor in the air (i.e.,  $AH$ ) is relatively constant. In that case, a misleading impact of  $RH$ , as sometimes deduced in previous works, is solely caused by temperature changes. For this reason, further investigations of the skewed  $RH$  metric are left out from now on. However, the influence of  $AH$ , particular in the long-term scale, is still unclear and will be investigated in the next section. Moreover, we conclude that  $T$  primarily affects hardware components of low-cost transceivers (which also results from the temperature-dependent oscillator characteristics, cf. Reference [11]) and that its impact is correctable. An effect of  $T$  concerning the attenuation through the medium was not observed.



(a) Representative RSSI time series in comparison to its assumed main influencing factors, namely both meteorological variables  $T$  and  $AH$  as well as the LAI.

(b) Selection of additional meteorological variables that might also influence LQ and its time series provided by the weather station of DWD.

Fig. 11. Long-term overview of LQ, weather conditions, and crop growth over the relevant period of the growing season.

## 7 LONG-TERM ENVIRONMENTAL EFFECTS

### 7.1 Observation

The long-term observation comprises the full period of the wheat growing season that was relevant for LAI monitoring, i.e., DOY 126–161. Figure 11 provides an overview on the collected sensor data during this observation period. Besides the RSSI curve of a representative link (Figure 11(a)), the main meteorological factors from the previous section gathered by the sensor deployment, except  $RH$ , are visualized (red). Additionally, external DWD data is included again (black) and also the corresponding LAI trajectory (from Reference [5], green), representing the vegetative development. Here, values are smoothened by an SMA with a larger window size than used in the short-term analysis ( $\pm 3$  days), which eliminates the diurnal pattern and focuses on the long-term influences. Apparently, all curves show, to some extent, more or less similarities. Again, there is a large discrepancy of  $AH$  readings from both sensors  $B$  and  $GW$ , indicated by the red-filled area in the graph that is caused by their different positions (ground vs. above), cf. Section 6.1. Note that the long-term temperature change is more moderate than extreme diurnal temperature fluctuations. Thus, the impact of temperature on the RSSI is less crucial on an averaged long-term scale and has only a little relevance here, as the following analysis will show. Hence, a possible temperature correction using the factor evaluated in the previous section is omitted consequently.

Figure 11(b) shows complementary meteorological variables that are occasionally included in related studies on LQs [8, 9, 19, 22, 24, 31]. Other variables, e.g., wind speed, despite its influence on transmittance-based LAI estimation in wheat cultivars [5], could reasonably be neglected. The time series of three selected variables, solar radiation ( $SR$ ), cloud coverage ( $CC$ ), and precipitation ( $Rain$ ) are obtained from the DWD. Mean curves are computed by an SMA over  $\pm 3$  days, except for precipitation, which is averaged over a period of only  $\pm 6$  hours to address its sporadic occurrence.

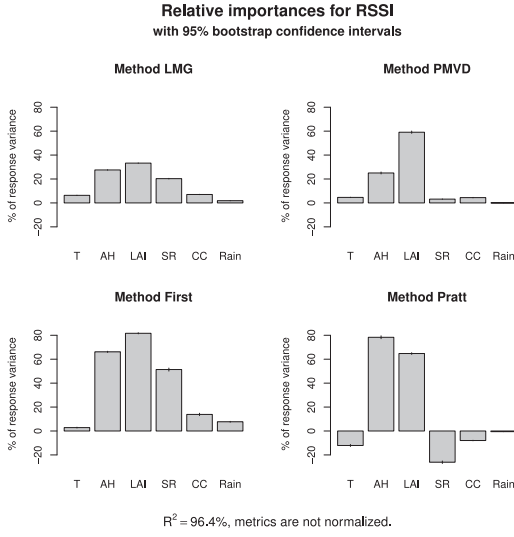


Fig. 12. Relative importance of various meteorological factors and crop growth (LAI) for long-term RSSI variance determined with four different decomposition metrics using the extended RSSI model (Equation (3)).

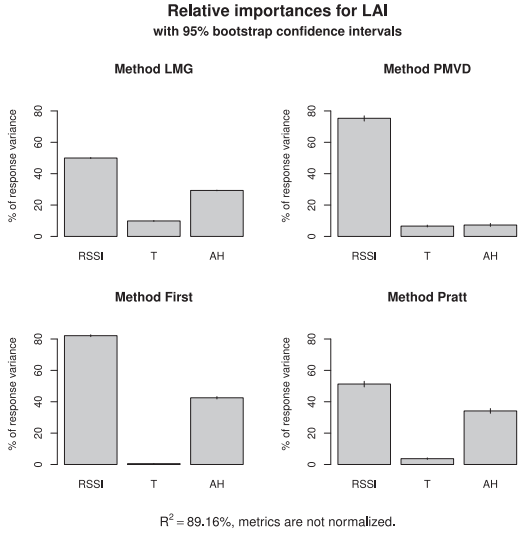


Fig. 13. Relative importance of predictors, RSSI,  $T$ , and  $AH$ , that were identified to be essential for the long-term LAI estimation, determined using the RSSI-based LAI model (Equation (5)).

The first two subplots show  $SR$ , particularly the short-wave irradiance that is further differentiated into direct and diffuse irradiance, first, at hourly scale, and second, at daily means curves.  $CC$ , measured in one-eighths increments, shows mostly mixed conditions with only a few stable periods. Generally, there was scarce rainfall as the precipitation subplot indicates.

## 7.2 Analysis

For the analysis of relevant variables having a long-term impact on RSSI, their relative importance is investigated once again. Therefore, the initial model (Equation (2)) is extended as follows:

$$RSSI \sim T + AH + LAI + SR + CC + Rain. \quad (3)$$

Again, the results show that also for long-term observations, almost all RSSI variances can be explained by the model ( $R^2 \approx 0.96$ , cf. Figure 12). The significant influences of LAI and  $AH$  are clearly emphasized by both metrics,  $PMVD$  and  $Pratt$ . In contrast,  $SR$ ,  $CC$ , and  $Rain$  were found to have no substantial effects on RSSI. The negligible effect of precipitation might, however, be caused by its rare presence. Likewise,  $T$  has very little relevance in the long-term observation, as opposed to its high importance observed for a short-term period (Section 6). This can be explained by the larger SMA window that eliminated diurnal variances.

In addition to the multiple regression above, the relative importance is evaluated in some different setups, i.e., with different link selections or different models, as listed in Table 2. The results obtained by these different setups demonstrate that, if only CAT 1 links are included, then the overall  $R^2$  and also the importance of  $AH$  decreases (Setup  $b$  compared to  $a$  and  $c$ ). The latter indicates that, as opposed to the impact of  $T$  on radio hardware, an  $AH$  increase seems to negatively affect the communication medium, thus, to amplify the path loss. Thus, shorter distances of CAT 1 links apparently decrease the impact of  $AH$ . For the sake of simplicity, regressors that are found

Table 2. Various Setups for the Multiple Linear Regression and Different Contributions of Regressors

Setup	RSSI Input	Resulting contribution to linear model (%) <sup>*</sup>							Model
		$R^2$	$T$	$AH$	$LAI$	$SR$	$CC$	$Rain$	
<i>a</i>	CAT 1+2	96.4	5	25	59	3	4	0	Equation (3) (Figure 12)
<i>b</i>	CAT 1	75.8	1	1	66	1	6	0	Equation (3)
<i>c</i>	CAT 2	95.8	7	41	41	4	3	0	Equation (3)
<i>d</i>	CAT 1+2	92.9	5	30	57	–	–	–	Equation (4)

<sup>\*</sup> based on PMVD relative importance model (regressors are not normalized).

to be barely relevant are reasonably dropped from the model:

$$RSSI \sim T + AH + LAI, \quad (4)$$

which only leads to a slightly lower overall  $R^2$  (Setup *d* in Table 2) and, thus, seems suitable for the model inversion in the next section.

## 8 POTENTIAL OF RSSI-BASED LAI ESTIMATION

### 8.1 LAI Modeling

The previous analysis of variables having a long-term impact on signal strength has shown that there is a correlation between RSSI and LAI. That means RSSI is evidently affected by vegetation. But there is also a meteorological influence. For this reason, a simplified model (Equation (4)) was proposed that includes both meteorological variables that were identified to have significant impacts, namely,  $T$  and  $AH$ . Now, this model is inverted as follows to predict LAI values using RSSI measurements together with  $T$  and  $AH$  information:

$$LAI \sim RSSI + T + AH. \quad (5)$$

Using the same data set as in the previous section, namely, the average of all CAT 1 and 2 links, the relative importance analysis of this multiple linear regression model reveals an overall  $R^2$  of 0.89. As expected, the analysis also confirms RSSI as the major predictor for LAI variances across all decomposition metrics (Figure 13). In contrast, a significant lower  $R^2$  of 0.82 is achieved by the linear model ( $LAI \sim RSSI$ ) that neglects the meteorological impact, as can be derived from the *First* method that determines the direct contribution of each regressor (cf. Section 6.2.2).

### 8.2 Application to RSSI Data

By using ordinary least squares (OLS) regression, this linear model is fitted to the collected data of our experiments. To this end, firstly, an SMA was applied to all CAT 1 and 2 links. Then, the average over all links is used. As a result, coefficients for each predictor and an intercept parameter are obtained that allow estimating LAI values. As those coefficients are expected to be highly crop- and site-specific and further depend on the actual deployment, they are probably not fully transferable. Thus, the determined coefficients are not explicitly listed here. However, they are used to visualize the potential of the RSSI-based LAI model in this section.

Figure 14 compares the modeled LAI trajectories (blue) with the reference data (red), i.e., the actual LAI observation assessed by the sensor deployment. Overall, as already found by the relative importance analysis, a good agreement ( $R^2 = 0.88$ ) is achieved using the mean RSSI-based model, averaged over all CAT 1 and 2 links. However, Figure 14 also shows a certain deviation at the beginning of the experiment (DOY 126–128). Here, the reference LAI is overestimated by the model. The reasons for overestimation are unclear and demand further investigations and experimental



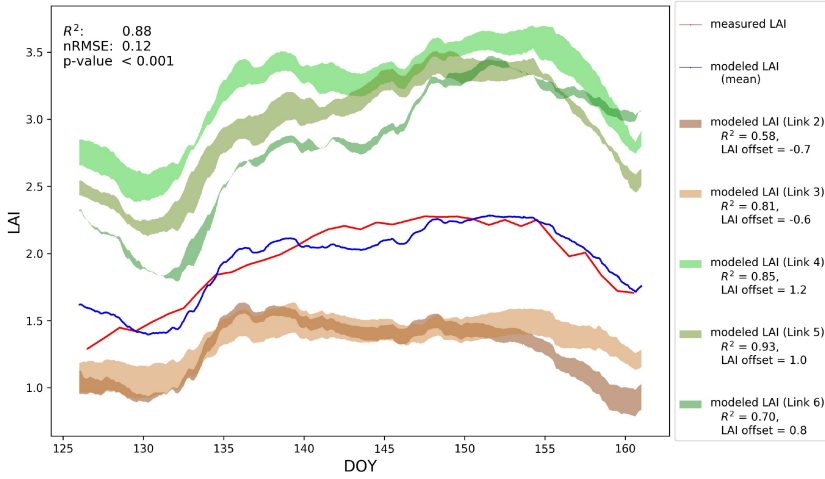


Fig. 14. Modeled mean RSSI-based LAI estimates (blue) in comparison with conventional transmittance-based LAI trajectories (red) over the wheat growing season. Besides, also modeled LAI trajectories of individual directed links (CAT 1 shaded in brown, CAT 2 in green colors) are included, which show similar trends but differ in their estimated LAI offsets.

replications. Moreover, the figure also comprises modeled LAI trajectories derived from RSSI time series of individual links from both categories, CAT 1 (shaded brown) and CAT 2 (shaded green).

On the one hand, it can be seen that all curves have a similar trend and more or less correlate with the trajectory of the reference data ( $0.58 \leq R^2 \leq 0.93$ ). A rough LAI estimation is thus found to be already possible with a single pair of transmitters.

On the other hand, there is an evident difference between curves from both categories, because CAT 1 links generally underestimate (LAI offset  $\approx -0.65$ ), whereas CAT 2 links generally overestimate (LAI offset  $\approx +1$ ) the actual LAI. This is caused by the joint OLS fitting (CAT 1 and 2 links) and the LQ, that is higher (CAT 1) or lower (CAT 2) than the average LQ. If the model is fitted to either only CAT 1 or only CAT 2 links, then that would not be the case. As a consequence, this difference points out that a generic calibration of the proposed model appears to be challenging.

### 8.3 Model Evaluation

**8.3.1 Impact of Sampling Rate.** The vast oversampling (2 min intervals) in our experiments was particularly helpful for the investigations of short-term dynamics. In practice, energy is usually the limiting resource. Because of the comparably high energy demand of radio communication (and radio sensing as well), a trade-off arises between energy consumption and sampling rate and, thus, probably also monitoring accuracy. Furthermore, for the potential integration of RSSI-based crop monitoring into existing agricultural systems, data transmissions, which are anyway necessary for the communication of other sensor data between the devices, might be exploited. Those transmissions usually are less frequent.

Hence, for the sake of practicability, the impact of the sampling rate on the quality of the proposed model is evaluated. For this purpose, the sampling rate is subsequently reduced step by step by removing corresponding samples from the raw data before continuing with the data processing and finally applying the model. For each reduced sampling rate, ten replications were performed, generated by biasing the periodic sampling in the time of day. The results regarding the correlation of the modeled with reference LAI trajectories are shown in Figure 15(a). It can be seen that a drastic reduction of the sampling rate is possible without a relevant loss of quality, even if an LQ

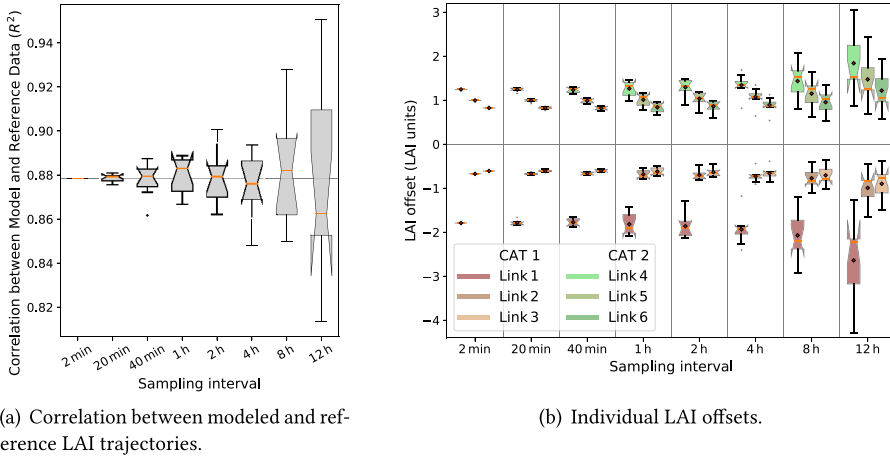


Fig. 15. Impact of sampling rate on the quality of the proposed RSSI-based LAI model (Equation (5)).

sensing is only conducted twice a day. Merely the  $R^2$  variance increases slightly due to the lower sampling.

The offsets of estimated LAI trajectories result from the difference between the mean signal strength of a respective link and the mean signal strength of all links. They were already observed in Figure 14 and now found to be hardly affected by the sampling reduction, as well (Figure 15(b)). Offsets remain almost constant, but the variance increases in a similar way. Overall, a sampling interval of 8 h achieves a reasonable trade-off and is also considered compatible with the typical network traffic of agricultural monitoring systems.

**8.3.2 Cross-Validation.** So far, the model was trained with the averaged RSSI time series of all selected link pairs (see Section 8.2). To investigate the robustness of the model, its performance is now evaluated in the case when only a certain link is used for training. For this purpose, the data set, i.e., the set of (bidirectional) links, was divided into training and test data. The quality of the results (in terms of  $R^2$  between model and reference data) is compared, also with regard to different sampling rates. Again, ten replications were performed using time biases for the reduced sampling rates.

The results of all combinations of the considered link pairs are visualized in Figure 16 for (a) the original sampling rate and (b) for an exemplarily reduced rate of 3 samples/day. Here, the means of individual determination coefficients are shown. Each row in the table represents the results achieved by the model, trained with a different link pair applied to different validation links in the columns. The first row in Figure 16(a) summarizes the results of Figure 14. The diagonal entries correspond to results where the training and evaluation set are identical, whereas the other entries show the actual results of the cross-validation.

Overall, in many cases, the cross-validation reveals a good agreement between the modeled LAI trajectories and the reference data, regardless of the sampling rate, which has only a minor impact on the observed results, as the comparison of the two figures shows. Occasionally, there are some exceptions, i.e., links 1 and 5, that do not always achieve sufficient scores. However, as expected, CAT 2 links are found to perform very well (highlighted by the red box). Those links are particularly suited for both training and application of the model.

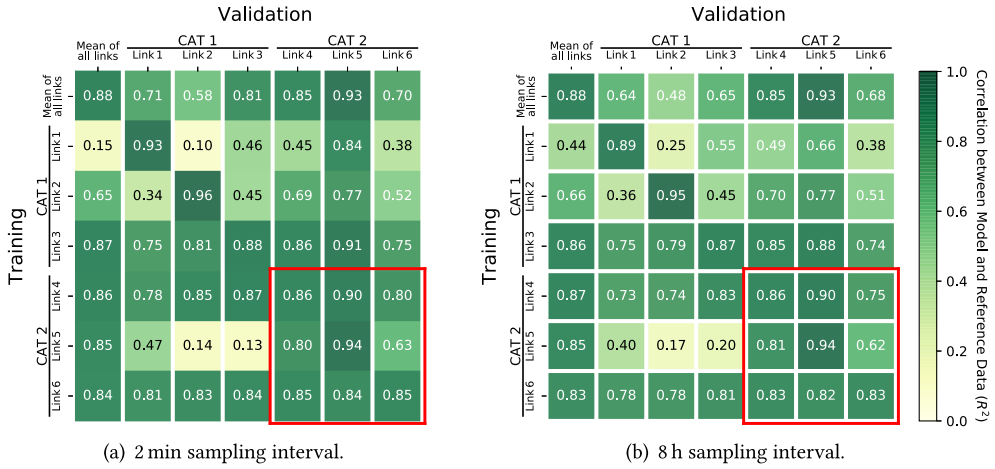


Fig. 16. Cross-validation of the RSSI-based LAI model (Equation (5)).

Hence, we conclude that in practice, the model should be trained with averaged RSSI time series of several medium-quality links. Also, for estimating reliable LAI trajectories, the model should ideally be applied to averaged LQs of multiple links. High sampling rates are not necessary, instead, moderate rates are quite sufficient, which makes our approach feasible for existing monitoring systems.

## 9 DISCUSSION

The empirical investigations using LQ and LAI data, collected in a long-term wheat field experiment, demonstrated the general feasibility of the proposed approach and showed first promising results. It unveiled the great potential of LQ sensing as enabling technology towards a complementary crop monitoring. The key findings of our analysis are:

- temperature ( $T$ ) has a significant impact on low-cost transceivers;
- absolute humidity ( $AH$ ) significantly affects the communication medium;
- both  $T$  and  $AH$  are the main influencing (meteorological) factors, others appear to be almost non-influential;
- vegetation has strong influence on *in situ* communication, particularly of ground sensors;
- RSSI is evaluated to be an appropriate LQ metric that reliably responds even to small changes induced by those factors;
- the integration of  $T$  and  $AH$  as influencing meteorological factors is important for RSSI-based crop monitoring;
- medium distance links between ground and above sensors (i.e., CAT 2) are appropriate for this purpose.

The first four findings are consistent with many studies in the literature and have various implications for network planning and protocol design. Reasonable recommendations for outdoor deployments, which are confirmed by our findings, have already been presented in Rankine et al. [24] and in Wennerström et al. [31].

### 9.1 Challenges

In contrast to photography- and transmittance-based approaches, RSSI-based crop monitoring does not require additional sensor hardware. Thus, existing agricultural WSNs, e.g., for soil

monitoring, could be easily retrofitted by simple software upgrades. Even existing radio transmissions of these monitoring systems could be leveraged by this approach, as low sampling rates proved to be sufficient. Hence, there are no extra costs, neither for hardware nor in terms of energy consumption.

However, our analysis also shows that the calibration of our model is challenging. The reason is that the perceived vegetation loss induced by crop growth depends on transmitter-receiver distances, transceivers positioning, and antenna orientations as experienced throughout our investigations. Also, many technical and transceiver-specific parameters, such as transmission power, antenna gain, receiver sensitivity, and signal modulation schemes, will certainly have calibration impacts. Future research in the direction of LQ-based crop monitoring should, therefore, consider such parameters for the design of sensors and systems. Also, utilizing frequency diversity, which is supported by typical IoT transceivers, could improve the robustness of LQ estimation and, thus, the introduced approach. Moreover, the calibration is also expected to be partially cultivar- and site-specific. In fact, the quantitative monitoring of absolute LAI values is challenging. However, deriving a rough LAI estimation and qualitative LAI trends that are useful for many applications was demonstrated to be possible.

## 9.2 Limitations of this Study

A precise evaluation of the accuracy of our approach is very challenging, because a real LAI ground truth as only provided by destructive measurements was not available for our experiments and is generally extremely rare. Instead, our analysis relies on transmittance-based LAI estimates, that are expected to contain a certain measurement uncertainty. Therefore, this study basically presents the general feasibility and can be considered as an initial proof-of-concept for RSSI-based wheat monitoring. Based on our research, more real-world experiments and replications are needed that, amongst others, also capture a more extensive range of different weather conditions, more transmitter-receiver distances and positioning setups, and also different transceivers. Although rainfall and fog appeared to be not influential, their sparse presence does not allow for credible conclusions regarding their impact. Moreover, our experiment was limited to the warm season (May/June). Thus, seasonal differences of meteorological effects, e.g., in winter months as shown by Luomala and Hakala [19], were neglected, although they might be relevant for the monitoring of other cultivars grown as catch crops.

Furthermore, our model is based on multiple linear regression. Possible future work, therefore, includes more sophisticated models and techniques of machine learning to further improve the potential of the proposed approach. For example, Gaussian Process Regression (GPR) could be considered, which is already applied in the context of remote sensing to combine *in situ* with space-borne LAI retrieval in a promising way [20, 32]. Finally, the data set collected in the long-term deployment used in this work has been released for the community.<sup>2</sup>

## 10 CONCLUSION

The experimental study conducted in wheat fields has shown that in-field radio communication using low-cost low-power IoT transceivers is significantly affected by vegetation. In the context of precision agriculture, this observation is exploited for a novel approach that inversely derives crop growth information from radio signal or link quality (LQ) variations, respectively. By using the leaf area index (LAI) as a representative metric for crop growth and conducting a comparative analysis, we evaluated the received signal strength indicator (RSSI) to be a suitable LQ metric. We found a strong correlation between both variables. However, our analysis also shows the influences of

<sup>2</sup>Data set available under <https://sys.cs.uos.de/wsn4crop>.

meteorological factors on LQ, particularly on RSSI. Two main influencing factors could be identified, which is the temperature that has an impact on transceivers' hardware and the absolute humidity that, in contrast, is assumed to affect the communication medium, similar to vegetation. Finally, a multiple linear model for RSSI-based LAI estimation was developed that consequently takes temperature and humidity impacts into account. As a proof-of-concept of our novel approach, this model was subsequently fitted to empirical data using OLS regression. Applying the model to collected RSSI, temperature, and humidity measurements revealed a promising and robust correlation of modeled and reference LAI estimates. This correlation successfully demonstrated the feasibility of our proposed approach for qualitative crop monitoring. However, the model is expected to be deployment-dependent. Particularly, link distances have an evident influence. Thus, for a quantitative LAI estimation, its calibration is essential and, at the same time, challenging in practice. In our future work, further deployments in different crops are planned to investigate the transferability of the model, and, as far as possible, also destructive LAI measurements will be conducted.

## ACKNOWLEDGMENTS

The authors thank Thomas Hänel for substantial LQ discussions, Siegfried Schittenhelm (JKI) for the great opportunity to realize the deployment at experimental crop fields of the JKI, Thomas Jarmer and Bastian Siegmann for deployment support and application insights that greatly assisted this research, and all anonymous reviewers for their valuable feedback, comments, and suggestions that helped improve this work.

## REFERENCES

- [1] Gregory Asner, Jonathan Scurlock, and Jeffrey Hicke. 2003. Global synthesis of leaf area index observations: Implications for ecological and remote sensing studies. *Global Ecol. Biogeogr.* 12, 3 (2003), 191–205. DOI : <https://doi.org/10.1046/j.1466-822X.2003.00026.x>
- [2] Nouha Baccour, Anis Koubâa, Luca Mottola, Marco Antonio Zúñiga, Habib Youssef, Carlo Alberto Boano, and Mário Alves. 2012. Radio link quality estimation in wireless sensor networks: A survey. *ACM Trans. Sensor Netw.* 8, 4 (2012), 1–34. DOI : <https://doi.org/10.1145/2240116.2240123>
- [3] Kenneth Bannister, Gianni Giorgetti, and Sandeep K. S. Gupta. 2008. Wireless sensor networking for “hot” applications: Effects of temperature on signal strength, data collection and localization. In *Proceedings of the 5th Workshop on Embedded Networked Sensors (HotEmNets'08)*. ACM, 1–5.
- [4] Jan Bauer and Nils Aschenbruck. 2018. Design and implementation of an agricultural monitoring system for smart farming. In *Proceedings of the IEEE IoT Vertical and Topical Summit for Agriculture (IOT'18)*. IEEE, 1–6. DOI : <https://doi.org/10.1109/IOT-TUSCANY.2018.8373022>
- [5] Jan Bauer, Thomas Jarmer, Siegfried Schittenhelm, Bastian Siegmann, and Nils Aschenbruck. 2019. Processing and filtering of leaf area index time series assessed by *in situ* wireless sensor networks. *Comput. Electron. Agric.* 165 (2019). DOI : <https://doi.org/10.1016/j.compag.2019.104867>
- [6] Jan Bauer, Bastian Siegmann, Thomas Jarmer, and Nils Aschenbruck. 2016. On the potential of wireless sensor networks for the *in situ* assessment of crop leaf area index. *Comput. Electron. Agric.* 128 (2016), 149–159. DOI : <https://doi.org/10.1016/j.compag.2016.08.019>
- [7] Beecham Research Ltd. 2016. *Enabling The Smart Agriculture Revolution—The Future of Farming through the IoT Perspective*. Technical Report. Retrieved from <http://www.beechamresearch.com/download.aspx?id=1051>.
- [8] Carlo Alberto Boano, James Brown, Zhitao He, Utz Roedig, and Thiemo Voigt. 2010. Low-power radio communication in industrial outdoor deployments: The impact of weather conditions and ATEX-compliance. In *Proceedings of the 1st International Conference on Sensor Applications, Experimentation, and Logistics (SENSAPPEAL'10)*. Springer, 159–176. DOI : [https://doi.org/10.1007/978-3-642-11870-8\\_11](https://doi.org/10.1007/978-3-642-11870-8_11)
- [9] James Brinkhoff and John Hornbuckle. 2017. Characterization of wifi signal range for agricultural WSNs. In *Proceedings of the 23rd Asia-Pacific Conference on Communications (APCC'17)*. IEEE, 1–6. DOI : <https://doi.org/10.23919/APCC.2017.8304043>
- [10] Qiang Chen, DaeHee Won, Dennis M. Akos, and Eric E. Small. 2016. Vegetation sensing using GPS interferometric reflectometry: Experimental results with a horizontally polarized antenna. *IEEE J. Select. Top. Appl. Earth Observ. Remote Sens.* 9 (2016), 4771–4780. DOI : <https://doi.org/10.1109/JSTARS.2016.2565687>



- [11] Datasheet: CC2420. 2017. 2.4 GHz IEEE 802.15.4 / ZigBee-ready RF Transceiver. Technical Report. Texas Instruments. Retrieved from <https://www.ti.com/lit/ds/symlink/cc2420.pdf>.
- [12] Björn Gernert, Jan Schlichter, and Lars Wolf. 2019. PotatoScanner—A mobile delay tolerant wireless sensor node for smart farming applications. In *Proceedings of the 15th IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS'19)*. IEEE, Santorini Island, Greece, 106–113. DOI : <https://doi.org/10.1109/DCOSS.2019.00037>
- [13] Ulrike Grömping. 2006. Relative importance for linear regression in R: The package relaimpo. *J. Stat. Softw.* 17 (2006), 1–27. Issue 1. DOI : <https://doi.org/10.18637/jss.v017.i01>
- [14] ITU-R. 2016. *Recommendation P.833-9—Attenuation in vegetation*. Technical Report. International Telecommunication Union (ITU). Retrieved from <https://www.itu.int/rec/R-REC-P.833/en>.
- [15] ITU-R. 2017. *Recommendation P.618-13—Propagation Data and Prediction Methods Required for the Design of Earth-space Telecommunication Systems*. Technical Report. International Telecommunication Union (ITU). Retrieved from <https://www.itu.int/rec/R-REC-P.618/en>.
- [16] Ulf Kulau, Stephan Rottmann, Sebastian Schildt, Johannes van Balen, and Lars Wolf. 2016. Undervolting in real world WSN applications: A long-term study. In *Proceedings of the 12th IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS'16)*. IEEE, 9–16. DOI : <https://doi.org/10.1109/DCOSS.2016.15>
- [17] Siyu Li and Hongju Gao. 2011. Propagation characteristics of 2.4 GHz wireless channel in cornfields. In *Proceedings of the 13th IEEE International Conference on Communication Technology (ICCT'11)*. IEEE, 136–140. DOI : <https://doi.org/10.1109/ICCT.2011.6157849>
- [18] Ji Luo, Xing Xu, and Qian Zhang. 2011. Understanding link feature of wireless sensor networks in outdoor space: A measurement study. In *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM'11)*. IEEE, 1–5. DOI : <https://doi.org/10.1109/GLOCOM.2011.6134117>
- [19] Jari Luomala and Ismo Hakala. 2015. Effects of temperature and humidity on radio signal strength in outdoor wireless sensor networks. In *Proceedings of the Federated Conference on Computer Science and Information Systems (FedCSIS'15)*. IEEE, 1247–1255. DOI : <https://doi.org/10.15439/2015F241>
- [20] Manuel Campos-Taberner and Francisco Javier García-Haro and Gustau Camps-Valls and Gonçál Grau-Muedra and Francesco Nutini and Alberto Crema and Mirco Boschetti. 2016. Multitemporal and multiresolution leaf area index retrieval for operational local rice crop monitoring. *Remote Sens. Environ.* 187 (2016), 102–118. DOI : <https://doi.org/10.1016/j.rse.2016.10.009>
- [21] Ramona Marfievici, Amy L. Murphy, Gian Pietro Picco, Federico Ossi, and Francesca Cagnacci. 2013. How environmental factors impact outdoor wireless sensor networks: A case study. In *Proceedings of the 10th IEEE International Conference on Mobile Ad-Hoc and Sensor Systems (MASS'13)*. IEEE, 565–573. DOI : <https://doi.org/10.1109/MASS.2013.13>
- [22] Andrew Markham, Niki Trigoni, and Stephen Ellwood. 2010. Effect of rainfall on link quality in an outdoor forest deployment. In *Proceedings of the International Conference on Wireless Information Networks and Systems (WINSYS'10)*. IEEE, 1–6.
- [23] Yonghua Qu, Yeqing Zhu, Wenchao Han, Jindi Wang, and Mingguo Ma. 2014. Crop leaf area index observations with a wireless sensor network and its potential for validating remote sensing products. *IEEE J. Select. Topics Appl. Earth Observ. Remote Sens.* 7, 2 (Feb. 2014), 431–444. DOI : <https://doi.org/10.1109/JSTARS.2013.2289931>
- [24] Cassidy J. Rankine, G. Arturo Sanchez-Azofeifa, and Mike H. MacGregor. 2014. Seasonal wireless sensor network link performance in boreal forest phenology monitoring. In *Proceedings of the 11th IEEE International Conference on Sensing, Communication, and Networking (SECON'14)*. IEEE, 302–310. DOI : <https://doi.org/10.1109/SAHCN.2014.6990366>
- [25] Jürgen Richter, Rafael F. S. Caldeirinha, Miqdad O. Al-Nuaimi, Andy Seville, Neil C. Rogers, and Nick Savage. 2005. A generic narrowband model for radiowave propagation through vegetation. In *Proceedings of the 61st IEEE Vehicular Technology Conference (VTC'05)*, Vol. 1. IEEE, 39–43. DOI : <https://doi.org/10.1109/VETECS.2005.1543245>
- [26] Siegfried Schittenhelm, Lorenz Kottmann, Martin Kraft, Katja Matschiner, and Tina Langkamp-Wedde. 2018. Agronomic performance of winter wheat grown under highly divergent soil moisture conditions in rainfed and water-managed environments. *J. Agron. Crop Sci.* 205, 3 (2018), 283–294. DOI : <https://doi.org/10.1111/jac.12322>
- [27] Florian Schmidt, Matteo Ceriotti, Niklas Hauser, and Klaus Wehrle. 2015. If you can't take the heat: Temperature effects on low-power wireless networks and how to mitigate them. In *Proceedings of the 12th European Conference on Wireless Sensor Networks (EWSN'15)*. Springer, 266–273. DOI : [https://doi.org/10.1007/978-3-319-15582-1\\_19](https://doi.org/10.1007/978-3-319-15582-1_19)
- [28] Kannan Srinivasan and Philip Levis. 2006. RSSI Is Under-appreciated. In *Proceedings of the 3rd Workshop on Embedded Networked Sensors (EmNets'06)*. 1–5. Retrieved from <https://sing.stanford.edu/site/publications/14>.
- [29] John Thelen, Daan Goense, and Koen Langendoen. 2005. Radio wave propagation in potato fields. In *Proceedings of the 1st Workshop on Wireless Network Measurement (WinMee'05)*. IEEE, 1–5.
- [30] Deepak Vasisht, Zerina Kapetanovic, Jong-ho Won, Xinxin Jin, Ranveer Chandra, Ashish Kapoor, Sudipta N. Sinha, Madhusudhan Sudarshan, and Sean Stratman. 2017. Farmbeats: An IoT platform for data-driven agriculture.

- In *Proceedings of the 14th USENIX Symposium on Networked Systems Design and Implementation (NSDI'17)*. USENIX Association, 515–529. Retrieved from <https://www.usenix.org/system/files/conference/nsdi17/nsdi17-vasisht.pdf>.
- [31] Hjalmar Wennerström, Frederik Hermans, Olof Rensfelt, Christian Rohner, and Lars-Åke Nordén. 2013. A long-term study of correlations between meteorological conditions and 802.15.4 link performance. In *Proceedings of the 10th IEEE International Conference on Sensing, Communications and Networking (SECON'13)*. IEEE, 221–229. DOI:<https://doi.org/10.1109/SAHCN.2013.6644981>
- [32] Gaofei Yin, Aleixandre Verger, Yonghua Qu, Wei Zhao, Baodong Xu, Yelu Zeng, Ke Liu, Jing Li, and Qinhua Liu. 2019. Retrieval of high spatiotemporal resolution leaf area index with gaussian processes, wireless sensor network, and satellite data fusion. *Remote Sens.* 11, 244 (2019), 1–18. DOI:<https://doi.org/10.3390/rs11030244>

Received April 2019; revised January 2020; accepted April 2020