



REGULAR ARTICLE

WILEY

Kriging-based robotic exploration for soil moisture mapping using a cosmic-ray sensor

Jaime Pulido Fentanes¹ | Amir Badiee³ | Tom Duckett¹ | Jonathan Evans² | Simon Pearson³ | Grzegorz Cielniak¹

¹Lincoln Centre for Autonomous Systems, School of Computer Science, University of Lincoln, Lincoln, United Kingdom

²Centre for Ecology & Hydrology, CEH Wallingford, Oxfordshire, United Kingdom

³Lincoln Institute for Agri-food Technology, University of Lincoln, Lincoln, United Kingdom

Correspondence

Jaime Pulido Fentanes, Lincoln Centre for Autonomous Systems, School of Computer Science, University of Lincoln, Brayford Campus, LN6 7TS, Lincoln, UK.
Email: jpulidofentanes@lincoln.ac.uk

Funding information

Science and Technology Facilities Council (STFC) Newton Fund programme, Grant/Award Number: ST/N006836/1

Abstract

Soil moisture monitoring is a fundamental process to enhance agricultural outcomes and to protect the environment. The traditional methods for measuring moisture content in the soil are laborious and expensive, and therefore there is a growing interest in developing sensors and technologies which can reduce the effort and costs. In this work, we propose to use an autonomous mobile robot equipped with a state-of-the-art noncontact soil moisture sensor building moisture maps on the fly and automatically selecting the most optimal sampling locations. We introduce an autonomous exploration strategy driven by the quality of the soil moisture model indicating areas of the field where the information is less precise. The sensor model follows the Poisson distribution and we demonstrate how to integrate such measurements into the kriging framework. We also investigate a range of different exploration strategies and assess their usefulness through a set of evaluation experiments based on real soil moisture data collected from two different fields. We demonstrate the benefits of using the adaptive measurement interval and adaptive sampling strategies for building better quality soil moisture models. The presented method is general and can be applied to other scenarios where the measured phenomena directly affect the acquisition time and need to be spatially mapped.

KEYWORDS

agriculture, environmental monitoring, exploration, terrestrial robotics

1 | INTRODUCTION

Management of water resources is of considerable concern in different parts of the world, with many areas facing prolonged droughts, while others experience devastating floods. The availability of water in the soil is essential for vegetation. In an agricultural setting, crop health depends greatly on soil moisture. It is precisely for this reason that soil moisture monitoring is key to improving agricultural processes. Perhaps the most obvious advantage of technologies for obtaining high-resolution soil moisture maps is that they would enable highly efficient irrigation planning, for example,

providing an accurate estimate of the quantity of water that should be put into a field and its required spatial distribution across the field.

Soil moisture is typically assessed either by a direct but lengthy procedure involving collecting physical soil samples followed by lab measurements, or by hand-held instruments used to measure moisture indirectly through proxies such as surface tension (manometers), or changes in soil conductivity (e.g., time-domain reflectometry; Noborio, 2001). All of these methods are very laborious, time-consuming and expensive. Recent advances in sensing technology introduced a new, noncontact method for measuring soil moisture using fast neutron detectors (Zreda, Desilets, Ferr, & Scott, 2008). The neutrons are

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Journal of Field Robotics* Published by Wiley Periodicals, Inc.

generated by cosmic rays and are reflected from the soil. The reflected neutron count is directly proportional to soil moisture content. Such sensors were successfully deployed at static locations covering large areas of land (Evans et al., 2016) but also as high-resolution variants with a reduced field of view and increased sensitivity (Schrön et al., 2017).

The most common method for creating soil moisture maps is to use data that are manually collected at predetermined locations in the field and extrapolate the expected measurements for unvisited regions using kriging or Gaussian process regression (Matheron, 1963; Williams & Rasmussen, 2006). This is a costly and laborious process, especially in the case of soil moisture monitoring, where the methods and instruments used to take measurements across the field require a high amount of labor and postprocessing. For this reason, there is a growing interest in developing instruments and methodologies to help reduce the effort and costs while improving the quality of the resulting soil moisture models.

In this work, we propose to use an autonomous mobile robot equipped with a noncontact soil moisture sensor that builds soil moisture maps on the fly and automatically selects the most optimal sampling locations. The robot is guided by an autonomous exploration strategy driven by the quality of the soil moisture model (i.e., kriging variance [KV]) which indicates areas of the field where the information is less precise, improving overall model quality. The employed fast neutron counting sensors provide a special category of measurements in which the acquisition time directly depends on the intensity of the phenomenon: In our case, the sensor registers more neutrons in drier soils. We model the sensor using the Poisson distribution and use a special kriging variant for this type of measurements. As a result, the exploration strategy plans not only the optimal sampling location but also the required acquisition time at each sampling location.

The contributions of this work are as follows:

- Application of a novel fast neutron counting sensor for robotic-assisted spatial mapping of soil moisture;
- Integration of the Poisson measurement model into the kriging estimation and exploration framework, which devises optimal spatial locations and measurement intervals, improving the resulting moisture models;
- Evaluation and validation of the proposed framework on data collected from two different field environments.

The remainder of the paper is structured as follows: Section 2 presents related work in soil moisture surveying and robotic exploration, followed by Section 3, which details our approach to Poisson kriging (PK) and exploration for soil moisture mapping using a mobile robot. The experimental framework is presented in Section 4, followed by results and their analysis in Section 5, and final conclusions in Section 6.

2 | RELATED WORK

Robotic environmental monitoring applications have attracted a lot of attention in the last few years (Dunbabin & Marques, 2012). One of the advantages of using robots for environmental modeling and

monitoring is that they can build models on the fly. At the same time, many authors have discussed how to use the model itself to plan new observations for data acquisition that improve the overall model. For example, Kerry, Oliver, and Frogbrook, (2010) demonstrated that kriging semivariograms are highly useful for sampling planning in precision agriculture. They proposed to use ancillary information to estimate a semivariogram and thus determine the spatial frequency of sampling based on the semivariogram parameters.

Other researchers (Oliver & Webster, 1986) propose the generation of an initial set of samples to obtain a semivariogram that can be extrapolated to find new sample positions. B. Marchant and Lark (2007) proposed an adaptive approach for optimizing reconnaissance surveys. They sampled at preplanned positions and calculated the probability density function of the sampling density required for the main survey in a Bayesian framework. If the requirements were not met, the number and location of observations within further phases were selected to reduce the uncertainty of the required sampling density. However, the effort required to survey a soil variable and simultaneously build and analyze the variance of the kriging model of the soil meant that these authors stopped short of planning the entire sampling procedure based on kriging models.

Robots, on the other hand, are able to create and update models of their operational environments through robotic exploration. A common approach is to plan trajectories that completely cover the area assuming some prior knowledge of the environment (Rodias et al., 2017). Other well-known exploration techniques drive the robot towards unmapped areas of the environment. For example, greedy approaches such as Koenig, Tovey, and Halliburton (2001) drive the robot towards the nearest location where new information can be gained. In frontier-based exploration (Yamauchi, 1997), the robot is driven towards the boundary between the known and unknown parts of the environment, while information-driven “next-best-view” methods use reward functions to predict the utility of an unexplored location (Pulido Fentanes, Zalama, & Gomez-Garcia-Bermejo, 2011). Authors like O’Callaghan and Ramos (2011); Vasudevan, Ramos, Nettleton, and Durrant-Whyte (2009) propose the use of environmental representations that are based in Gaussian processes, they argue that these representations overcome many of the limitations of occupancy grid maps such as scale and provide information about model quality which can be extremely useful for robotic exploration (Jadidi, Miró, Valencia, & Andrade-Cetto, 2014).

Many authors have proposed informative path planning (IPP) techniques for modeling physical phenomena with an unknown spatial distribution. These techniques address how to plan a path that maximizes sensor information (Binney, Krause, & Sukhatme, 2013) and can be classified into two approaches: Those that depend only on a priori information about the environment (Hollinger & Sukhatme, 2013) and adaptive sampling techniques that can be modified depending on the observations made (Sadat, Wawerla, & Vaughan, 2015). Yang, Keat Gan, and Sukkarieh (2013) propose to use a Gaussian process to model the occupancy of a cluttered environment and use Randomly-exploring Random Trees to guarantee a collision-free full exploration of the

environment. Martinez-Cantin, de Freitas, Doucet, and Castellanos, (2007) propose to use Gaussian process variance to plan paths that increase robot knowledge of the environment whilst minimizing position uncertainty. In Ghaffari Jadidi, Valls Miro, and Dissanayake (2019) an approach to generate dense maps using incremental information gathering is proposed, in this work the authors present a framework that uses an automatic stopping criteria based on information gain versus information gathering cost and simultaneously considers the uncertainty of the robots position and environment model to plan the robot trajectory.

Some authors have opted to use IPP to model different variables to plan robot actions. For example, Gao et al. (2018) propose the use of an informative sampling technique to minimize the total distance traveled by a fleet of phenotyping robots. To do this, they model the environment using Gaussian processes and use the model variance to plan the most informative paths for the fleet. R. Marchant and Ramos (2014) use Gaussian processes to plan the paths that guarantee both to observe the phenomenon of interest and improve the modeling of the same phenomenon for environmental monitoring applications such as ozone concentration across the USA. More recently, Popovic et al. (2017) proposed an adaptive IPP methodology to map green biomass in an agricultural setting.

Other authors have chosen to use OK to model in-field phenomena. Glaser, Schaefer, and Burgard (2018) use it to model soil properties perceived with a multispectral camera, and then use the resulting model to improve the robot localization. Diggle, Tawn, and Moyeed (1998) demonstrated that using kriging methods designed for Gaussian variables with Poisson processes can over-smooth the data and underestimate the spatial extremes of the intensity, for this reason, they proposed a new distributional framework which allows embedding nonlinear data in a linear kriging methodology. An alternative solution for this problem (Goldberg, Williams, & Bishop, 1998; Kersting, Plagemann, Pfaff, & Burgard, 2007) is to use Gaussian processes that model variables and its incertitude as independent Gaussian processes, these methods are known as heteroscedastic Gaussian process regression.

Within the kriging family, Kim and Shell (2014) proposed augmentation of OK to enable modeling of ocean current dynamics which they use for adaptive path planning in the field in ocean multi-robot scenarios. Pulido Fentanes, Gould, Duckett, Pearson, and Cielniak (2018a) proposed a robotic exploration methodology aimed at building soil condition maps using ordinary KV as a reward function for exploration. The current work builds upon this approach to model soil moisture measured with a novel sensor that does not follow a normal distribution. To achieve this we combine PK with a kriging-based exploration methodology.

3 | METHODOLOGY

In this work, we propose a kriging-based exploration pipeline for agricultural mobile robots to facilitate efficient mapping of soil moisture. The framework combines a unique sensor model, an online

spatial mapping component and an exploration strategy to guide the robot to the next best sampling location.

We consider a special category of measurements which are based on counting, and hence follow a Poisson distribution. An inherent property of such measurements is that their uncertainty directly depends on the length of the measurement interval. In our scenario, we use a robot-mounted soil moisture sensor (see Section 3.1) which counts low energy neutrons as a proxy for soil moisture. Therefore, the soil moisture level will affect the amount of time the robot spends at each sampling location. For the spatial mapping, we use a version of ordinary kriging (OK) which incorporates measurements following a Poisson distribution (see Section 3.3). We use the KV as a reward function for the exploration strategy to plan the optimal location for each subsequent measurement. Section 3.4 discusses the different exploration strategies that have been applied in this work.

The original kriging framework was presented in our previous work for mapping soil compaction (Pulido Fentanes et al., 2018a). In this paper, we generalize and extend the approach to take into account measurements following a Poisson distribution. This results in exploration strategies which not only consider the optimal sampling location but also adjust the measurement duration for each reading to ensure a high-quality model.

3.1 | Soil moisture measurement using a cosmic-ray sensor

The main sensor used in this work is based on measuring fast neutrons, which are generated by cosmic rays and reflected from the soil (Zreda et al., 2008). The intensity of the reflected neutrons is affected by the hydrogen in the soil, and hence provides an indication of the soil moisture content. A neutron detector is a tube containing a gas that can convert thermal neutrons into detectable electrons by ionization. Since the detectors are sensitive to fast neutrons only, the low energy neutrons (after colliding with the hydrogen atoms) are not counted. As a result, a higher neutron count means more fast neutrons and corresponds to dryer soil. To improve the sensitivity of the detector to fast neutrons, a polyethylene shield is used as a moderator.

Several correction procedures need to be applied on the acquired neutron counts (which we refer to as the raw neutron count N_{raw}) to account for variations in background cosmic-ray intensity, atmospheric pressure, and humidity (Evans et al., 2016). The reference values for the corrections are established during a calibration procedure which requires reference soil moisture values to be established by direct soil moisture measurements using traditional equipment. The correction factors include:

- Cosmic-ray intensity:

$$F_C = \frac{C_0}{C}, \quad (1)$$

where C is the measured neutron count rate (from the nearest monitoring station) and C_0 is the value measured during calibration.

- Pressure:

$$F_P = \exp[\beta(P - P_0)], \tag{2}$$

where P is the measured barometric pressure (using a barometer), P_0 is an arbitrary reference value (e.g., 1,010 hPA) and β is the barometric pressure coefficient established during calibration.

- Humidity:

$$F_Q = 1 + 0.00054(Q - Q_0), \tag{3}$$

where Q is the measured humidity (derived from temperature measurements) and Q_0 is the average humidity during calibration.

The corrected neutron count N_{crr} is obtained by multiplying the raw neutron counts by the correction factors:

$$N_{crr} = N_{raw} \cdot F_P \cdot F_Q \cdot F_C. \tag{4}$$

N_{crr} can then be used to calculate volumetric water content (VWC), which provides the final measure of the soil moisture. Since in this paper we mainly work with the corrected neutron counts N_{crr} , we refer the interested readers to Evans et al. (2016) for further detail of the exact conversion procedure.

The summarized methodology for measuring soil moisture has been used successfully by Evans et al. (2016), who have established a network of soil moisture monitoring stations in the UK covering an area of 12 ha. Although this coverage is useful for large scale soil moisture assessment, its application to individual fields in agriculture is limited. To achieve higher spatial resolutions, we have employed a high-sensitivity version of the sensor consisting of 12 neutron detectors with a bespoke polyethylene shield to limit the detection footprint of the sensor to approximately 10 m. The sensor mounted on our agricultural mobile robot Thorvald can be seen in Figure 3.

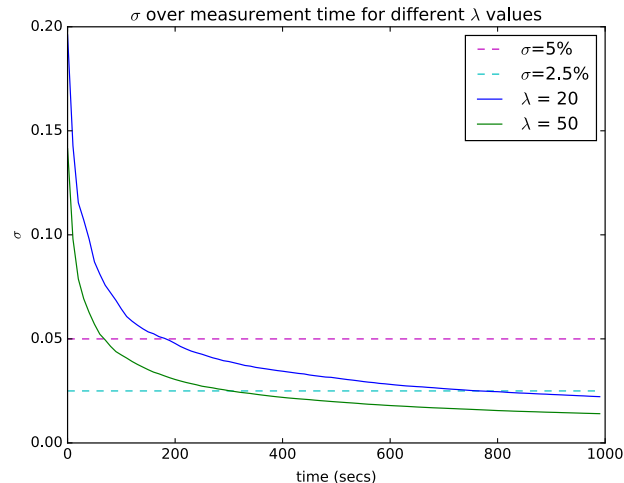
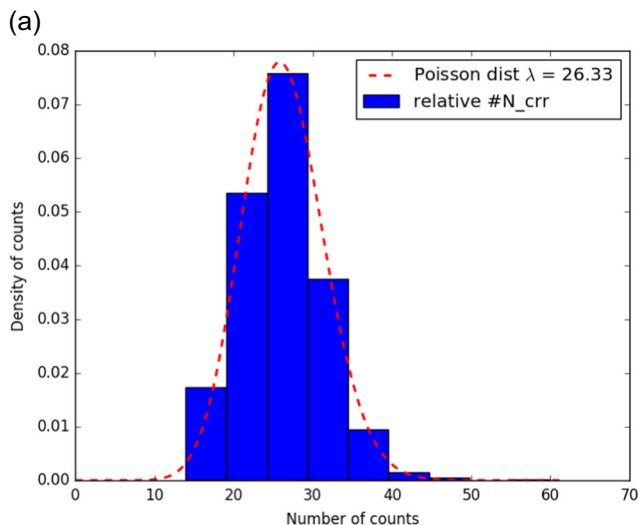


FIGURE 2 Measurement uncertainty σ over time for different λ values and sample thresholds for the adaptive measurement interval regime [Color figure can be viewed at wileyonlinelibrary.com]

3.2 | Poisson distribution measurements and sampling regime

Our soil moisture sensor provides the corrected neutron counts N_{crr} . The appropriate probabilistic model for modeling count data and events is the Poisson distribution, with parameter λ representing the average count rate over a period of 10 s. However, the uncertainty σ in the measurement depends directly on total neutron count over the measurement time, and is calculated as follows:

$$\sigma = \frac{\sqrt{N_{crr}}}{N_{crr}}. \tag{5}$$

Figure 1 shows the histogram reading and the evolution of the λ and σ parameters for the same measurement over time. Figure 1 shows how the standard error and variance decrease

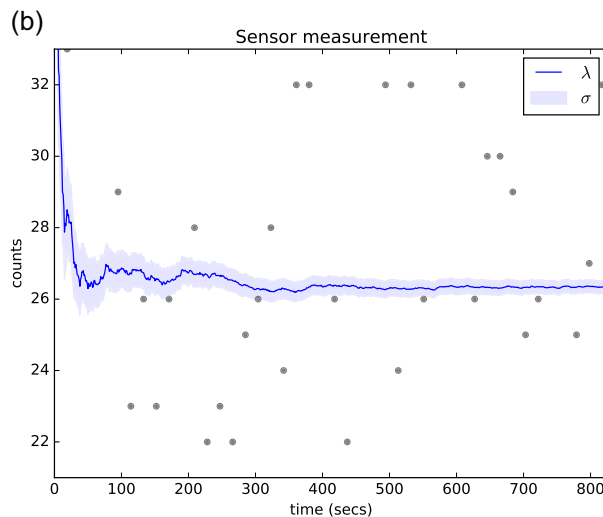


FIGURE 1 An example measurement from the cosmic-ray sensor (a) distribution of fast neutron counts, (b) evolution of the count rate and measurement uncertainty over time, black dots denote 1 out of 20 sensor readings for illustrative purposes [Color figure can be viewed at wileyonlinelibrary.com]

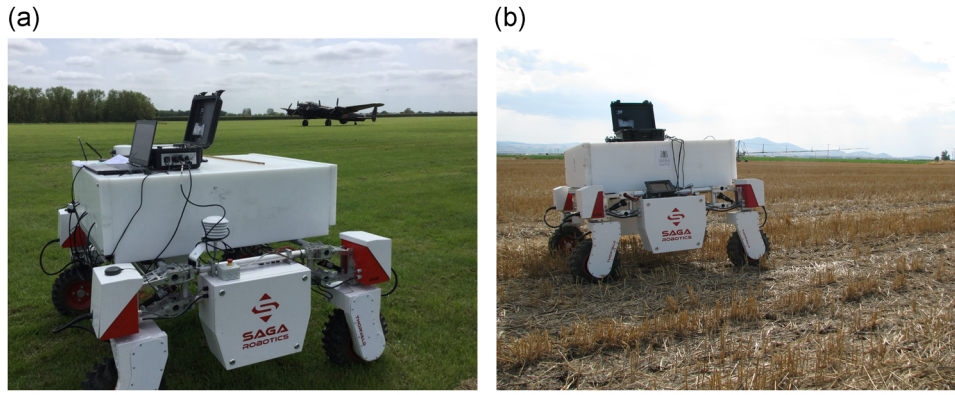


FIGURE 3 The Thorvald robot equipped with a cosmic-ray sensor during data collection at an airfield at the Lincolnshire Aviation Heritage Center in East Kirkby, UK (left); a wheat stubble field near Volos, Greece (right) [Color figure can be viewed at wileyonlinelibrary.com]

over time, meaning that readings with longer duration achieve higher quality.

The sampling regime is the criterion used to decide how long each measurement should last. In this scenario, the quality of the measurement is directly correlated to the number total number of observed events (N_{corr}). For this reason, we propose to use two different methodologies using fixed measurement intervals (FMI), in which each measurement lasts for a predetermined amount of time, or adaptive measurement intervals (AMIs), under which each measurement will last until a minimum level of quality is obtained. This paper compares both regimes and analyses what happens to the exploration process with each sampling regime, and more specifically, what is their effect on the final model quality.

The AMIs regime uses a threshold typically defined in terms of σ_m (see Equation (5)) to determine the duration of a measurement. In practice, this means that in this case, the robot will stay at each location until the normalized standard error falls below a predetermined percentage of the total amount of counts, so that the robot will stay longer in places where the count rates are lower (or the soil is wetter in this scenario) and spend less time in locations with higher count rates.

Figure 2 illustrates the evolution of the normalized standard error (σ_m) over time for different rates (λ), where the dashed lines indicate thresholds that can be used for this sampling regime, the time at which the threshold lines intersect the standard error lines, represents the point at which the measurement is considered complete. This guarantees a maximum uncertainty limit for each measurement which adapts to the actual neutron rate forcing the robot to stay longer at places where the rate of events is lower than usual or to leave as soon as possible in places with higher rates.

3.3 | Poisson kriging

OK has proven to be an effective method for interpolating spatial data when the data's main source of error is intrinsic to the measurement technique, for example, when it depends on the precision of an instrument. However, when the variance of the

measurement depends on the phenomenon itself, as in the case of events that can be modeled using a Poisson distribution, OK does not have a way to incorporate the different variances from each data point.

For this reason, different authors have proposed specific implementations of kriging methods that deal with data that is not normally distributed. Monestiez, Dubroca, Bonnin, Durbec, and Guinet (2006) presented a kriging methodology to model whale populations using data from observers on ferries and cargo ships, which can be modeled using a Poisson distribution. This approach is known as PK and has since been used to model phenomena as diverse as Cancer mortality (Goovaerts & Gebreab, 2008, Feb 04) and gamma-ray spectral mapping (Reinhart, 2013). For this reason, we have chosen this methodology for the current work.

PK provides an estimate $\hat{Z}(\mathbf{x}_0)$ for a variable Z at unknown location \mathbf{x}_0 while assuming a constant unknown mean over its neighborhood, although in this case the observations $Z(\mathbf{x}_i)$ are dependent on some underlying mean count rate and the amount of time spent at each location. The estimate is a weighted linear combination of the available observation $z_i = Z(\mathbf{x}_i)$ and the amount of time spent at each location t_i from a set of locations \mathbf{x}_i . The estimator is thus described as follows:

$$\hat{Z}(\mathbf{x}_0) = \sum_{i=1}^n z_i \frac{w_i}{t_i}, \quad i = 1, \dots, n, \quad (6)$$

where $\sum_{i=1}^n w_i = 1$ to ensure unbiased estimates. To correctly estimate the values at \mathbf{x}_0 the weights $\mathbf{w} = [w_1, \dots, w_n]^T$ must be calculated. This can be achieved by solving the PK system, which is a linear system of $n + 1$ equations.

$$\sum_{j=0}^n w_{ij} C_{ij} + w_i \frac{\hat{m}}{t_i} + \mu = C_{i\mathbf{x}_0} \quad \text{for } i = 1, \dots, n, \quad (7)$$

where C_{ij} is the covariance of the observed values, $C_{i\mathbf{x}_0}$ is the covariance at the prediction location \mathbf{x}_0 , and μ is a Lagrange factor which ensures the optimal solution. Finally, \hat{m} is estimated from the data as a weighted average of the count rates, where the weights correspond to the observation times.

Once this system is solved, the estimated values at a location x_0 can be found using Equation (6), and the associated variance of the prediction σ^2 can be calculated using the same equation as in OK

$$\sigma^2(x_0) = \sum_{i=1}^n W_i C_{ix_0}. \quad (8)$$

3.3.1 | Semivariogram

In empirical scenarios, it is possible to use a semivariogram created from the real-world data to express the relation between locations and estimate the weights for each observation. However, unlike OK, in this case, it is necessary to account also for the observation times for each data point. For this reason PK uses a weighted variogram estimator, which takes into account the different observation times

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{ij} \left(\frac{t_i t_j}{t_i + t_j} \left(\frac{z_i}{t_i} - \frac{z_j}{t_j} \right)^2 - \hat{m} \right) I_{d_{ij} \sim h}, \quad (9)$$

where h is the distance between points i and j , \hat{m} is the same mean as in Equation (7) and $I_{d_{ij} \sim h}$ is a gating function that takes a value of 1 when i and j are roughly distance h apart, and 0 otherwise. $N(h)$ is a normalizing factor calculated as follows:

$$N(h) = \sum_{ij} \frac{t_i t_j}{t_i + t_j} I_{d_{ij} \sim h}. \quad (10)$$

The semivariograms $\gamma(h)$ can take multiple forms but are generally characterized by an equation that can be parameterized. We use the following Gaussian semivariogram model in our work:

$$\gamma(h) = p_0 + (p_2 - p_0) \left(1 - \exp\left(-\frac{h^2}{p_1^2}\right) \right), \quad (11)$$

with the following three parameters: nugget p_0 , range p_1 and sill p_2 (Pulido Fentanes, Gould, Duckett, Pearson, & Cielniak, 2018b).

The parameters for this equation are automatically fitted from the semivariogram of the sampled data using the soft L^1 norm minimization scheme (Murphy, 2018).

3.4 | Exploration strategies

Our proposal is to use the variance of the kriging (KV) process (see Equation (8)) as a measurement of information gain. The use of KV as a reward function for robotic exploration has been previously studied in (Pulido Fentanes et al., 2018a, 2018b). In this work, we compare some well-known exploration strategies and how they interact with the sampling regime. The methods to be tested can be classified into next-best-view (NBV) and adaptive sampling methods. In addition, we also added a random strategy where the next sampling location is randomly chosen from a set of unexplored cells, which serves as a baseline for comparisons.

3.4.1 | Next-best-view

NBV methods update the environment model every time a new sample is acquired and then choose a new location depending on the distribution of the KV across the field. Location selection is done using one of the following strategies:

- *Greedy*: The next sampling point is the point with the highest KV in the set of candidate locations.
- *Monte Carlo*: A set of candidate sampling locations is generated each time, and each candidate location is allocated a weight depending on its KV. The next sampling location is selected randomly, but in a way that guarantees that the probabilities are distributed according to the weight of each candidate.

3.4.2 | Adaptive sampling

This strategy generates an initial plan that is modified depending on the reward function after each model update. In this case, the robot plans an initial sampling regime based on a random trajectory and a mission time horizon, which depends on the minimum expectations of measurements to be made in each case. Every new sample taken is used to update the model, which is used to remove sampling points with low KV, so the targets whose KV is below the overall KV mean of the model are removed, afterwards, as many new points as necessary to meet the minimum expectation of measurements in the remaining mission time are added by choosing new candidates using a Monte Carlo method. Finally, a new route is re-planned through the new set of points using a traveling salesman problem algorithm.

4 | EXPERIMENTAL FRAMEWORK

4.1 | Hardware setup

Our experimental setup consists of an autonomous mobile robot Thorvald (Grimstad & From, 2017) equipped with a custom-made, high-sensitivity soil moisture sensor based on fast neutron counting principle manufactured by Hydroinnova (see Figure 3). The 12 neutron detectors are accompanied by temperature and humidity sensors which are used for providing the corrected neutron counts every 10 s. The sides and top of the sensor are shielded by using a 50-mm polyethylene shield to limit the detection footprint of the sensor to 10 m. The total weight of the sensor is around 300 kg. The sensor is interfaced with the robot through an Ethernet link. The robot is controlled through an in-built PC running Linux OS and Robot Operating System. The platform is equipped with a GNSS sensor, which enables robot localization and geotagging of the collected data samples. The navigation component uses a graph-based representation, allowing the robot to move between a predetermined set of waypoints.



FIGURE 4 Location and layout of two data collection sites: An airfield (0.3 ha) at the Lincolnshire Aviation Heritage Center in East Kirkby, UK (left); a wheat stubble field (7 ha) near Volos, Greece (right) [Color figure can be viewed at wileyonlinelibrary.com]

4.2 | Datasets

Evaluating the performance of robotic exploration strategies is inherently difficult and previous work in that domain often relies on simulated experiments (e.g., Santos, Krajnik, Pulido Fentanes, & Duckett, 2016). In our case, we propose to use the “surrogate” models of soil moisture, based on data collected from two real fields with the described equipment. We used the collected data in off-line “simulations” to compare different exploration strategies and understand their overall performance. Simulations using a surrogate model are a useful tool to compare exploration methods (Pulido Fentanes et al., 2011, 2018a), providing the “ground truth” for the exploration results.

The two data collection sites include an airfield at the Lincolnshire Aviation Heritage Center in East Kirkby, UK and a wheat stubble field near Volos, Greece. Both fields were prepared in such a way so that they had equal parts of dry and wetland. Such an arrangement enabled us to systematically test the effectiveness of kriging-driven exploration strategies under a significant gradient between dry and wet areas akin to a step response.

The airfield site (see Figure 4) features a hard border between the grass field and concrete airstrip. Since concrete contains low levels of hydrogen, the airstrip provides a perfect replacement for dry conditions (5% VWC). The data collection took part in March 2018 and, therefore, the grass field was in a relatively wet condition (20% VWC). Thirteen measurement locations were selected along a parallel line to the wet/dry border at 1, 2, 4, 8, 15, and 30 m away from the border on both sides and a single point at the border itself. The measurement interval for all the points was set to 10 min.

The wheat stubble field in Greece (see Figure 4) covered a rectangular area of approximately 7 ha. The data collection took part in June 2018 under dry weather conditions. To create a wet area, the field was irrigated before data collection resulting in a wet/dry border with VWC of 18% for the dry part and 24% for the wet area, representing a fairly low gradient between the two parts. The whole field meshed into a grid of 72 sampling locations with a spatial resolution of 30×30 m. The measurement interval for all the points was set to 10 min.

Both datasets were used to create a set of testing models which were used to verify multiple hypotheses presented in Section 5. Each one of

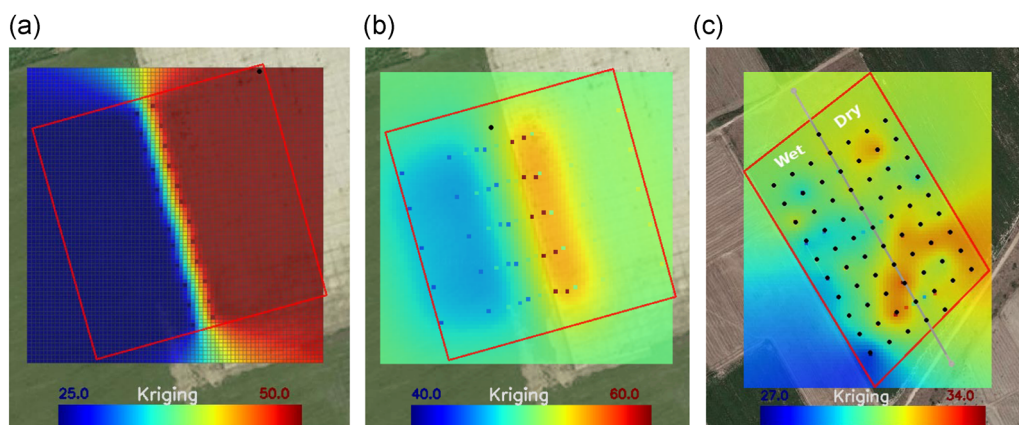


FIGURE 5 The high-gradient synthetic model generated from the airfield (a), the simulated model (b), and the validation model generated from the wheat stubble field (c) [Color figure can be viewed at wileyonlinelibrary.com]

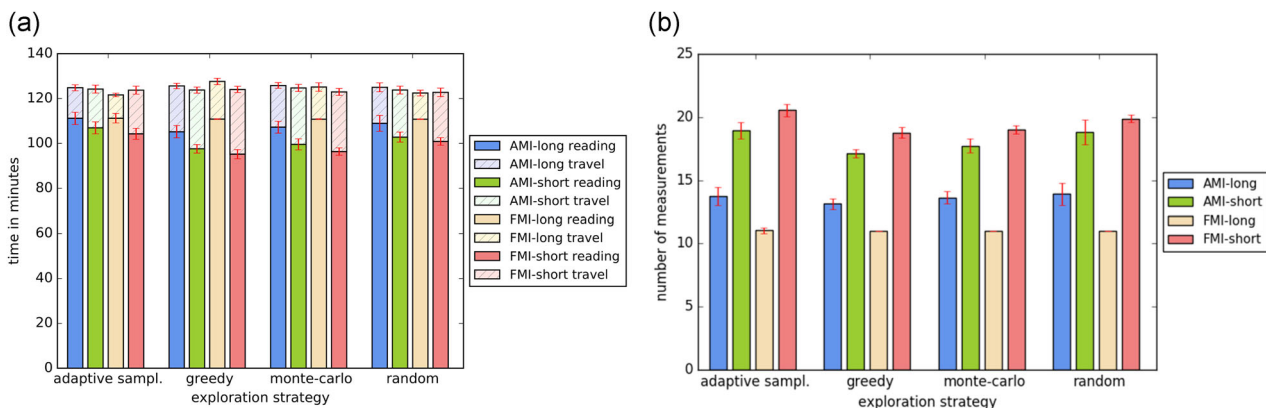


FIGURE 6 High-gradient synthetic scenario. Comparison of different sampling regimes and exploration strategies: (a) Measurement and travel times, (b) Number of samples taken. Average results over 30 runs with, error bars representing standard deviation for each case [Color figure can be viewed at wileyonlinelibrary.com]

these testing models has neutron rates as inputs for the measurement model which were then extrapolated across the testing area using OK. This way an estimated rate can be produced for every location on the field. This extrapolated rate is used as λ to produce simulated counts that follow the Poisson distribution (as seen in Figure 3) every 10 s (real sensor's update rate) at every cell in the environment, resulting in high-density models used as a reference. The models include:

- *Synthetic model* is based on real sensor rates recorded from the airfield (see Figure 8). To obtain these rates the sensor was left for 4 hr at the center of both wet and dry areas, using that data two synthetic models representing a high and low gradient between wet and dry soil were generated. The high-gradient synthetic (HGS) model was generated from the highest recorded rates in both readings, which were 2.5 and 5.0 counts/s for the wet and dry parts, respectively. The low-gradient synthetic (LGS) model

represents the average values for both readings which were 3.0 and 4.0 counts/s for the wet and dry halves, respectively.

- *Simulated model* is based on the real data recorded in the airfield and extrapolated into multiple lines covering a rectangular area (see Figure 9). To generate this model data from 10 min sensor readings at 13 different data points was used, all data points were captured in a straight line at the center of the field perpendicular to the division between both areas, six readings were made at 30, 15, 8, 4, 2, and 1 m from the center of the field into each half and one additional reading at the center of the field. These readings were copied into four additional parallel lines 10 m apart evenly spaced across the field.
- *Validation model* in which the real data from the wheat stubble field is used (see Figure 10). This model represents the most realistic soil moisture conditions and is used to validate the proposed algorithms.

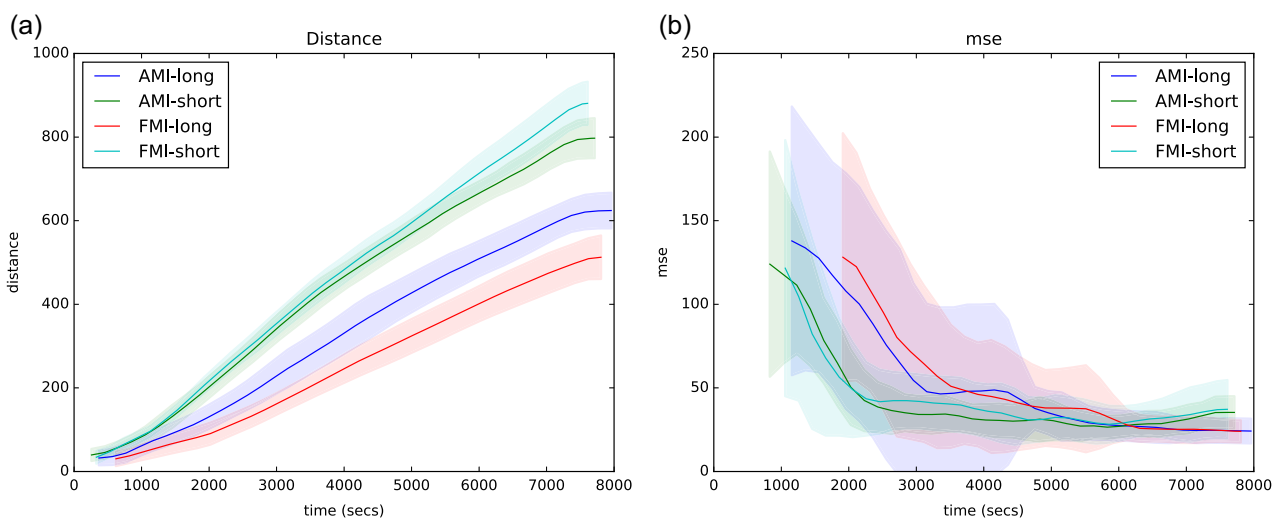


FIGURE 7 High-gradient synthetic scenario. Comparison of performance for methods using fixed versus adaptive measurement interval in terms of (a) travel distance and (b) mean square error. Average results over 30 runs, shaded areas represent standard deviation for each case [Color figure can be viewed at wileyonlinelibrary.com]

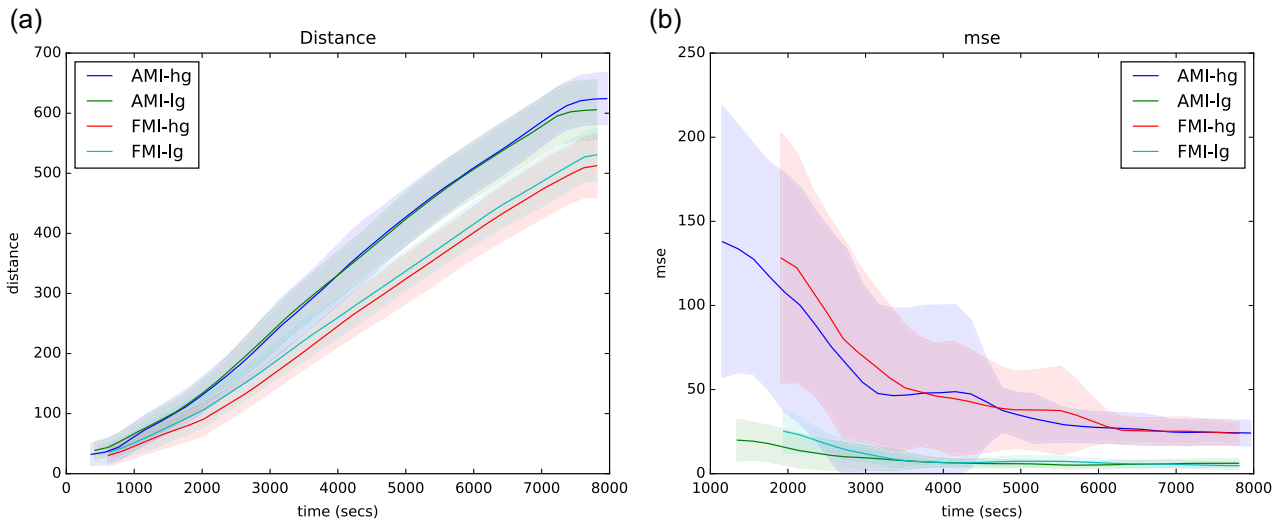


FIGURE 8 Comparison of performance of long fixed measurement interval and adaptive measurement interval in synthetic scenarios low-gradient synthetic and high-gradient synthetic in terms of (a) distance, and (b) mean square error. Average results over 30 runs, shaded areas represent standard deviation for each case [Color figure can be viewed at wileyonlinelibrary.com]

The cells size for both synthetic and simulated models based on the airfield dataset was 1 m^2 and for the validation model 25 m^2 . The cell size values were chosen considering the size of the environment and the computational cost; in both cases, the kriging model calculation and candidate location evaluation and the selection were kept under 10 s on an 8-core Intel i7-3770 CPU with 16GB of RAM running Ubuntu 16.04.

To indicate the overall variability of soil moisture in each model, we also present the estimated parameters for the Gaussian semivariogram model used for their generation:

- HGS model: $p_0 = 25.29$, $p_1 = 69.32$, and $p_2 = 316.36$;
- LGS model: $p_0 = 4.04$, $p_1 = 69.32$, and $p_2 = 50.61$;
- Simulated model: $p_0 = 20.56$, $p_1 = 21.71$, and $p_2 = 29.44$;
- Validation model: $p_0 = 0.67$, $p_1 = 144.14$, and $p_2 = 2.76$.

5 | EXPERIMENTS

To evaluate our framework, we have devised a set of experiments to test multiple hypotheses. First, the robot will focus on sampling the area with the highest uncertainty, that is, the border between the soil and concrete parts of the field and borders of the field. Second, we want to verify how much does the rate difference between the wet/dry parts of the field influence the exploration process (we call this a step response). Finally, we want to analyze the different impact of having an FMI and an AMI which warrants a minimum measurement uncertainty before moving on to the next sampling point. Because our sensor follows the Poisson distribution model, we believe that the robot will require less time to sample the dry area of the field as it would have observed a higher number of events in the same time reducing the measurement σ .

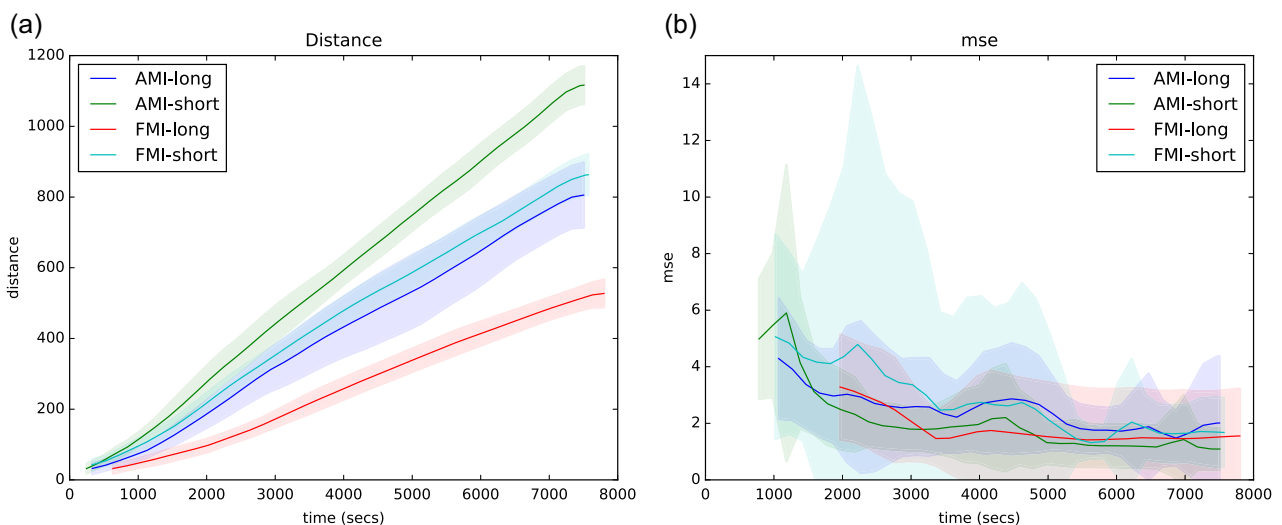


FIGURE 9 Simulated scenario. Comparison of performance for methods using fixed versus adaptive measurement interval in terms of (a) travel distance, and (b) mean square error. Average results over 30 runs, shaded areas represent standard deviation for each case [Color figure can be viewed at wileyonlinelibrary.com]

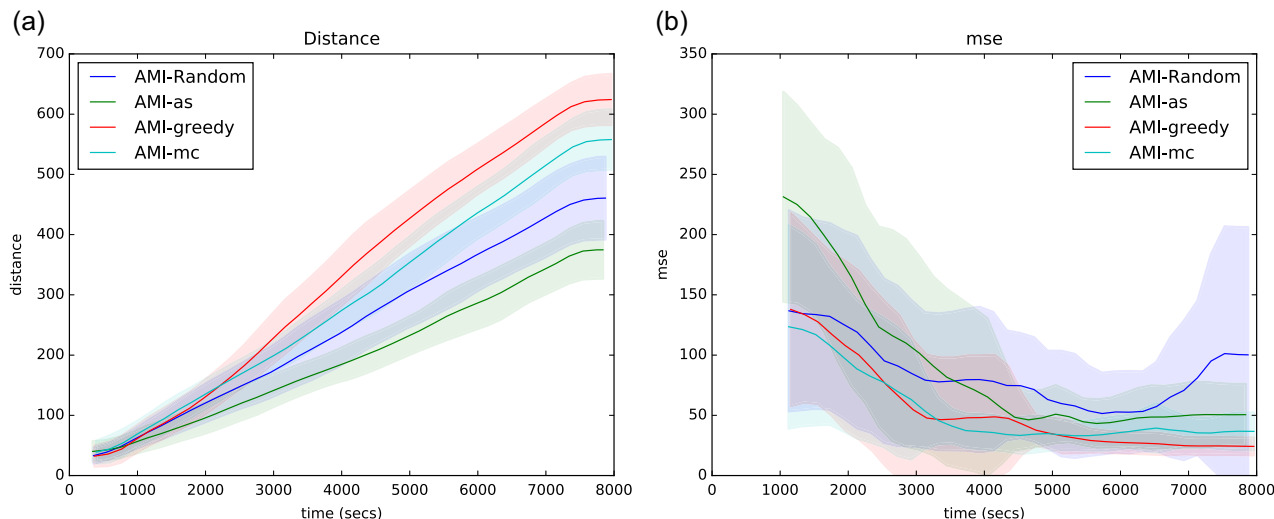


FIGURE 10 High-gradient synthetic scenario. Performance for different strategies using adaptive measurement intervals in terms of (a) distance, and (b) mean square error. Average results over 30 runs, shaded areas represent standard deviation for each case [Color figure can be viewed at wileyonlinelibrary.com]

The simulated robot was set to travel at 0.6 m/s, similar to the speed of the actual robot used for data collection.

The results presented in this section were obtained using 30 simulated runs over the testing models presented in Section 4.2. The performance of the exploration methods presented in this section is evaluated in terms of traveled distance and model error. For assessing the quality of the resulting model, we compare the model produced against the surrogate model used for the exploration. To compare any two resulting models A and B we use mean square error (MSE)

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (A_i - B_i)^2, \quad (12)$$

where A_i and B_i are the corresponding cells in the generated model B and the surrogate model A , and n is the total number of cells in both matrices. Variogram estimation requires a number of initial samples and hence the kriging results are not immediately available at the beginning of the exploration process. This is manifested with graphs representing the model quality starting at times different than 0 in all figures presented in the following sections.

5.1 | FMI versus AMI

To compare the influence of the sampling regime on the exploration process, all strategies were tested in the synthetic experimental setup following four different sampling regimes: two FMI and two AMI experiments. For the FMI case, one experiment was set to 10-min intervals (FMI-long) and the other 1–5 min intervals (FMI-short). For the AMI case, one experiment was set to a 2.5% measurement σ threshold (AMI-long) and the other one to a 3% threshold (AMI-short). Short and long cases should have comparable measurement times between them. The stopping criteria was the mission time which was set for the synthetic and simulated models to 2 hr and for the validation to 4 hr.

Figure 6 shows the impact that different strategies and sampling regimes have over the amount of time the robot spends gathering data and the total amount of samples it can gather within the specified constraints. The sampling regime plays a much bigger role in how the robot spends its time between reading and travel than the exploration strategy, hence the choice of sampling regime is critical to the performance.

Figure 7a shows that the total distance driven in the HGS scenario depends mainly on the measurement time. This was predictable given that the amount of time that the robot spends reading data is inversely proportional to the amount of time the robot spends navigating from one location to another. In Figure 7b it can be seen that AMI regimes lead to faster convergence than their FMI counterparts.

AMI strategies achieve better quality in shorter times because they can optimize the sampling time and drive exploration considering the conditions of the field (e.g., the robot will spend less time in drier places as it will observe a higher number of events and achieve higher levels of confidence for the readings). These gains are highly dependant on the variability of the soil moisture in the field, for example, in a predominantly wet field the gains from adaptive sampling interval strategies will be less noticeable. To verify this hypothesis, this analysis was also performed in a simulation with a lower gradient between the wet and dry parts (LGS model).

Figure 8 shows a comparison of both sampling regimes in synthetic scenarios with different gradients. The difference in performance between both regimes is relatively low in the scenario with the lower gradient (LGS). However, the traveled distance, for the adaptive strategy is slightly higher in both cases, indicating that sampling regimes are not important for controlling the traveled distance and that this is a factor that is mainly driven by the exploration strategy.

Figure 9 presents a comparison between both sampling regimes in the simulated model. Comparing these results to the ones obtained with the synthetic model (Section 5.1), it is possible to see that the results are

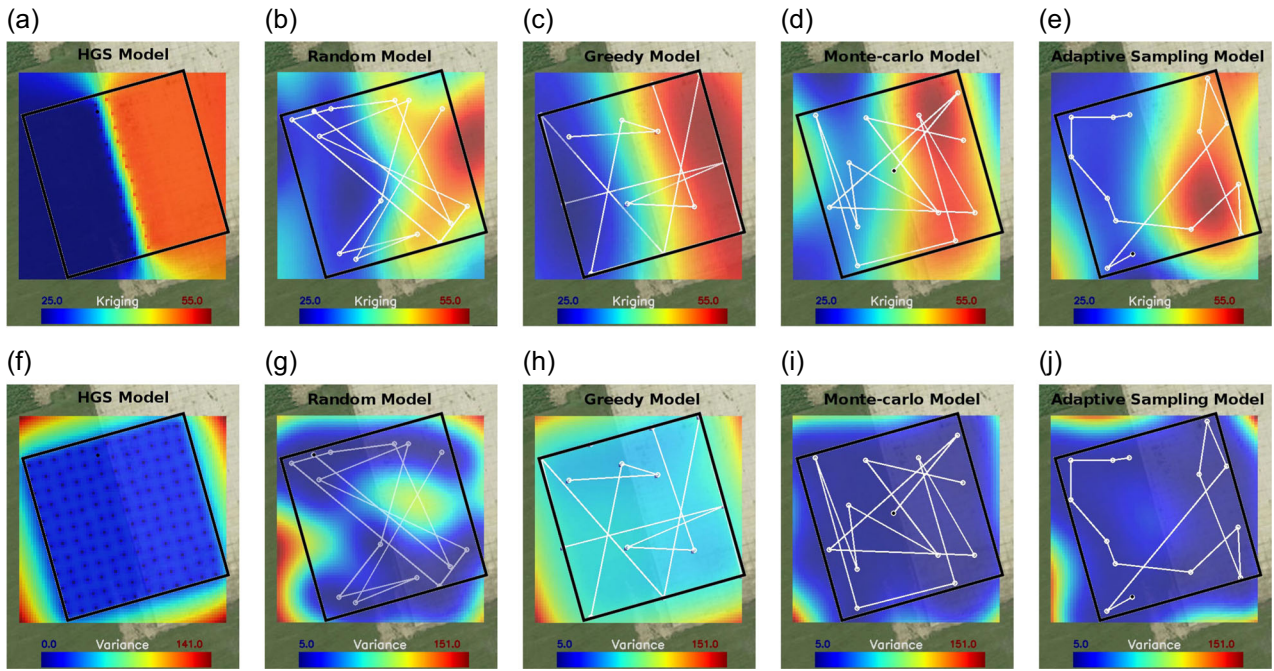


FIGURE 11 High-gradient synthetic model. Exploration outputs and robot trajectories. The kriging output (top row) and variance (bottom row) for the (a,e) full model, (b,f) Random, (c,g) Greedy, (d,h) Monte Carlo, and (e,j) Adaptive sampling strategies [Color figure can be viewed at wileyonlinelibrary.com]

almost identical in all cases. This indicates that, despite the fact that variability in the simulated model is just slightly higher than in the LGS scenario, the sampling regime has an influence over the variability of the results. In particular, FMI regimes are much more unstable than their adaptive counterparts indicating that it is generally preferable to use an AMI regime as it is more stable with medium gradients.

5.2 | Comparison of the exploration strategies

To verify the influence of different exploration strategies over the exploration process, we ran a series of simulations with four

different strategies namely: Random, Greedy, Monte Carlo, and Adaptive Sampling. In all cases, we used AMI as the measurement interval regime to isolate the effects of the exploration strategy only.

Figure 10 shows the performance of the different exploration strategies indicating their high influence on the distance traveled by the robot. In particular, it can be noticed that an adaptive sampling strategy achieves models that are slightly worse than those resulting from other strategies. This trade-off is a result of shorter travel distances and can indicate that this strategy might compare better in larger fields than those considered in this scenario, due to the fact

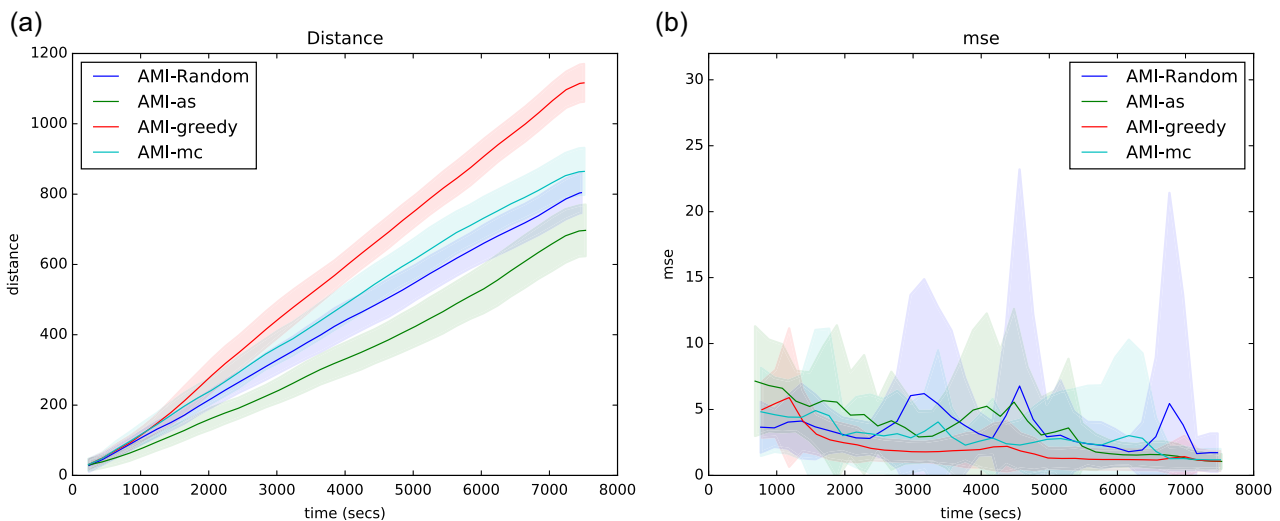


FIGURE 12 Simulated scenario. Performance for different strategies using adaptive measurement intervals in terms of (a) distance and (b) mean square error. Shaded areas represent standard deviation over 30 runs [Color figure can be viewed at wileyonlinelibrary.com]

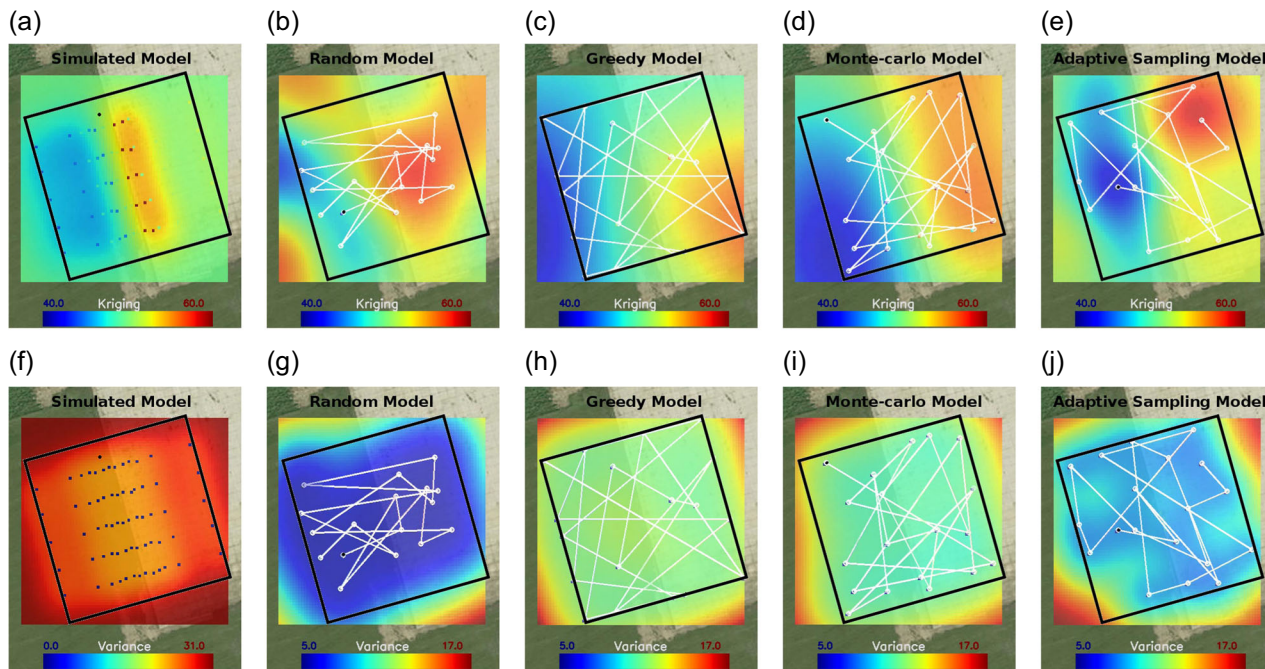


FIGURE 13 Simulated model: exploration outputs and robot trajectories. The kriging output (top row) and variance (bottom row) for the (a,f) full model, (b,g) Random, (c,h) Greedy, (d,i) Monte Carlo, and (e,j) Adaptive sampling strategies [Color figure can be viewed at wileyonlinelibrary.com]

that long travel distances can translate into a significant amount of time not spent on gathering data.

Figure 11 presents the outputs of the exploration result for the HGS model. The figure shows the resulting models for a field after 2 hr of autonomous exploration with the trajectories followed by

the robot. One interesting thing is that the greedy strategy drives the robot mostly to the edges of the field. This is mainly because the kriging methods are better at interpolation than extrapolation, so the highest variances are always around the limit areas. This has the advantage that it can drive the model's variance down very quickly.

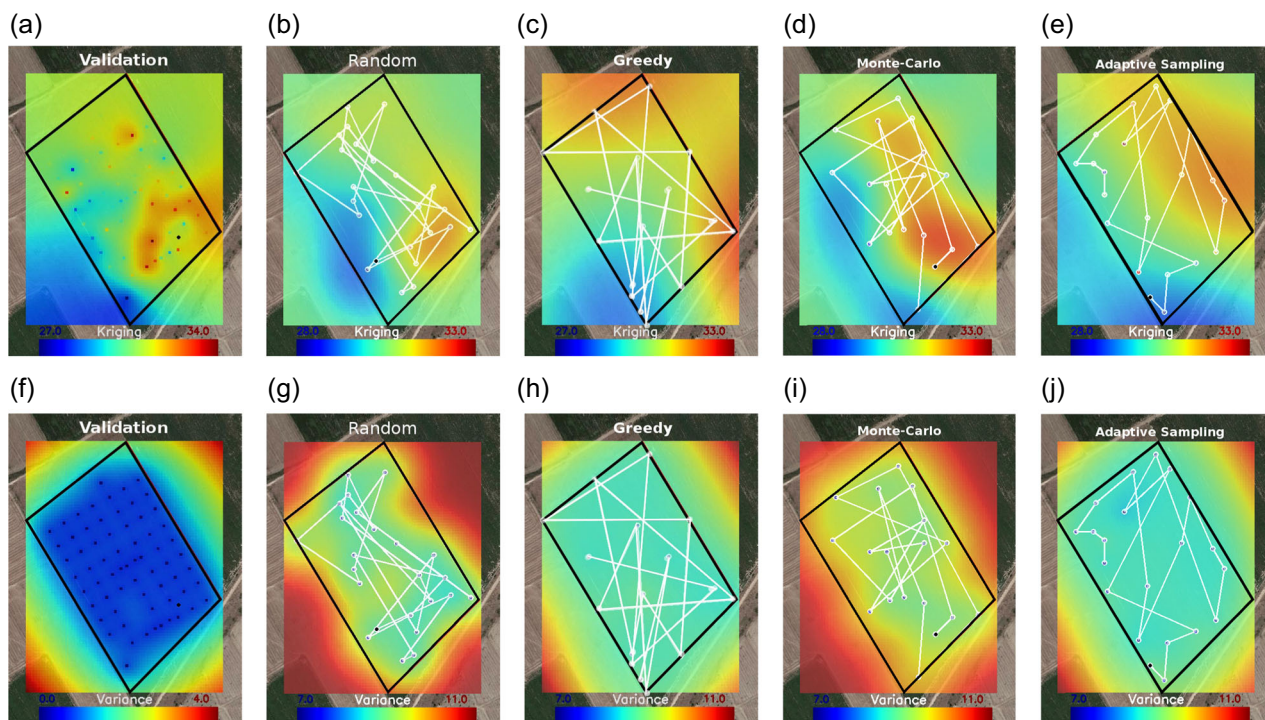


FIGURE 14 Validation model. Exploration outputs and robot trajectories. The kriging output (top row) and variance (bottom row) for the (a, f) full model, (b,g) Random, (c,h) Greedy, (d,i) Monte Carlo, and (e,j) Adaptive sampling strategies [Color figure can be viewed at wileyonlinelibrary.com]

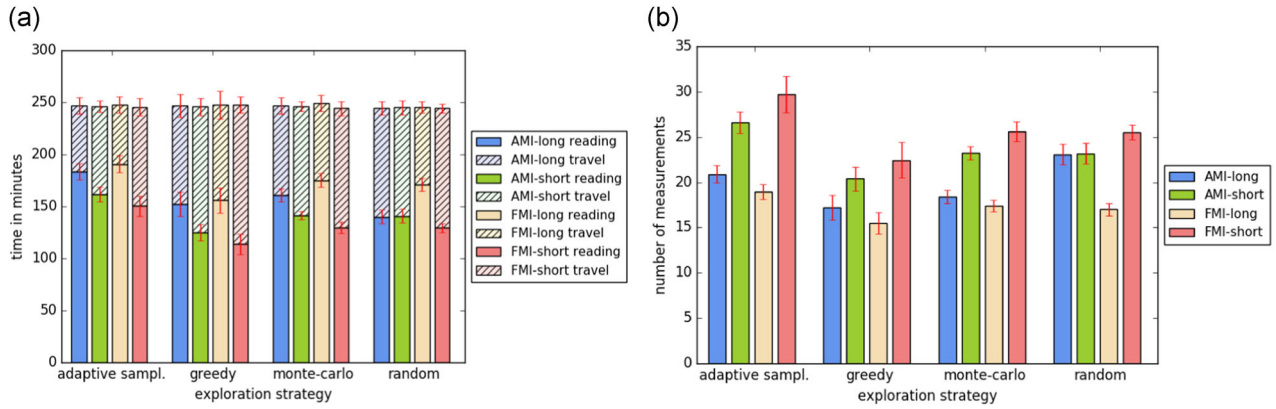


FIGURE 15 Validation scenario. Comparison of different sampling regimes and exploration strategies (a) measurement and travel times (b) number of samples taken. Average results over 30 runs with, error bars representing standard deviation over 30 runs [Color figure can be viewed at wileyonlinelibrary.com]

It might also mean, however, that it can miss relevant in-field information. In comparison, the adaptive sampling took samples that were evenly distributed across the field but visited them in a more organized way producing much smoother and shorter trajectories.

To verify these findings we performed the same test on the simulated airfield scenario. Figure 12 shows that the performance of the different strategies is similar to that exhibited in Figure 10. This indicates that the behavior of each strategy is consistent and does not tend to vary much across testing scenarios.

The outputs (see Figure 13) show again that greedy strategies follow very long paths and outer sampling points contrasting to the adaptive sampling method which follows a more balanced approach, that seems to linger around areas that are either drier or wetter than usual. The Monte Carlo approach shows an interesting behavior, it appears to be going back and forwards around the border between the grass and concrete, this seems to be because there higher

variances around the border area, however, the paths are very random and this increases the traveling distance. In that sense, the adaptive sampling strategy has a big advantage over Monte Carlo because it follows the same principle for choosing targets but at the same time, it reduces travel distance.

5.3 | Validation on the surrogate model

To validate the methodology, several experiments were executed simulating an exploration task of 4 hr. Figure 14 presents the resulting models for four experiments using different exploration strategies and AMI as the sampling regime.

It is possible to see by simple visual inspection that the resulting models do not reflect perfectly the reference validation model. We believe that this is mainly due to two factors: First, the gradient between wet and dry parts in this environment was very low affecting the resulting KV leading to less effective sampling. And

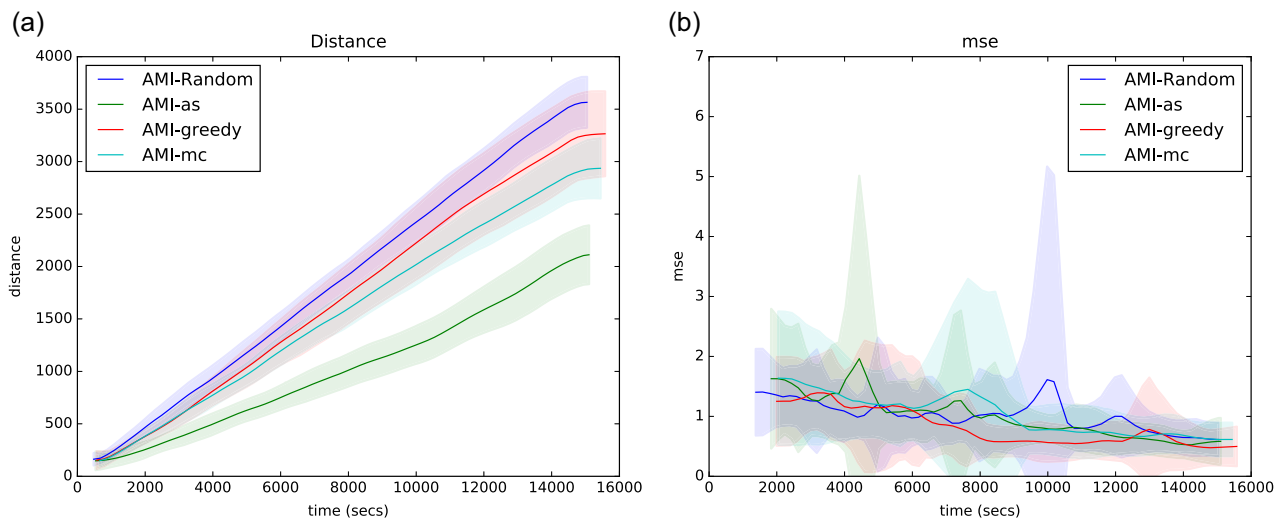


FIGURE 16 Validation scenario. Performance for different strategies using adaptive measurement intervals in terms of (a) distance and (b) mean square error. Shaded areas represent standard deviation over 30 runs [Color figure can be viewed at wileyonlinelibrary.com]

second, the size of the environment limits how many samples per hectare the robot can achieve. This means in practice that the maps had much lower resolution than the validation model, hence each sample represents a much broader area.

Figure 15 shows the impact the different strategies and sampling regimes have over the number of measurements the robot can take and the distribution of the time spent taking measurements against the time spent traveling between locations. Comparing these results to those in the smaller field (see Figure 7) it can be seen that the difference in the amount of samples captured using adaptive sampling with respect to the other strategies increases as well as the percentage of time spent capturing data.

The performance of the exploration strategies in terms of traveled distance and model error in the validation scenario is aligned with the results in the synthetic and simulated scenarios (see Figure 16). One special case is the adaptive sampling strategy that seems to have results that are more consistent in this scenario than in the HGS case. This could be a further indication that this method performance improves comparatively to other strategies as the field size grows. Overall, the validation model resembles the low-gradient synthetic LGS scenario where all methods converged quickly but with a high degree of variability (relatively high standard deviation of MSE) especially for strategies such as monte-carlo MC or random which are not information-driven.

It is worth noting that all the strategies-generated models whose wetter areas and dryer areas correspond to those of the validation model. Also, the soil moisture maps produced to provide a very good estimation of the areas where water deficit and concentration are in the field. Most likely, the miss-alignment between the validation model and the model outcome could have been overcome by having a longer mission. The fact that the resulting model can discriminate wet and dry areas in such a short time (the validation model required more than 60 hr of work) is very encouraging.

6 | CONCLUSIONS AND FUTURE WORK

In this paper, we proposed an exploration framework for autonomous mobile robots equipped with a soil moisture sensor to create high-quality soil moisture maps. The sensor is a novel device based on fast neutron counting which enables noncontact measurements of soil moisture. Such a class of sensors can be modeled by the Poisson distribution and we demonstrated how to integrate such measurements into the kriging framework. We also investigated a range of different exploration strategies and assessed their usefulness in different scenarios. The proposed framework was evaluated on a range of datasets based on real soil moisture data collected from two different fields.

One of the important findings of the paper is the fact that the sampling regime's contribution to the overall exploration process is highly dependant on the characteristics of the field. In fields with high variability and less uniform distribution of soil moisture, the use of AMI shows significant improvements in model quality compared with a fixed measurement time regime. We also demonstrated that adaptive sampling

strategies guarantee lower navigation times and allocate more time obtaining samples leading to more consistent and faster-converging models compared with the nonadaptive strategies. This might be especially important in large fields where traveling takes a significant proportion of the exploration time. Greedy methods tend to sample the outer border of the environments, which is where the KV is usually higher. They tend to miss localized patches, although their overall model quality is comparable. For small fields with uniform soil moisture distributions, these might be preferable exploration strategies.

Although the presented framework was demonstrated for the soil moisture mapping, it is a general approach which can be used to map other soil properties such as compaction, chemical composition, and so forth. It is a framework that would be particularly suitable in scenarios where the measured phenomena directly affect the acquisition time and need to be spatially mapped. This includes applications such as rainfall measurements, people and animal counting, gas detection and so forth. One of the follow-up questions arising from this research is if changing the time measurement regime on the fly could improve the resulting models even further. Future work could also address the additional path planning constraints caused by the layout of typical agricultural fields which feature soil beds and rows and utilizing information based stopping criteria such as the one proposed in (Ghaffari Jadidi et al., 2019) instead of standard mission times as in the current framework. Finally, the framework will be extended to map multiple soil properties at the same time.

It is worth noting that whilst this work addressed the study of Poisson KV as a reward function for exploration, other types of Gaussian processes such as heteroscedastic Gaussian process regression (e.g., Kersting et al., 2007) models both the spatial distribution of soil moisture and its noise dependant uncertainty and could also be applied to this case. For this reason, the study of different modeling methodologies performance for soil properties mapping and exploration is an interesting line for future work.

ACKNOWLEDGMENT

This work was supported by the Science and Technology Facilities Council (STFC) Newton Fund programme, project ID ST/N006836/1.

ORCID

Jaime P. Fentanes  <http://orcid.org/0000-0001-7505-6533>

Grzegorz Cielniak  <http://orcid.org/0000-0002-6299-8465>

REFERENCES

- Binney, J., Krause, A., & Sukhatme, G. S. (2013). Optimizing waypoints for monitoring spatiotemporal phenomena. *The International Journal of Robotics Research*, 32(8), 873–888.
- Diggle, P. J., Tawn, J. A., & Moyeed, R. (1998). Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47(3), 299–350.
- Dunbabin, M., & Marques, L. (2012). Robots for environmental monitoring: Significant advancements and applications. *IEEE Robotics & Automation Magazine*, 19(1), 24–39.

- Evans, J. G., Ward, H. C., Blake, J. R., Hewitt, E. J., Morrison, R., Fry, M., & Jenkins, A. (2016). Soil water content in southern England derived from a cosmic-ray soil moisture observing system-COSMOS-UK. *Hydrological Processes*, 30(26), 4987–4999. <https://doi.org/10.1002/hyp.10929>
- Gao, T., Emadi, H., Saha, H., Zhang, J., Lofquist, A., Singh, A., & Bhattacharya, S. (2018). A novel multirobot system for plant phenotyping. *Robotics*, 7(4), 61.
- Ghaffari Jadidi, M., Valls Miro, J., & Dissanayake, G. (2019). Sampling-based incremental information gathering with applications to robotic exploration and environmental monitoring. *The International Journal of Robotics Research*, 38(6), 658–685.
- Glaser, S., Schaefer, A., & Burgard, W. (2018). Mapping and localization using multispectral imaging of the soil. Paper presented in International Conference on Intelligent Robots and Systems (IROS), Unconventional Sensing and Processing for Robotic Visual Perception Workshop, Madrid, Spain.
- Goldberg, P. W., Williams, C. K., & Bishop, C. M. (1998). Regression with input-dependent noise: A Gaussian process treatment. In *Advances in neural information processing systems*, 493–499.
- Goovaerts, P., & Gebreab, S. (2008, Feb 04). How does poisson kriging compare to the popular bym model for mapping disease risks? *International Journal of Health Geographics*, 7(1), 6. <https://doi.org/10.1186/1476-072X-7-6>
- Grimstad, L., & From, P. J. (2017). The Thorvald II agricultural robotic system. *Robotics*, 6(4), 24. <https://doi.org/10.3390/robotics6040024>
- Hollinger, G. A., & Sukhatme, G. S. (2013). Sampling-based motion planning for robotic information gathering. In *Robotics: Science and systems*, 3.
- Jadidi, M. G., Miró, J. V., Valencia, R., & Andrade-Cetto, J. (2014). Exploration on continuous gaussian process frontier maps. In *2014 IEEE International Conference on Robotics and Automation*, pp. 6077–6082.
- Kerry, R., Oliver, M. A., & Frogbrook, Z. (2010). Sampling in precision agriculture. In Oliver, M. (Ed.), *Geostatistical applications for precision agriculture* (pp. 35–63). Dordrecht: Springer.
- Kersting, K., Plagemann, C., Pfaff, P., & Burgard, W. (2007). Most likely heteroscedastic Gaussian process regression. In *Proceedings of the 24th International Conference on Machine Learning*, pp. 393–400.
- Kim, Y.-H., & Shell, D. A. (2014). Distributed robotic sampling of non-homogeneous spatiotemporal fields via recursive geometric subdivision. In *Proceedings of the IEEE International Conference on Robotics & Automation*, pp. 557–562.
- Koenig, S., Tovey, C., & Halliburton, W. (2001). Greedy mapping of terrain. In *Proceedings of the IEEE International Conference on Robotics & Automation*, Vol. 4, pp. 3594–3599.
- Marchant, B., & Lark, R. (2007). Optimized sample schemes for geostatistical surveys. *Mathematical Geology*, 39(1), 113–134.
- Marchant, R., & Ramos, F. (2014). Bayesian optimisation for informative continuous path planning. In *Proceedings of the IEEE International Conference on Robotics & Automation*, pp. 6136–6143.
- Martinez-Cantin, R., de Freitas, N., Doucet, A., & Castellanos, J. A. (2007). Active policy learning for robot planning and exploration under uncertainty. *Robotics: Science and systems*, 3, 321–328.
- Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, 58(8), 1246–1266.
- Monestiez, P., Dubroca, L., Bonnin, E., Durbec, J.-P., & Guinet, C. (2006). Geostatistical modeling of spatial distribution of *Balaenoptera physalus* in the northwestern Mediterranean sea from sparse count data and heterogeneous observation efforts. *Ecological Modelling*, 193(3–4), 615–628.
- Murphy, B. (2014–2018). *Kriging toolkit for python*. <https://github.com/bsmurphy/PyKrig>
- Noborio, K. (2001). Measurement of soil water content and electrical conductivity by time domain reflectometry: A review. *Computers and Electronics in Agriculture*, 31(3), 213–237.
- O'Callaghan, S. T., & Ramos, F. T. (2011). Continuous occupancy mapping with integral kernels. In *25th AAAI Conference on Artificial Intelligence*.
- Oliver, M. A., & Webster, R. (1986). Combining nested and linear sampling for determining the scale and form of spatial variation of regionalized variables. *Geographical Analysis*, 18(3), 227–242.
- Popovic, M., Vidal-Calleja, T., Hitz, G., Sa, I., Siegwart, R., & Nieto, J. (2017). Multiresolution mapping and informative path planning for UAV-based terrain monitoring. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*.
- Pulido Fentanes, J., Gould, I., Duckett, T., Pearson, S., & Cielniak, G. (2018a). 3D soil compaction mapping through kriging-based exploration with a mobile robot. *IEEE Robotics and Automation Letters*, 3(4), 3066–3072. <https://doi.org/10.1109/LRA.2018.2849567>
- Pulido Fentanes, J., Gould, I., Duckett, T., Pearson, S., & Cielniak, G. (2018b). Soil compaction mapping through robot exploration: A study into kriging parameters. In *ICRA 2018 Workshop: Robotic Vision and Action in Agriculture*.
- Pulido Fentanes, J., Zalama, E., & Gomez-Garcia-Bermejo, J. (2011). Algorithm for efficient 3D reconstruction of outdoor environments using mobile robots. In *Proceedings of the IEEE International Conference on Robotics & Automation*, pp. 3275–3280.
- Reinhart, A. (2013). *An integrated system for gamma-ray spectral mapping and anomaly detection* (Unpublished doctoral dissertation). The University of Texas at Austin, Austin, TX.
- Rodias, E., Berruto, R., Busato, P., Bochtis, D., Srensen, C. G., & Zhou, K. (2017). Energy savings from optimised in-field route planning for agricultural machinery. *Sustainability*, 9(11), 1956. <https://doi.org/10.3390/su9111956>
- Sadat, S. A., Wawerla, J., & Vaughan, R. (2015). Fractal trajectories for online non-uniform aerial coverage. In *Proceedings of the IEEE International Conference on Robotics & Automation*.
- Santos, J., Krajcnik, T., Pulido Fentanes, J., & Duckett, T. (2016). A 3D simulation environment with real dynamics: a tool for benchmarking mobile robot performance in long-term deployments. In *ICRA 2016 Workshop: AI for Long-term Autonomy*.
- Schrön, M., Rosolem, R., Köhli, M., Piusi, L., Schröter, I., Iwema, J., & Zacharias, S. (2017). The Cosmic-Ray Neutron Rover - Mobile Surveys of Field Soil Moisture and the Influence of Roads. *ArXiv e-prints*.
- Vasudevan, S., Ramos, F., Nettleton, E., & Durrant-Whyte, H. (2009). Gaussian process modeling of large-scale terrain. *Journal of Field Robotics*, 26(10), 812–840.
- Williams, C. K., & Rasmussen, C. E. (2006). *Gaussian processes for machine learning*. Cambridge, MA: The MIT Press.
- Yamauchi, B. (1997). A frontier-based approach for autonomous exploration. In *Computational Intelligence in Robotics and Automation, 1997. cira'97., Proceedings., 1997 IEEE International Symposium on*, pp. 146–151.
- Yang, K., Keat Gan, S., & Sukkarieh, S. (2013). A gaussian process-based RRT planner for the exploration of an unknown and cluttered environment with a UAV. *Advanced Robotics*, 27(6), 431–443.
- Zreda, M., Desilets, D., Ferr, T. P. A., & Scott, R. L. (2008). Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons. *Geophysical Research Letters*, 35(21), <https://doi.org/10.1029/2008GL035655>

How to cite this article: Pulido Fentanes J, Badiie A, Duckett T, Evans J, Pearson S, Cielniak G. Kriging-based robotic exploration for soil moisture mapping using a cosmic-ray sensor. *J Field Robotics*. 2020;37:122–136. <https://doi.org/10.1002/rob.21914>