2012 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 7-12, 2012. Vilamoura, Algarve, Portugal

Bayesian Optimisation for Intelligent Environmental Monitoring

Roman Marchant and Fabio Ramos

Abstract— **Environmental Monitoring (EM) is typically performed using sensor networks that collect measurements in predefined static locations. The possibility of having one or more autonomous robots to perform this task increases versatility and reduces the number of necessary sensor nodes to cover the same area. However, several problems arise when making use of autonomous moving robots for EM. The main challenges are how to build an accurate spatial-temporal model while choosing locations for measuring the phenomenon. This paper addresses the problem by using Bayesian Optimisation for choosing sensing locations, and presents a new utility function that takes into account the distance travelled by a moving robot. The proposed methodology is tested in simulation and in a real environment. Compared to existing strategies, our approach exhibits slightly better accuracy in terms of RMSE error and considerably reduces the total distance travelled by the robot.**

I. Introduction

Environmental concerns have topped the agenda in the past decades. Problems such as water and air pollution, climate change and resource depletion are recognised by the scientific community as major challenges. Meanwhile machine learning and robotics research has seen significant developments to take on problems with large quantities of data, and allowing them to expand their potential to address real-world problems. These create great opportunities to tackle fundamental environmental issues [1].

A plausible long-term aspiration is having a group of robots capable of monitoring the environment and executing actions that maintain it suitable for humans. However, this goal is far from being completed and the first challenge to be solved is *Intelligent Environmental Monitoring* (IEM), i.e. robots operating autonomously and deciding where to gather samples from a natural phenomenon to best model it. For instance, an environmental monitoring challenge would be to supervise the quality of the water in a lake used as water reservoir for a big city. This requires building a model of the presence of pollutants over the whole lake based on previously sampled areas. Measurements can be the concentration of chemicals or other related variables such as temperature or ambient light. Other examples are monitoring air pollution in cities, tracking ozone concentration, studying vegetation growth in problematic areas, among many others. The technique presented in this paper can be applied to any of these situations. Properly solving these problems requires creating a reliable spatial-temporal model of a phenomenon and, highly important, find the path that maximises the robot's understanding of the environment. Both areas have received lots of attention over the years.

Roman Marchant and Fabio Ramos are with the School of Information Technologies, University of Sydney, NSW 2006 Australia. {r.marchant,f.ramos}@acfr.usyd.edu.au

This paper presents a novel approach for active sampling, using *Bayesian Optimisation* (BO) [2–4]. The BO framework allows a mobile robot to choose sensing locations taking into account the uncertainty of the understanding of the phenomenon, the expected value of the studied variable (e.g. temperature, humidity, ambient light), and the cost to travel to specific locations. This is achieved by using a *Gaussian Process* (GP) [5] model for the phenomenon and choosing an appropriate acquisition function that automatically deals with the exploration-exploitation tradeoff. The main contributions of this paper are the extension of the BO framework to environmental monitoring with mobile robots and a new acquisition function for the BO algorithm, that considers the distance between sampling locations. This last innovation is beneficial for robotic systems as it reflects the cost of moving in the environment, mainly determined by the power consumption.

The remainder of this paper is structured as follows. Section II reviews the existing literature related to environmental monitoring and spatial-temporal models. Section III presents the theory behind GP regression for environmental monitoring, while section IV details the proposed intelligent sampling methodology using BO. Experimental setup, results and analysis are shown in section V. Finally, section VI draws some conclusions and suggests directions for future work.

II. Related Work

Environmental monitoring and informative path planning have been studied by many researchers. Existing research has focused mainly in two areas: modelling spatial-temporal phenomena and choosing the most informative sampling locations.

A. Spatial-Temporal Modelling

The idea of predicting the value of a process in space and time, based on a set of samples, is not new and has been widely studied. GP regression [5] has become popular in the machine learning community due to its known good results with spatially correlated data. For example, GP regression has been used successfully for modelling the signal strength of a wireless signal [6], or gas concentration in indoor and outdoor environments [7].

Singh et al. [8] use GPs for spatial regression and address the problem of defining appropriate covariance functions to model spatial and temporal dependences for environmental monitoring. They evaluate the performance of six different classes of covariance functions combining stationarity, separability and temporal consideration. The main conclusion

978-1-4673-1736-8/12/S31.00 ©2012 IEEE

of their study was that separable, non-stationary covariance functions give better results when modelling spatial temporal dynamic variables.

Other strategies for learning spatial models have been studied, such as Gaussian Process Mixtures [7, 9], Space-Time Process Convolutions [10, 11] or Kernel DM+V/W algorithm [12]. As this paper focuses on an efficient sampling algorithm, we use a standard GP modelling approach where time is considered as an additional dimension in the input space.

B. Informative Sensing Locations

Initial solutions for informative sampling of environmental phenomena were based on a fixed sensing network [13]. This approach evolved rapidly into the use of sensing robots, especially in large open environments, where one or more mobile robots can access larger areas using less sensing equipment.

The most popular approach to choose sampling locations is based on Information Theory principles [14]. Sampling in higher entropy locations [13, 15] or sampling in locations with higher information gain [8, 16–18] are the known strategies for choosing key points as sampling locations. Krause and Guestrin [16] use a greedy algorithm for walk selection maximising sub-modular functions to generate a sequential path plan for each moving robot.

These approaches chose the path of one or more robots based only on the existing uncertainty of the expected value over the whole domain. However, in most environmental monitoring applications, the places of interest for sampling are associated with extreme values of the sampled variable. For example, areas of high ozone concentration at lower altitude, or areas of high pollutant concentration. Therefore, the use of the BO framework fits logically because it prioritises higher or lower values of the prediction in the unsampled space. The theory behind this approach is explained in section IV.

III. Environment Model using Gaussian Process Regression

A mathematical model of a phenomenon is very important for making correct decisions using a quantitative analysis. Scientists have established physical and chemical laws that describe how environment phenomena behave and how relevant variables relate to each other. Although this deterministic approach may be valid for controlled situations, it is not useful for most environmental monitoring applications. For example, the spatial-temporal distribution of air-pollutant concentration in a city depends on traffic, temperature, humidity, wind speed, air pressure and building configuration among many others. There is no exact model that can capture the cross dependencies between all of these variables and achieve an accurate prediction of what the concentration will be in any given location.

GPs are an elegant solution to perform regression in stochastic processes, as they are a powerful, non-parametric tool for non-linear regression. This section reviews GP theory for regression. For deeper insights and a more extensive theoretical description, the reader can refer to [5].

Essentially, a GP places a multivariate Gaussian distribution over the space of functions mapping the input to the output. It is completely defined by a mean function $m(\mathbf{x})$ and a covariance function $k(\mathbf{x}, \mathbf{x}')$ i.e. $f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$.

Using supervised learning to build a GP model involves gathering a set of observations $S = \{\mathbf{x}_i, y_i\}_{i=1}^{N}$, where $\mathbf{x}_i \in \mathbb{R}^D$ are the N sampling locations (inputs) in a D dimensional space and $y_i \in \mathbb{R}$ are the corresponding noisy outputs. The learnt GP model can be used to predict a Gaussian distribution over $f(\mathbf{x}^*)$ at any new sampling location $\mathbf{x}^*$.

Samples from the real process are assumed to be noisy, i.e. $y = f(\mathbf{x}) + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$. The joint distribution of the observed target values $\mathbf{y} = \{y_i\}_{i=1}^{N}$ and the test locations $\mathbf{y}^* = \{y_i^*\}_{i=1}^{M}$ is given by

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}^* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K_X & K(X, X^*) \\ K(X^*, X) & K(X^*, X^*) \end{bmatrix}\right), \quad (1)$$

where $K(X', X'')$ is the covariance matrix, defined in a component-wise fashion as $K(X', X'')_{(i,j)} = k(x_i', x_j'')$, X, X^* are the groups of training and testing locations respectively, and $K_X = K(X, X) + \sigma_n^2 I$. Given the training locations, the predictive distribution over the test locations is given by

$$\begin{aligned} \mathbf{y}^* | X, \mathbf{y}, X^* &\sim \mathcal{N}(\boldsymbol{\mu}^*, \Sigma^*)\,, \quad (2) \\ \boldsymbol{\mu}^* &= K(X^*, X) K_X^{-1} \mathbf{y}\,, \\ \Sigma^* &= K(X^*, X^*) - K(X^*, X) K_X^{-1} K(X, X^*)\,. \end{aligned}$$

In practice, each test location $\mathbf{x}^* \in X^*$ is evaluated separately. Therefore, for the case of a unidimensional output, each testing point generates a unidimensional mean and variance values as an output

$$\begin{aligned} \mu(\mathbf{x}^*) &= K(\mathbf{x}^*, X) K_X^{-1} \mathbf{y}\,, \\ \sigma(\mathbf{x}^*) &= K(\mathbf{x}^*, \mathbf{x}^*) - K(\mathbf{x}^*, X) K_X^{-1} K(X, \mathbf{x}^*)\,. \end{aligned} \quad (3)$$

Choosing an appropriate covariance function is important, as it influences the GP's output directly. Several covariance functions have been studied throughout the literature, such as the Neural Network, Squared Exponential and Matern covariance functions. In this work, we use the Squared Exponential covariance function with automatic relevance determination,

$$k(\mathbf{x}', \mathbf{x}'') = \sigma_f^2 \exp\left(-\frac{1}{2}(\mathbf{x}' - \mathbf{x}'')^T M (\mathbf{x}' - \mathbf{x}'')\right), \quad (4)$$

where M is a square matrix of size D, with each element defined as:

$$M_{i,j} = \begin{cases} \ell_i & \text{if } i = j \\ 0 & \sim \end{cases} \quad (5)$$

The parameter σ_f is called signal variance and the vector $\boldsymbol{\ell} = \{\ell_i\}$ contains a length scale parameter for each dimension.

Although, this covariance function may be considered as simple, the proposed methodology is capable of using a more complex, non-stationary and/or non separable covariance functions, as described in [8].

The optimal set of parameters $\boldsymbol{\theta}^* = \{\sigma_f, \boldsymbol{\ell}, \sigma_n\}$ is found by maximising the log marginal likelihood (LML) goal function,

$$\boldsymbol{\theta}^* = \max_{\boldsymbol{\theta}} \ \mathrm{LML}(\mathbf{y}, \mathbf{X}, \boldsymbol{\theta}) \,, \tag{6}$$

with,

$$\mathrm{LML}(\mathbf{y}, \mathbf{X}, \boldsymbol{\theta}) = -\frac{1}{2}\mathbf{y}^T K_X^{-1}\mathbf{y} - \frac{1}{2}\log|K_X| - \frac{n}{2}\log 2\pi \,. \tag{7}$$

By finding the optimal hyper-parameters, the structure of the covariance function is adapted to achieve an optimal representation of the sampled data.

IV. Intelligent Environment Monitoring

Using a GP model by itself is not enough for creating a system for intelligent environmental monitoring. As it was stated earlier, it is necessary to choose sensing locations wisely to build a high quality model from the phenomenon in each time step.

The goal of environmental monitoring is to detect abnormal or extreme values of interesting variables. Therefore, it can be addressed as an optimisation problem. However, not every optimisation strategy is useful, as the goal is not only trying to find the extreme values (maximum and minimum), but also understand what happens across the studied area. This is a non trivial problem considering that we are dealing with an unknown and complex function modelling a time dependent process.

An increasingly popular tool in machine learning that helps solving all of the above stated problems is BO. This section gives a short description of BO, as well as the benefits and challenges of applying it to IEM. A complete and extensive theoretical treatment of the general BO algorithm can be found in Brochu et al. [2].

A. Bayesian Optimisation

The goal of BO is to find the maximum of an unknown function g. This is equivalent to finding $\mathbf{x}^\star = \arg\max_{\mathbf{x}} g(\mathbf{x})$. If a particular problem requires finding the minimum, BO can maximise the negative of the original function.

The only assumption over g is that it is Lipschitz-continuous, i.e.

$$\exists \ C \in \mathbb{R} \ \mid \ \|g(\mathbf{x}_1) - g(\mathbf{x}_2)\| \leq C\|\mathbf{x}_1 - \mathbf{x}_2\| \ \ \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}.$$

BO is an iterative algorithm that makes use of Bayes' theorem to combine prior knowledge with evidence to produce a new estimation of the statistical model of g. The prior is a GP model of the phenomena, whose mean captures the estimated value of the function, and whose variance determines the level of uncertainty of the prediction in that particular location, as described by 3. Evidence are the noisy samples from g. A clever use of this information may result in a quantitative estimation of the most probable location of the optimum.

Algorithm 1 Bayesian Optimisation

1: Let $\mathbf{x}_t$ be the chosen sampling point at iteration t.
2: Let s be an acquisition function.
3: **for** $t = 1, 2, 3, \ldots$ **do**
4: Evaluate $s(\mathbf{x})$ over the sampling domain.
5: Find $\mathbf{x}_t = \arg\max_{\mathbf{x}} s(\mathbf{x})$
6: Sample of the unknown process.
7: Update the GP model with the new sample.
8: **end for**

The BO algorithm pseudo code is shown in Algorithm 1. In each iteration the BO algorithm must maximise the *expected utility*, i.e. a function that quantifies the benefit of choosing a specific location to be sampled. This function is called *acquisition function*, s, and its general form will be discussed in section IV-B. In brief, BO uses a quantitative measure given by the acquisition function for taking informed decisions and choose the most convenient locations to sample the unknown function over its domain.

It can be noted that the problem of maximising g has now been moved to finding the maximum of s in each iteration, another non-convex optimisation. However, considering an appropriate acquisition function, optimising in the domain of the acquisition function is easier than the original problem.

There are several characteristics that make BO a powerful algorithm for efficiently finding the maximum of a function, particularly in IEM:

1) It considers a completely unknown function, where no information about the explicit expression of the function or its derivatives is present. Extremely helpful for environmental applications where no closed form is present for the phenomena.
2) It is aims to keep the number of samples to a minimum, acknowledging a costly-to-evaluate function. Specifically useful for the application of IEM using mobile robots, were saving energy is imperative.
3) Lastly, it makes use of a GP model to handle the prior and posterior probability density functions over the function space, using the uncertainty and expected value to choose sampling locations.

B. Acquisition Function for IEM

The key for success of the BO approach for optimisation is the acquisition function. The combination of an appropriate acquisition function and a GP allows an elegant trade off between exploration and exploitation while searching for the optimum.

An often used approach is to sample in areas where the robot does not have a good understanding of the phenomenon by maximising the *Information Gain* (IG). The IG is defined as the difference in entropy when adding a new point to the training set,

$$\mathrm{IG}(\mathbf{x}) \triangleq H[X] - H[X \cup \mathbf{x}], \tag{8}$$

where $H[X]$ is the entropy over the entire domain, using X as training set. IG is a monotonic function that has its maximum located where the variance is highest. The IG approach would then correspond to an acquisition function that is equivalent to the variance of the prediction over the domain. Although this is a popular approach, we believe that in most applications of environmental monitoring it is also important to sample in areas where the studied phenomenon may reach extreme values. This is because relevant information is associated with data away from the norm, as it may potentially cause problems to the environment.

The BO literature proposes several functions that make use of the mean and variance, such as the probability of improvement, the expected improvement and the *Upper Confidence Bound* (UCB). We choose the UCB acquisition function because it will not completely discard previously sampled locations. The UCB acquisition function evaluated at location $\mathbf{x}$ with uni-dimensional mean value for the prediction $\mu(\mathbf{x})$ and variance $\sigma(\mathbf{x})$ takes the form

$$UCB(\mathbf{x}) \triangleq \mu(\mathbf{x}) + \kappa \cdot \sigma(\mathbf{x}) \,. \tag{9}$$

The parameter κ is related to the exploration-exploitation trade off. While high values of κ lead to an explorative behaviour of the algorithm, lower values of κ favour exploitation near known sampled locations.

The approach studied in this paper uses BO to choose sampling locations, by using the analogy that the costly-to-evaluate function is the dynamic phenomenon present in the environment. From the robotics point of view, it is useful to get as much information as possible from the environment while sampling in a small number of locations. If these locations are placed and ordered smartly, the robot will save energy while constructing a reliable spatial-temporal model of the phenomenon.

All of the above is achieved by using an acquisition function that takes into account the mean of the prediction, the variance and the distance to the last sampled location in the domain. We call this acquisition function *Distance-based Upper Confidence Bound* (DUCB), due to its similarities with UCB. Its expression is given by

$$DUCB(\mathbf{x}|\mathbf{x}^-) \triangleq \mu(\mathbf{x}) + \kappa \cdot \sigma(\mathbf{x}) + \gamma \cdot d(\mathbf{x}, \mathbf{x}^-), \tag{10}$$

where the term $d(\mathbf{x}, \mathbf{x}^-)$ is the euclidian distance between the last sampled location $\mathbf{x}^-$ and the candidate location $\mathbf{x}$.

In the existing literature the acquisition function is evaluated over the whole domain of each dimension of $\mathbf{x}$. However, for the IEM application, the acquisition function is evaluated over spatial dimensions of $\mathbf{x}$, disregarding the time dimension and setting it to a fixed value. This means that for every iteration, the next sampling location is chosen over the spatial domain for that particular instant of time.

A one dimensional example of the BO algorithm using DUCB acquisition function, with $\kappa = 28.6$ and $\gamma = 0.72$, is illustrated in Figure 1. It is possible to see how the mean of the GP regression model quickly converges to the unknown function and the variance is reduced near sampling locations.

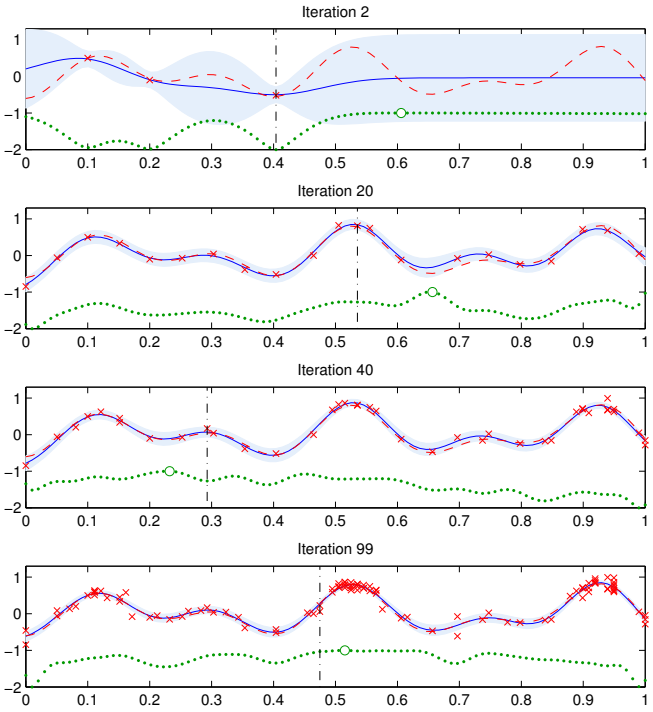


Fig. 1. One dimension example of active sampling based on BO. The continuous line and shade represent the GP mean an variance respectively. The dashed line is the unknown function, and noisy samples from this function are shown as crosses. The dotted line is the DUCB acquisition function, with a circular mark on its maximum. This function is scaled and with an offset for visualisation purposes. The last sampling locations is shown with a vertical dash-dot line.

The behaviour of the acquisition function changes through time[1]. Initially, when the variance is high, the sampled locations are automatically chosen by the algorithm to get an initial approximation of the unknown function. In the long term, samples concentrate near higher values of the unknown function. The effect of the distance-penalty in DUCB is clear, as the next sampled location (circle) is chosen to be close to the last sampled location (vertical line).

V. Experiments

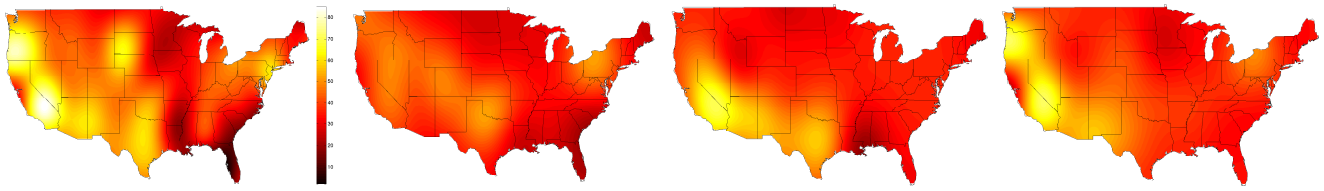
The proposed active sampling technique was evaluated in two experiments. The first one is a simulated scenario using a real dataset obtained across the entire US territory. The second is a real-world experiment using an autonomous mobile robot to model ambient conditions indoors.

A. Simulated Experiment

This experiment uses part of a real-world environment dataset, made available by the United States Environmental Protection Agency[2]. This is a considerably large dataset, covering the U.S territory with hourly samples dating back

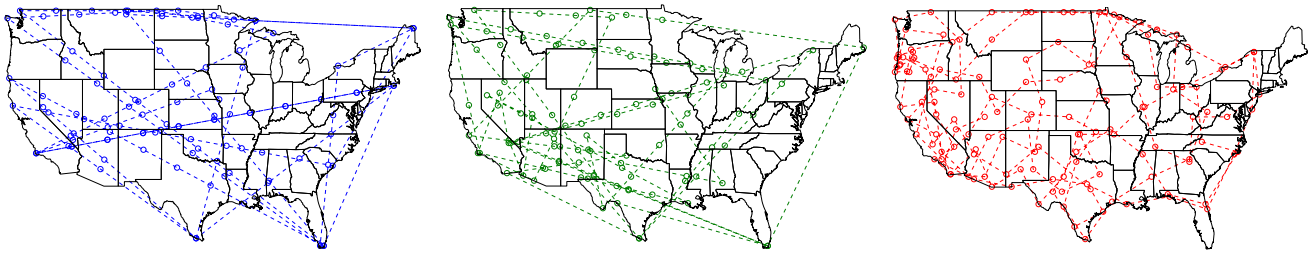
[1]A video showing all iterations is available at `http://www.it.usyd.edu.au/~rmar5258/IROS2012/1d\_example.html`

[2]Dataset web access: `http://java.epa.gov/castnet/reportPage.do`



(a) Ground Truth (b) Information Gain (c) BO with UCB (d) BO with DUCB

Fig. 2. Estimated mean of a GP model of ozone concentration [ppb]. Color-scale showed in (a) and instant of prediction ($t = 0.62[day]$) is the same for all figures.



(a) Information Gain (b) BO with UCB (c) BO with DUCB

Fig. 3. Trajectories followed by robots, samples are only allowed inside the US territory.

to 1987, including meteorological variables such as temperature, humidity, solar radiation and ozone concentration among many others.

The environmental phenomenon to be monitored is the ozone concentration in $[ppb]$ with raw data provided by $N = 103$ static monitoring stations across the US territory for the 1st of August 2009. A spatial-temporal GP was trained with the data, using the Squared Exponential with automatic relevance determination covariance function (4), whose optimal hyper-parameters were found by maximising the log marginal likelihood ($\{\sigma_f, \ell_{lat}, \ell_{long}, \ell_t, \sigma_n\} = \{14.02, 4.1, 4.1, 0.1336, 2.09\}$). A special restriction of having the same value for both spatial length scales is enforced during the optimisation process, because of the known isotropic behaviour of gas concentration in space. The mean of this GP (μ_{gt}), shown in Figure 2a, is used as ground truth for the entire continuous domain in space and time.

The sampling performed by a mobile robot is simulated by noisy measurements from the ground-truth GP mean, with $\sigma_n = 2$. Three different sampling algorithms are ran over the entire temporal domain, i.e. 24 hours. The first algorithm choses sensing locations based on IG criterium, the second one picks sensing locations using the BO algorithm with UCB as acquisition function (with $\kappa = 8$), and the third uses the BO algorithm with our proposed DUCB as acquisition function (with $\kappa = 8$ and $\gamma = 7$).

The same starting point $\mathbf{x}_0 = (\text{Long}, \text{Lat}) = (-95, 40)$ was selected for all three algorithms. The simulation process considers a hypothetical UAV that travels at a speed of $600[km/h]$. Given the short length scale in the time dimension, the high speed helps dealing with the large distances that need to be covered by the robot in a small amount of time. This may seem quite extraordinary, however, the results are equally valid for lower scale problems if the size of the domain and speed are reduced proportionally.

An iterative process is executed for the 24 hours of simulated data[3]. In each iteration, a robot moves towards the last selected goal. Each robot uses its own criteria for selecting a new goal after reaching the previous one. The concentration of ozone gas particles is sampled every $10[min]$ while the robot moves towards the goal. This ensures that robots take a similar number of samples over the experiment, the sole difference being the places where measurements are taken, allowing a proper comparison of the three different algorithms. The estimated GP model of the phenomenon is recalculated after a new sample is acquired. Figure 2 shows the estimated mean of the GP model at an arbitrary time step ($t = 0.62[day]$) of the simulation for the three sampling strategies. The trajectories followed by each of the robots during the 24 hours sampling window are shown in Figure 3.

A visual inspection over the IG strategy (Figure 2b) allows stating that it fits the ground truth data poorly. The IG strategy shows similar performance for the overall space, not showing any preference for higher or lower ozone concentrations. The BO approach with UCB (Figure 2c) and the strategy using BO with DUCB (Figure 2d) present a better fit to the ground truth data. Both of them model correctly the highest spike in ozone concentration on the west of the U.S. However, using the DUCB seems to capture in a better way the two high concentration areas on the west of the United States of America.

The trajectories followed by each robot during the 24 hours sampling window are shown in Figure 3. It can be seen that the IG sampling strategy picks locations far away

[3]A video showing the ground truth, mean estimation, the acquisition function and the trajectory for each approach can be found at `http://www.it.usyd.edu.au/~rmar5258/IROS2012/results.html`

TABLE I

RESULTS FOR SIMULATED EXPERIMENT, MEAN AND STD ARE IN [PPM]

Indicator	Method	Mean	Std	Distance $10^3[km]$
RMSE	IG	13.78	2.31	717.38
RMSE	UCB	13.04	2.26	718.25
RMSE	DUCB	12.33	3.47	639.06
WRMSE	IG	8.21	1.78	717.38
WRMSE	UCB	7.10	1.65	718.25
WRMSE	DUCB	6.51	1.21	639.06

from each other and usually the extremes of the territory. A similar behaviour is observed for the BO with UCB strategy, although the sampling locations tend to concentrate on the south west of the U.S. territory. This is because the mean of the ozone concentration is higher in that area over the whole day. Finally the approach using BO with DUCB presents a more elegant solution, where each sampling point is near to the last one. This strategy distributes measurements over the whole domain, and at the same time presents dense sampling in high ozone concentration areas.

The three different algorithms are quantitatively evaluated using two different performance indicators, depending on the error between the ground truth and the model after each time step. The performance indicators are calculated over the whole domain, using a fine grid resolution with M samples. The first one is the *Root Mean Squared Error* (RMSE), that reflects the error in estimation independent of the predicted value. This means that RMSE indicator gives the same importance to the error where the ozone concentration is low, than where the ozone concentration is high. Therefore, we believe RSME is not the best indicator for EM applications. The second performance indicator is the *Weighted Root Mean Squared Error* (WRMSE),

$$\text{WRMSE} = \sqrt{\frac{\sum_{i=1}^{M}\left[\frac{(\mu(\mathbf{x}^*)-\mu_{gt}(\mathbf{x}^*))(\mu_{gt}(\mathbf{x}^*)-\min\mu_{gt}(\mathbf{x}^*))}{\max\mu_{gt}(\mathbf{x}^*)-\min\mu_{gt}(\mathbf{x}^*)}\right]}{M}}. \tag{11}$$

WRMSE is essentially identical to RMSE but the error is multiplied by a factor that depends on the mean of the predicted value, normalised between the minimum and the maximum over the entire domain. This performance indicator gives more importance to the error in areas with higher values of the studied phenomenon, a logical assumption if we are interested in potentially dangerous areas for humans. We believe that WRMSE is a suitable performance indicator for this particular application, considering that ozone is a pollutant at ground level.

Table I presents the results for each sampling method using the two different indicators. For each method, we calculate the mean and standard deviation of the indicator through time, and the total distance the robot travelled.

Overall, it can be seen that the DUCB based method has the smallest error in RMSE and in WRMSE. The standard deviation of DUCB for the RMSE indicator is higher than the other strategies because DUCB avoids sampling in areas where the ozone concentration is known to be low. Therefore, it presents small error in high ozone concentration areas and bigger error in lower ozone concentration areas, increasing total standard deviation. The IG approach has the bigger error but smaller standard deviation because it distributes measurements uniformly over the entire domain. Regarding WRMSE, the DUCB method has the best results in terms of mean and standard deviation, because the error in lower ozone concentration areas is less relevant (11). The distance travelled by the DUCB sampling strategy is 11% smaller than the other approaches, due to the distance penalty in the acquisition function. This means that our proposed DUCB acquisition function results in lower error and at the same time reduces energy consumption.

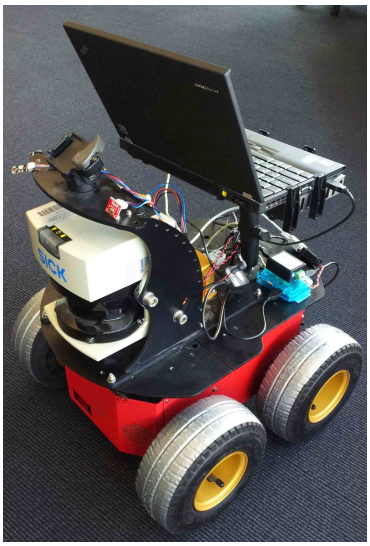


Fig. 4. Mobile robot used for indoor experiment.

B. Real World Experiment

The studied approach for IEM was tested in a real environment using an autonomous robot (Figure 4). The robot is equipped with a laser scanner, odometry, and an ambient light sensor. It uses a laptop running ROS[4] as main processing unit. The built-in packages in ROS deal with localisation and mapping using the laser scanner and odometry. Therefore, the robot is assumed to be properly localised during the whole measurement process.

The monitored phenomenon is the ambient light distribution in an indoor office environment. Figure 6a shows the map built autonomously by the robot, where the shaded area represents the selected area for modelling the phenomenon. To create the ground truth of the intensity of light, the robot was driven manually over the whole domain taking a considerable amount of measurements. These were used to train a GP as the ground truth (shown in Figure 5a), with optimal hyper-parameters ($\{\sigma_f, \ell_{lat}, \ell_{long}, \sigma_n\} = \{40.01, 0.7, 0.7, 6.02\}$).

[4]Robot Operating System `http://www.ros.org`

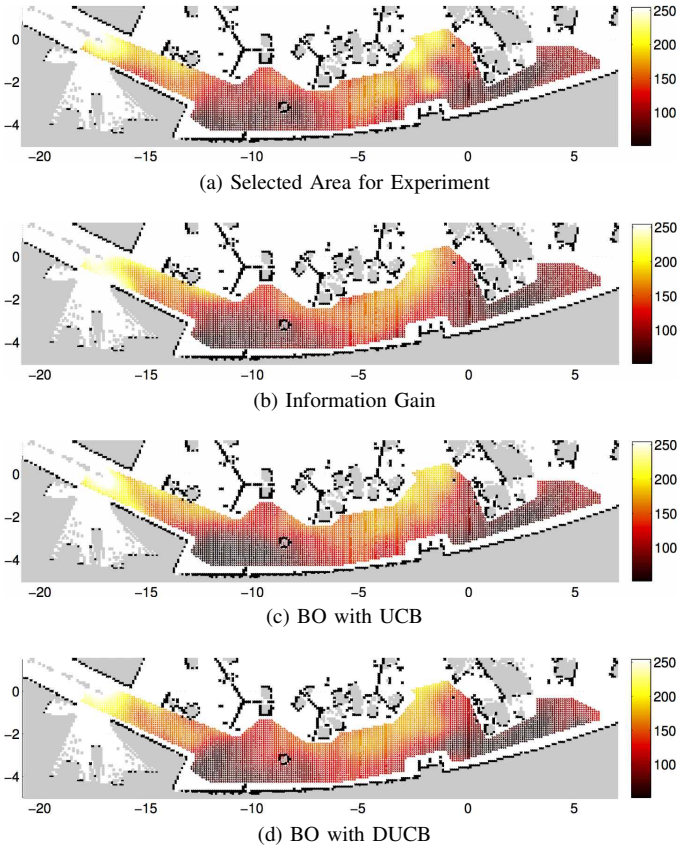


Fig. 5. Estimated mean of a GP fitted to training data of each sensing strategy. Light intensity with no SI units. Axis show distance in meters.

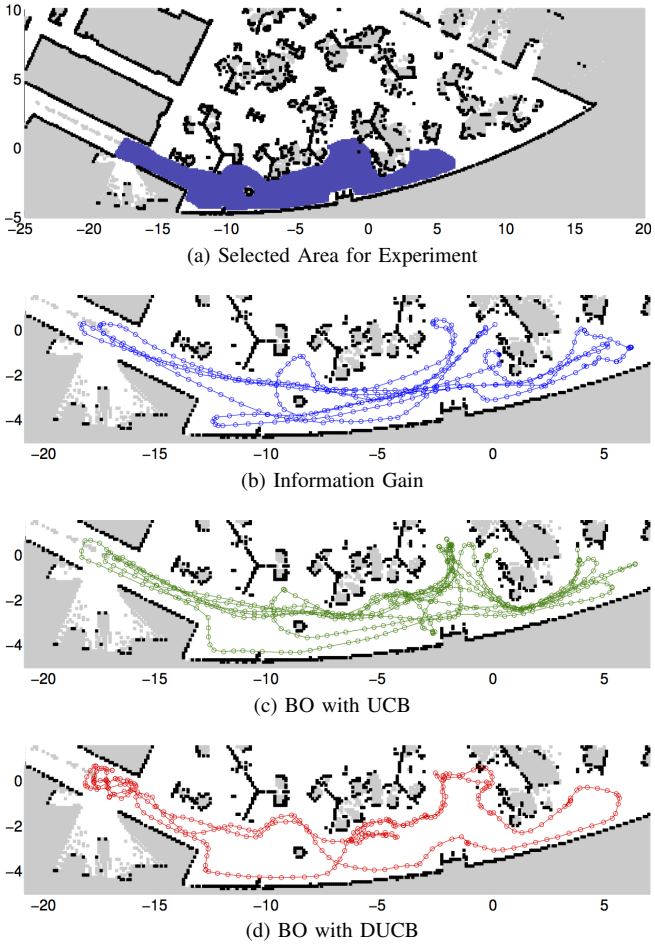


Fig. 6. The trajectories followed by the robot. Samples are only taken inside the allowed area defined by (a). Axis show distance in meters.

The three different approaches tested in V-A are now run in the described real environment using the autonomous robot. The intensity of the ambient light was kept constant during the execution of the experiment to allow repeatability (no variation of the process with time). Each sensing strategy was tested for 10 minutes, acquiring one measurement per second. The final model built by each policy is displayed in Figure 5, where higher light intensity is equivalent with high values in the plot (no known scale of luminosity can be provided due to the nature of the sensor). The three models are very similar because for this experiment the process does not vary with time, i.e. it is expected that all policies produce an accurate model if they have visited the entire domain.

Figure 6 shows the trajectories followed by the robot using the three sampling strategies. It can be seen that the resulting trajectories of the IG and UCB approaches are spread over the whole domain. This is because after reaching a goal the next selected sensing location is very far from the last one. A slight difference can be observed when using the UCB strategy (Figure 6b), because sensing locations are more concentrated in areas with higher luminosity values. The trajectory followed by the robot using our proposed acquisition function DUCB (Figure 6d) is clearly more concentrated in areas of higher ambient light. In contrast, areas of lower ambient light are only visited once during the experiment, which is enough for the underlying process in this experiment. Comparing against a distance penalised information gain criterium (DIG) may be possible; however, we believe this acquisition function is not suitable as it does not consider the mean of predictions.

TABLE II
RESULTS FOR REAL EXPERIMENT

Indicator	Method	Value	Distance $[m]$
RMSE	IG	16.44	157.01
RMSE	UCB	17.16	217.45
RMSE	DUCB	16.91	98.08
WRMSE	IG	5.41	157.01
WRMSE	UCB	4.60	217.45
WRMSE	DUCB	4.70	98.086

Table II details the error of prediction for each sensing method, using the RMSE and WRMSE over the whole spatial domain (No mean and standard deviation of the indicators provided as the process does not vary with time). It can be seen that the RMSE is similar for all three different strategies and that no relevant differences are present. However, our proposed policy using distance penalty reduces the travelled distance almost to half of the other strategies, while building a similar model. Having said that, it is clear that our sensing strategy exposes a more efficient way of sampling.

VI. Conclusion

This paper proposes a new planning strategy for environmental monitoring. The main contributions are:

1) Extending the BO framework to choosing sensing locations in environmental monitoring applications. This allows mobile robots to take informed decisions based on the spatial-temporal GP model of the phenomena, getting as much relevant information as possible.
2) Presenting a new acquisition function, called Distance based Upper Confidence Bound. This acquisition function addresses indirectly the exploration-exploitation tradeoff. In addition, it is energy efficient since it reduces the total distance travelled by the robot.

The proposed methodology was evaluated using simulation and a real experiment. The results show a notable improvement in terms of error when monitoring extreme areas of time varying processes (ozone concentration experiment). A considerably important result is that our proposed acquisition function helps reducing travel distances to achieve comparable results to other methods. This can be appreciated in the simulated and in the real experiments, where the travelled distance was reduced by 40%.

The contributions made in this paper enable an intelligent autonomous robot to efficiently monitor the state of an environmental phenomenon. We believe that using BO and designing new acquisition functions can help dealing with the difficulties when sampling complicated environmental processes.

As future work, we expect to include experiments using non-stationary covariance functions for dynamic processes. Admitting a variation in time dependancy over space can lead to interesting behaviour of autonomous robots, such as sampling more frequently in highly dynamic areas. Finally, we plan to study the maximisation of the acquisition function in a parametrized trajectory space, finding a continuos path that provides the highest information content for various types of acquisition functions.

References

[1] M. Dunbabin and L. Marques, "Robotics for environmental monitoring," *IEEE Robotics and Automation Magazine*, 2012.

[2] E. Brochu, V. M. Cora, and N. de Freitas, "A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning," University of British Columbia, Tech. Rep., 2010.

[3] M. Hoffman, E. Brochu, and N. de Freitas, "Portfolio allocation for bayesian optimization," in *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2011.

[4] M. Osborne, "Bayesian gaussian process for sequential prediction, optimisation and quadrature," Ph.D. dissertation, University of Oxford, 2010.

[5] C. E. Rasmussen and C. Williams, *Gaussian processes for machine learning*. The MIT Press, Cambridge, Massachuset, 2006.

[6] B. Ferris, D. Hahnel, and D. Fox, "Gaussian processes for signal strength-based location estimation," in *Robotics Science and Systems (RSS)*, 2006.

[7] C. Stachniss, C. Plagemann, and A. Lilienthal, "Gas distribution modeling using sparse Gaussian process mixture models," in *Robotics Science and Systems (RSS)*, 2008.

[8] A. Singh, F. Ramos, H. D. Whyte, and W. J. Kaiser, "Modeling and decision making in spatio-temporal processes for environmental surveillance," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2010.

[9] V. Tresp, "Mixtures of Gaussian Processes," in *Neural Information Processing Systems (NIPS)*, 2000.

[10] D. Higdon, "Space and space-time modeling using process convolutions," *Quantitative methods for current environmental issues*, 2002.

[11] C. Gaetan and X. Guyon, *Spatial Statistics and Modeling*. Springer New York, 2010.

[12] M. Reggente and A. J. Lilienthal, "Using local wind information for gas distribution mapping in outdoor environments with a mobile robot," in *IEEE Sensors*, 2009.

[13] C. Guestrin, A. Krause, and A. P. Singh, "Near-optimal sensor placements in Gaussian processes," in *International Conference on Machine Learning (ICML)*, 2005.

[14] D. Mackay, *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.

[15] R. Stranders, A. Rogers, and N. Jennings, "A decentralized, on-line coordination mechanism for monitoring spatial phenomena with mobile sensors," in *International Workshop on Agent Technology for Sensor Networks (ATSN)*, 2008.

[16] A. Krause and C. Guestrin, "Near-optimal Observation Selection using Submodular Functions," in *AAAI Conference on Artificial Inteligence*, 2007.

[17] A. Singh, A. Krause, C. Guestrin, W. Kaiser, and M. Batalin, "Efficient Planning of Informative Paths for Multiple Robots," in *International Joint Conference on Artificial Inteligence (IJCAI)*, 2007.

[18] A. Singh and A. Krause, "Nonmyopic adaptive informative path planning for multiple robots," in *International Joint Conference on Artificial Inteligence (IJCAI)*, 2009.