2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Bayesian Optimisation for Informative Continuous Path Planning

Roman Marchant and Fabio Ramos

Abstract— **Environmental monitoring with mobile robots requires solving the informative path planning problem. A key challenge is how to compute a continuous path over space and time that will allow a robot to best sample the environment for an initially unknown phenomenon. To address this problem we devise a layered Bayesian Optimisation approach that uses two Gaussian Processes, one to model the phenomenon and the other to model the quality of selected paths. By using different acquisition functions over both models we tackle the exploration-exploitation trade off in a principled manner. Our method optimises sampling over continuous paths and allows us to find trajectories that maximise the reward over the path. We test our method on a large scale experiment for modelling ozone concentration in the US, and on a mobile robot modelling the changes in luminosity. Comparisons are presented against information based criteria and point-based strategies demonstrating the benefits of our method.**

I. Introduction

One of the most significant impacts of mobile robotics is in environmental monitoring. Climate change; air, water, land and acoustic pollution; solar power intensity; tidal wave behaviour among many others, are highly complex processes that have a noticeable impact on human well-being. Scientists have shown a great deal of interest in understanding these phenomena, specially in areas such as heath, mining, energy generation, agriculture, forestry and many more. However, deterministic differential equations representing these processes are difficult to devise, and building mathematical models requires gathering large quantities of data distributed over space and time. The procedure can be tedious and extensive. Deploying an autonomous robot for this process is both cost-effective and convenient as,

1) Robots can build statistical models of the environment and choose sampling locations intelligently;
2) Robots can move to more informative sensing locations adding flexibility over static sensor networks;
3) Robots can automate the sampling procedure reducing human supervision;
4) Robots can access areas that are dangerous for humans.

The use of autonomous robots for environment monitoring has expanded massively over the past decade. Hardware capabilities have increased noticeably, giving robots the power of traversing over a wide range of environments and monitoring several phenomena. Information gathering techniques have also seen interesting developments. However, the problem of where and when to gather the most informative samples in an efficient manner is still an open question.

This paper addresses the problem by finding informative paths in a continuous domain, solving not only the question of *where* and *when* to sample, but *how* to get there. We developed a decision making algorithm to identify areas of interest of an initially unknown environmental phenomenon. The decisions are based on an incremental spatial–temporal model of the phenomenon using a *Gaussian Process* (GP) that take into account the uncertainty and predicted values in time. The method is based on *Bayesian Optimisation* (BO) [1] techniques that can naturally deal with the exploration-exploitation trade off. Each decision is chosen using an acquisition function that is maximised w.r.t. the parameters of an unknown path. We solve this maximisation using another layer of BO. This allows us to find the best set of parameters that determine a continuous path where the robot travels on while taking samples. We validate our algorithm on a large-scale environment for monitoring ozone concentration in the US, and on a mobile robot that monitors the dynamics of luminosity changes.

This work builds on earlier work by the authors [2], and presents the following contributions:

1) Generalisation of the BO algorithm to optimise along continuous trajectories instead of discrete locations;
2) A layered BO for informative path planning in spatial–temporal environmental monitoring.

The paper is structured as follows. In section II we review the current state-of-the-art in modelling and decision making for environmental monitoring under uncertainty. Section III describes GP regression for space-time models. Section IV presents the theory behind our decision making algorithm, describing the BO algorithm, and the optimisation procedure to determine sampling paths. Section V shows experimental results and comparisons against competing approaches. Finally, section VI draws conclusion and presents ideas for future work.

II. Related Work

Over the last decade, a vast amount of research efforts have been dedicated to field robotics. Particularly environmental monitoring using mobile robots is gaining popularity among a wide range of applications [3–5]. The two main areas of interest for researchers are how to create statistical models of environmental phenomena, and where to acquire more measurements to improve on current models.

A very popular Bayesian technique for modelling spatial-temporal phenomena are GPs [6]. In robotics, GPs have been

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

978-1-4799-3684-7/14/$31.00 ©2014 IEEE

used for gas concentration modelling [2, 4], terrain modelling [7], wireless strength signal modelling [8] and occupancy maps [9] among many others. Specific covariance functions have been found to favour learning spatial and temporal correlations in data [5] and several approximation techniques have been developed to handle large amount of data points [7, 10, 11].

Decision making under uncertainty has seen significant developments as well. Information gain strategies for placing static sensors were studied in [12]. Mobile sensing agents can use active learning to choose where to sample from the environment. For this purpose, Markov Decision Processes and Reinforcement Learning approaches have been used by [13] and [14]. [13] uses belief-state MDPs for selecting observations that minimise uncertainty and [14] uses GPs for modelling the state-action value function. The main limitation of these approaches is the limited action space and tractability to deal with real scenarios. Uncertainty-driven planning was also explored by [15] using the travel salesman problem and RRTs [16] to find paths. RRTs are also explored by [17], where a record the minimum cost cycle is considered to find cyclic trajectories. [18] combines Rapidly Exploring Random Graphs (RRGs) and Branch and Bound optimisation techniques to find informative paths. An optimisation approach has been explored by [19], that uses simulated annealing and swarm optimisation for planning energy-optimal paths for an AUV under strong currents. The main drawback of this work is that uncertainty is not estimated by the model used to derive decisions. Another cost aware path planner was presented in [20]. It assumes an already known cost map and is not useful for exploration purposes.

Recent decision making algorithms make use of sub-modularity properties for planning non-myopic, long-term way points for uncertainty reduction [12, 21–23]. These methods provide convergence guarantees and error bounds based on an exploration-only behaviour. While minimising the overall uncertainty of the model is important in some applications, for most of pollution monitoring tasks this is not sufficient. In such cases, it is desirable to be more accurate in areas of high pollution than in areas of low pollution. This introduces extra terms in the objective function (such as the mean of the predicted pollution concentration) making the sub-modularity assumption invalid.

The proposed method has the following advantages over the previous techniques: i) It is not a way-point greedy solution to acquiring new observations as it takes into account measurements obtained along a path with predictions propagated over time; ii) It considers a continuous action space; iii) It uses both the mean and the variance to define paths and addresses the exploration-exploitation trade off in principled Bayesian framework.

III. GPs for Spatial–Temporal Modelling

Gaussian Processes have been a popular tool for regression problems, particularly for space and time correlated data [24]. In this paper, GPs are used to model a spatial–temporal phenomenon from observations collected by a mobile robot, to predict the value of a unknown noisy function in space and time. In this section we briefly describe the theory behind GPs for regression and inference in space-time modelling.

A GP is a nonparametric Bayesian technique that places a prior distribution over the space of functions mapping inputs to outputs. A latent noisy function $f(\mathbf{s};t)$ representing the realisation of a spatial–temporal environmental phenomenon is modelled as $y = f(\mathbf{s};t) + \epsilon$, where $\mathbf{s} \in \mathbb{R}^D$ are the coordinates in a spatial D-dimentional space, $t > 0 \in \mathbb{R}$ represents time and $\epsilon \sim \mathcal{N}\left(0, \sigma_n^2\right)$ is the noise associated to each independent observation.

A GP is fully determined by a spatial-temporal mean function $m(\mathbf{s};t)$ and a positive semi-definite covariance function $k\left((\mathbf{s};t),(\mathbf{s};t)'\right)$, i.e.,

$$f(\mathbf{x}) \sim \mathcal{GP}\left(m(\mathbf{s};t), k\left((\mathbf{s};t),(\mathbf{s};t)'\right)\right). \tag{1}$$

In this paper we assume that the mean function is constant, $m(\mathbf{s};t) = \eta$, and learnt from data. Treatment for non-constant mean functions can be found in [6].

Denoting the set of all inputs as $X = \{(\mathbf{s};t)_i\}_{i=1}^N$ and corresponding outputs $\mathbf{y} = \{y_i\}_{i=1}^N$, the predictive distribution for a new query input $(\mathbf{s};t)^*$ is Gaussian, $f((\mathbf{s};t)^*) \sim \mathcal{N}\left(\mu, \sigma^2\right)$, where the mean, μ, and variance σ^2, are:

$$\begin{aligned} \mu((\mathbf{s};t)^*) = &\quad K((\mathbf{s};t)^*, X)K_X^{-1}(\mathbf{y} - m(X)), \\ \sigma^2((\mathbf{s};t)^*) = &\quad K((\mathbf{s};t)^*, (\mathbf{s};t)^*) - \\ &\quad K((\mathbf{s};\mathbf{t})^*, X)K_X^{-1}K(X, (\mathbf{s};\mathbf{t})^*). \end{aligned} \tag{2}$$

Here, $K_X = K(X, X)$ is the covariance matrix between all observations where each element (i, j) is calculated as $k_{i,j} = k((\mathbf{s};t)_i, (\mathbf{s};t)_j)$.

Covariance functions encode a degree of relationship between inputs. A covariance function usually has a set of hyper-parameters, $\boldsymbol{\theta}$, representing properties such as length scales, amplitude, etc . There are several covariance functions in the literature with different characteristics. Spatial–temporal data can be handled in various ways. In the simplest case, time can be considered as an extra dimension in the input space, which is not convenient if we wish to capture time-specific behaviour. The other option is to consider separable or non-separable covariance functions. Separable covariance functions are essentially the product of two independent covariance functions, one defined over space and the other defined over time,

$$k_{sep}((\mathbf{s};t),(\mathbf{s};t)'|\boldsymbol{\theta}) = k_1(\mathbf{s},\mathbf{s}'|\boldsymbol{\theta_1})\, k_2(t,t'|\boldsymbol{\theta_2}). \tag{3}$$

This type of covariance function is convenient when time-specific behaviour does not depend on the spatial location. On the other hand, non-separable covariance functions can capture more complex dependencies when space and time are coupled [5].

In this work we use separable covariance functions by combining the following components with sums and products in the time and space domain:

1) Matérn 3:

$$k_{\mathrm{m3}}(\mathbf{s},\mathbf{s}'|\sigma_f, L) = \quad \sigma_f\left(1 + \sqrt{3}r\right)\exp\left(-\sqrt{3}r\right), \tag{4}$$

where $r = (\mathbf{s} - \mathbf{s}')L(\mathbf{s} - \mathbf{s}')^T$, L is a diagonal matrix whose components are length scales $L_{ii} = l_i^{-2}$ for each dimension, and σ_f is the signal variance.

2) Matérn 5:

$$k_{\text{m5}}(\mathbf{s}, \mathbf{s}'|\sigma_f, L) = \quad \sigma_f \left(1 + \sqrt{5r} + \tfrac{5r}{3}\right) \exp\left(-\sqrt{5r}\right), \tag{5}$$

where $r = (\mathbf{s} - \mathbf{s}')L(\mathbf{s} - \mathbf{s}')^T$, σ_f is the signal variance and L is a diagonal matrix whose components are length scales $L_{ii} = l_i^{-2}$ for each dimension.

3) Periodic covariance function:

$$k_p(t, t'|\sigma_{f_2}, \gamma, \varphi) = \sigma_{f_2} \exp \frac{\left(-2 \sin^2\left(2\pi\varphi(t - t')\right)\right)}{\gamma}, \tag{6}$$

where φ is the frequency and γ is the smoothness of the periodic component.

SectionV details the components used for each experiment and the shows the set $\boldsymbol{\theta}$ of hyper-parameters.

Finding the best set of values for the hyper-parameters $\boldsymbol{\theta}^\star$ can be achieved by maximising the *log marginal likelihood* (LML) of the data,

$$\boldsymbol{\theta}^* = \max_{\boldsymbol{\theta}} \ \text{LML}(\mathbf{y}, \mathbf{X}, \boldsymbol{\theta}) \,, \tag{7}$$

with,

$$\text{LML}(\mathbf{y}, \mathbf{X}, \boldsymbol{\theta}) = -\frac{1}{2}\mathbf{y}^T K_X^{-1}\mathbf{y} - \frac{1}{2}\log|K_X| - \frac{n}{2}\log 2\pi \,. \tag{8}$$

In problems with large amounts of data, as those involving robots collecting measurements for long periods of time, an approximation is used for reducing the computational cost of inference and learning. In this paper we used the nearest neighbour approximation [7] where for a query point $(\mathbf{s}, t)^*$, only the m nearest neighbours in the covariance space are used to calculate the inversion of K_X. This means that K_X is no longer an $N \times N$ matrix, but an $m \times m$ matrix with $m \ll N$, reducing the computational complexity from $\mathcal{O}(N^3)$ to $\mathcal{O}(m^3)$.

IV. Continuous Path Planning

After building a probabilistic model of the studied phenomenon using the methodology in section III, a Bayesian optimisation method is derived to estimate continuous paths for sampling. In this section we describe the general BO algorithm, to later present details on the specific algorithm for planning over continuous paths.

A. Bayesian Optimisation

BO is used for finding the optimal (maximum or minimum) of an unknown and costly to evaluate function f, i.e. find $\mathbf{x}^\star = \arg\max_{\mathbf{x}} f(\mathbf{x})$. To achieve this, it builds a statistical model of the unknown function by collecting samples and a GP prior. An acquisition function h is evaluated over the statistical model and guides the search for the optimum. The procedure requires the maximisation of h at each iteration which is usually a much simpler optimisation problem. The generic algorithm for BO is shown in Figure 1. Line 2 is the inner optimisation conducted within BO and can be solved using gradient-based optimisers [1].

Inputs: f, h
Outputs: $\mathbf{x}^\star$, $f(\mathbf{x}^\star)$
1: **for** $j = 1, 2, 3, \ldots$ {Max iterations} **do**
2: Find $\mathbf{x}_j = \arg\max_{\mathbf{x}} h(\mathbf{x})$
3: $y_j \leftarrow f(\mathbf{x}_j)$ {Gather sample from f}
4: Augment training set with $(\mathbf{x}_j, y_j)$.
5: Update GP
6: **if** $y_j > \mu(\mathbf{x}^\star)$ **then**
7: $\mathbf{x}^\star \leftarrow x_j$ {Update location of optimum}
8: **end if**
9: **end for**

Fig. 1. General Bayesian Optimisation algorithm.

The most common acquisition functions are:

- Probability of Improvement [25]

$$\text{PI}(\mathbf{x}) = P(f(\mathbf{x}) \leq f(\mathbf{x}^+) + \xi) \tag{9}$$

$$= \Phi\left(\frac{\mu(\mathbf{x}) - g(\mathbf{x}^+) - \xi}{\sigma(\mathbf{x})}\right); \tag{10}$$

- Expected Improvement [25]

$$\text{EI}(\mathbf{x}) = \sigma(\mathbf{x})\left[Z\Phi(Z) + \phi(Z)\right] \,, \tag{11}$$

where

$$Z = \frac{\mu(\mathbf{x}) - g(\mathbf{x}^+) - \xi}{\sigma(\mathbf{x})} \,;$$

- Upper Confidence Bound [26]

$$\text{UCB}(\mathbf{x}) = \mu(\mathbf{x}) + \kappa\sigma(\mathbf{x}) \,, \tag{12}$$

where $\mathbf{x}^+$ is the location of the best sample gathered so far; ξ, κ are exploration-exploitation tunning parameters; and ϕ, Φ are the normal probability density function and normal cumulative distribution function respectively.

For space-time problems, inputs $\mathbf{x} = (\mathbf{s}; t)$ transform an ordinary acquisition function into a space-time acquisition function that considers time. Equation 2 is then used to calculate the mean and variance of the GP model.

B. Path Parametrisation

We use BO to find the path that provides the most useful information for an autonomous robot monitoring the environment. A path is here defined as a curve $\mathcal{C}$; a continuous function mapping $\mathbb{R}$ to $\mathbb{R}^D$, where D is the dimension of the spatial component. $\mathcal{C}(u|\boldsymbol{\beta})$ is dependent on a set of parameters $\boldsymbol{\beta}$, with $u \in [0, 1]$. In this paper we use cubic splines restricted to a two dimensional plane, $D = 2$. $\mathcal{C}$ can be defined as

$$\mathcal{C}_1(u) = a_x u^3 + b_x u^2 + c_x u + d_x \tag{13}$$

$$\mathcal{C}_2(u) = a_y u^3 + b_y u^2 + c_y u + d_y \tag{14}$$

and $\boldsymbol{\beta} = \{a_x, b_x, c_x, d_x, a_y, b_y, c_y, d_y\}$. When estimating a path to follow, the current state of the robot $\mathbf{x} = (p_x, p_y, \alpha)$ is assumed known. This defines the boundary condition for the curve: $\mathcal{C}(u = 0) = \mathbf{x}_i = (p_{x_i}, p_{y_i}, \alpha_i)$. Therefore:

$$\mathcal{C}_1(u = 0) = p_{x_i} = d_x, \tag{15}$$

$$\mathcal{C}_2(u = 0) = p_{y_i} = d_y, \tag{16}$$

$$\left.\frac{\partial \mathcal{C}_2/\partial u}{\partial \mathcal{C}_1/\partial u}\right|_{u=0} = \frac{c_y}{c_x} = \tan\alpha_i. \tag{17}$$

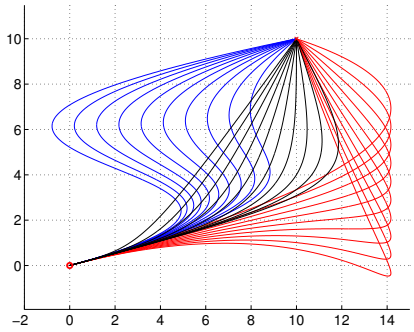


Fig. 2. Paths generated with $\mathbf{x}_i = (0, 0, 15°)$. $\mathcal{C}_1(u = 1) = p_{x_f} = 10$, $\mathcal{C}_2(u = 1) = p_{y_f} = 10$, $50 < a_x < 110$ (red), $0 < a_y < 50$ (blue) and $10 < c_x < 40$ (black)

This reduces the action space to five free parameters, $\boldsymbol{\beta} = \{a_x, a_y, b_x, b_y, c_x\}$ only. Different values of these parameters produce curves with different lengths and shapes. Figure 2 shows a small set of the possible curves drawn for different values of a_x, a_y, c_x and settings $\mathcal{C}_1(u = 1) = p_{x_f} = 10$ and $\mathcal{C}_2(u = 1) = p_{y_f} = 10$.

To take into account the temporal dimension over the curve we assume that the robot travels at a constant speed v. Note that our method accepts modifications to the curve parametrisation and is not strongly linked to this particular spline model.

C. BO for Continuous Path Planning

In environmental monitoring problems one is typically interested in providing accurate predictions in areas of high concentration or anomalies [2]. For example, gas concentration, smoke for fire-detection, temperature, pollution, among others. We assume that a robot will travel along paths that follow the parametrisation described in Section IV-B. To evaluate the score, r, of a path $\mathcal{C}(u, \boldsymbol{\beta})$, we integrate the acquisition function h over the path,

$$r(\mathcal{C}(u, \boldsymbol{\beta})|h) = \int_{\mathcal{C}(u,\boldsymbol{\beta})} h(v)\mathrm{d}v. \tag{18}$$

Considering $h = \text{UCB}$,

$$r(\mathcal{C}(u, \boldsymbol{\beta})|h) = \int_{\mathcal{C}(u,\boldsymbol{\beta})} h(v)\mathrm{d}v \tag{19}$$

$$= \int_0^1 \text{UCB}(\mathcal{C}(u, \boldsymbol{\beta}))\|\mathcal{C}'(u, \boldsymbol{\beta})\|\mathrm{d}u \tag{20}$$

$$= \int_0^1 \left[\mu(\mathcal{C}(u, \boldsymbol{\beta})) + \kappa\sigma(\mathcal{C}(u, \boldsymbol{\beta}))\right] \|\mathcal{C}'(u, \boldsymbol{\beta})\|\mathrm{d}u. \tag{21}$$

The integral in equation 18 does not always have an analytical solution, depending on the definition of the acquisition and covariance functions. In this case we use a rectangle rule quadrature-based approximation [27], which generally results in accurate approximations for the one dimensional case (since the integral is over a 1-D variable, u).

To use BO for finding continuous paths a modification to the general algorithm shown in Figure 1 is required. Instead of finding a discrete location for taking the next sample from f (Line 2 of Figure 1), we find the parameters

Inputs: f, h, q
1: **for** $j = 1, 2, 3, \ldots$ **do**
2: $\quad\boldsymbol{\beta}^\star \leftarrow \text{BO}(r(h), q)$ {Algorithm in Figure 1}
3: $\quad\{x, y\} \leftarrow$ Sample along $\mathcal{C}(u, \boldsymbol{\beta}^\star)$
4: $\quad$Augment training set with $\{x, y\}$.
5: $\quad$Update GP.
6: **end for**

Fig. 3. BO for continuous path planning.

$\boldsymbol{\beta}^\star$ that define a continuous path over space and time that cumulatively delivers the best return by integrating over the acquisition function (Line 2 of Figure 3). To find the *best* set of parameters $\boldsymbol{\beta}^\star$ that defines the curve that maximises the integral over h, the following optimisation problem is solved,

$$\boldsymbol{\beta}^\star = \arg\max_{\boldsymbol{\beta}} r(\mathcal{C}(u, \boldsymbol{\beta})|h). \tag{22}$$

In section IV-A, the optimisation required to find the highest value of the acquisition was performed using gradient-based optimisers. Alternatively, equation 22 can be solved using another layer of BO. Since the action space $\boldsymbol{\beta}$ is five dimensional and the function r is highly non-convex and expensive to evaluate, BO provides a natural solution. This is performed by placing a GP prior over r and using a second acquisition function q to decide which path parameters to evaluate over r. This second layer of BO (Line 2 in Figure 3) follows the classic algorithm described in Figure 1. Note this step is fast to evaluate because it does not require the robot to move and gather training samples as it uses the existing GP model of the phenomenon. The complete algorithm for finding the most informative path using two layered BO is shown in Figure 3. Line 3 consists of the data acquisition process, where the robot moves and gathers samples along the path found in the previous step.

V. Experiments

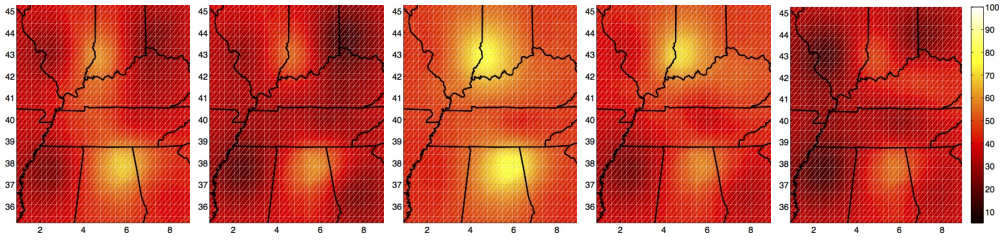
In this section the proposed method is tested in two scenarios: a large-scale experiment for ozone monitoring in the US, and real-time monitoring of illumination with a mobile robot.

A. Large-scale pollution monitoring

The first experiment simulates an Unmanned Air Vehicle (UAV) monitoring ozone concentration; considered a pollutant at ground level. To simulate the environment we use real data provided by the US Environment Protection Agency[1]. A large number of ozone concentration measurements, dating back to 1987, are available with one hour period for static sensing locations across the US. The UAV is forced to stay within the region specified in Figure 5.

The discrete data $\mathcal{S}$ from the database is used to create a simulated environment using GP regression, called *Ground Truth GP* (GTGP). A robot samples from the mean μ_{gt} of this continuous process in space and time. The GTGP uses a separable covariance function with an isotropic Matérn3 component for space (Equation 4), and the sum of a Matérn5

[1] http://java.epa.gov/castnet/reportPage.do



(a) $t = 0$ days (12am) (b) $t = 0.25$ days (6am) (c) $t = 0.5$ days (12pm) (d) $t = 0.75$ days (6pm) (e) $t = 1$ days (12am)

Fig. 4. Ground-truth for ozone concentration across the US. Axis are measured in $10^6 m$ and corresponding to UTM coordinates for section 16F.

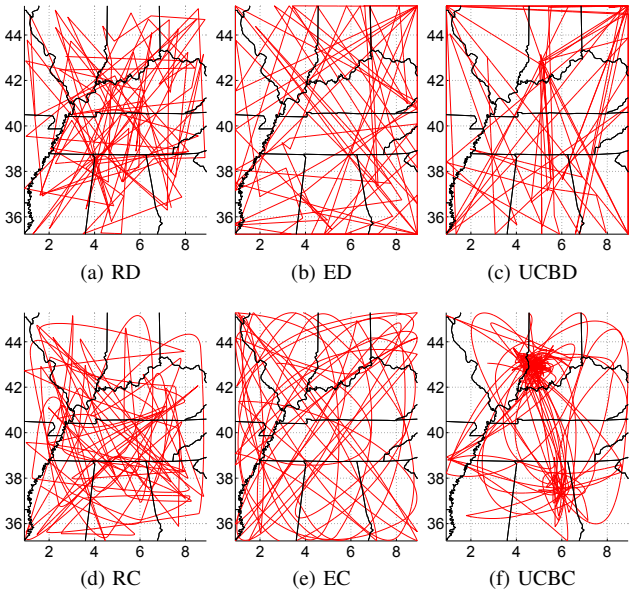
Fig. 5. Area for the experiment (16F in UTM coordinates)

and a Periodic component for time (Equation 5 and 6 respectively). The set of optimal hyper-parameters was found by maximising equation 7 using a gradient decent method. The optimal values found by the optimiser are: $\sigma_{fm3} = 1.862$, $l_{m3} = 1.195$, $\sigma_{fm5} = 0.201$, $l_{m5} = 5.84$, $\sigma_{fper} = 0.94$ and $\gamma = 5.75$. In order to deal with limited computational resources, the approximation parameter m, described in Section III, is set to 300 for all experiments. The concentration of ozone changes periodically with a period of one day (known a priori), and the time is measured in days. Figure 4 shows a GTGP regression over space for five timestamps within one day. It can be noted that two peaks appear around midday with values that can reach up to 100ppb. The pattern is repeated every-day with slight variations in amplitude due to unknown environmental factors.

Ideally, the robot should accurately capture changes in the areas where pollution is more densely concentrated. We compare six different techniques for planning the motion of the UAV while monitoring the environment:

a) Random Discrete Sampling (RD): Randomly pick discrete goal locations within the environment.
b) Entropy Discrete Sampling (ED): Pick discrete locations for sampling using the maximum variance (entropy) criterium [12].
c) UCB Discrete Sampling (UCBD): Pick discrete locations using UCB [2].
d) Random Continuous Sampling (RC): Select random paths using an uniform distribution over $\boldsymbol{\beta}$.
e) Entropy Continuous Sampling (EC): Find paths that maximise entropy reduction.
f) UCB Continuous Sampling (UCBC): Choose paths that maximise UCB acquisition function.

The method presented in the paper is indicated in f). UCBC and UCBD use the upper confidence bound acquisition function, with parameter $\kappa = 0.1$ manually tuned to balance



(a) RD (b) ED (c) UCBD (d) RC (e) EC (f) UCBC

Fig. 6. Paths for different methods. Axis are measured in $10^6 m$.

the exploration-exploitation trade off. Although UCB was chosen due to its particular exploring behaviour [1], other acquisition functions can be used. Future work can compare different acquisition functions and learn the parameters of the acquisition function within the optimisation procedure. To include realistic sampling from the GTGP, random noise with $\sigma_n = 5$ is added to every sample independently. The experiment takes place for 30 days and assumes the vehicle moves at an average speed of $60 km/h$. A sample of ozone concentration is collected every minute for all strategies, assuring that each method collects the same number of samples for predicting the values of the phenomenon. Given that all methods will acquire the same number of samples, the differences in error will only depend on the locations where the samples were acquired. The inner optimisation for maximising among paths (Line 2 of Figure 2) uses $q = UCB$ as acquisition function for strategies EC and UCBC. The GP model of the inner optimisation uses a Matérn3 (Equation 4) covariance function whose hyper-parameters are optimised on each iteration using gradient decent.

Figure 6 shows the paths travelled by the robot for each case. A quick visual inspection shows that all methods were able to cover the region of interest and explore the entire

environment. Random sampling strategies (RD and RC) do not present any interesting patterns and move chaotically across the studied area. Entropy based techniques (ED and EC) cover the region uniformly, reducing the uncertainty of the whole area. Finally, UCBD and UCBC concentrate their samples towards the areas of higher pollution.

A very important difference is the shape of paths for discrete and continuous sampling strategies. Even though κ has the same value for the acquisition function of UCBC and UCBD, the trajectories are much more concentrated over the high pollution areas for the continuous optimisation case (UCBC). The main reason is that this method takes into account the value of the acquisition function over the entire path that is being traversed. One way of seeing this is that if a method only takes into account a discrete goal location it will not necessarily collect useful information on its way to the target location. However, if the method does take into account the information gathered while reaching the target location, then the informativeness of gathered samples will increase noticeably.

We use four different error measures at M locations to evaluate the performance quantitatively:

i) *Root Mean Squared Error* (RMSE): Error without taking into account the value of the predicted variance.

ii) *Weighted Root Mean Squared Error* (WRMSE): Places weights depending on the magnitude of the ground truth output, giving more importance to errors in higher pollution areas [2].

iii) *Mean Log Loss* (MLL): Evaluates the negative log probability of the ground truth data point under the model. Takes into account not only the prediction error but also the associated uncertainty.

$$MLL = \frac{\frac{\sum_{i=1}^{M}\left(-\log p(\mu_{gt}(\mathbf{x}_i^\star)|\mathcal{S},\mathbf{x}_i^\star)\right)}{M}}{\frac{\sum_{i=1}^{M}\left(\frac{1}{2}\log\left(2\pi\sigma^2(\mathbf{x}_i^\star)\right)+\frac{(\mu_{gt}(\mathbf{x}_i^\star)-\mu(\mathbf{x}_i^\star))^2}{2\sigma^2(\mathbf{x}_i^\star)}\right)}{M}}. \quad (23)$$

iv) *Weighted Mean Log Loss* (WMLL): Similar to WRMSE, but weighted over the mean log loss. Gives more importance to error in high pollution areas, taking into account the variance of the predictive model.

$$MLL = \frac{\sum_{i=1}^{M}\frac{\left(-\log p(\mu_{gt}(\mathbf{x}_i^\star)|\mathcal{S},\mathbf{x}_i^\star)\right)\left(\mu_{gt}(\mathbf{x}^*)-\min\mu_{gt}(\mathbf{x})\right)}{\max\mu_{gt}(\mathbf{x})-\min\mu_{gt}(\mathbf{x})}}{M} \quad (24)$$

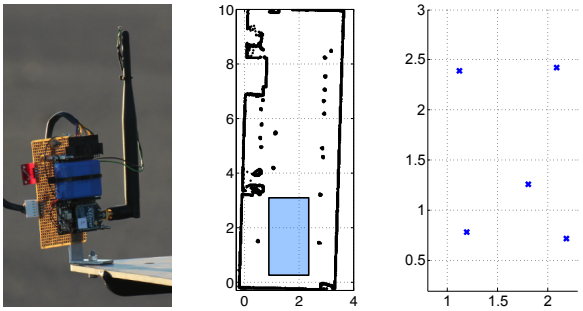
with $\mathbf{x} = (\mathbf{s}; t)$ for the space-time case.

Table I shows the error for each method evaluated w.r.t. the ground truth on a grid over space and time for the entire duration of the experiment. It can be seen that the proposed method (UCBC) delivers the best performance for all indicators. UCBC favours areas of high pollutant concentration, achieving more accuracy over the areas that account for the most relevant component of error. The difference in performance between strategies is remarkable for the weighted errors WRMSE and WMLL. The main reason for this is the extra importance to model areas with hight pollution (exploitative behaviour). For this experiment, UCBC also presents lower error for non weighted metrics, expected when the areas of interest account for the most important component of error. Therefore, when UCBC focuses on sampling from areas of higher concentration it will achieve lower error overall, compared to EC or RC that will model better areas that do not reduce the overall error importantly (because the output variable has lower values for non relevant areas).

It is also noticeable the difference between continuous and discrete sampling strategies. An improvement is revealed for all strategies as we are optimising over continuous paths rather than choosing discrete locations.

TABLE I
RESULTS FOR US OZONE MONITORING

Method	RMSE	WRMSE	LogLoss	WLogLoss
RD	7.3574	2.2324	3.4149	0.0956
ED	7.5225	2.2570	3.4332	0.0957
UCBD	7.1579	1.9764	3.3999	0.0937
RC	7.2238	2.0698	3.3951	0.0935
EC	7.1103	2.2958	3.3907	0.0956
UCBC	**6.7971**	**1.4981**	**3.3537**	**0.0863**



(a) Sensing Board. (b) Experiment area. (c) Position of sensing nodes for ground truth.

Fig. 7. Sensing board, map of the area and location of ground truth measurements. Axis in metres.

B. Luminosity monitoring

A small, wheeled mobile robot was used to monitor dynamic illumination changes in an indoor environment. The goal of this experiment is to compare different techniques for path planning and their impact on the abilities of the robot to learn the space-time patterns of a dynamic phenomenon. The idea is to create a real-world phenomenon under a controlled environment where the dynamics can be adjusted accordingly. Two light sources with variable intensity are dimmed electronically to expose patters with a periodic component and amplitude changes through time.

The robot is equipped with an on-board CPU running ROS[2], an environmental sensing electronic board, shown in Figure 7a, and a laser scanner for localisation in an previously built map. Samples from the phenomenon are gathered

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

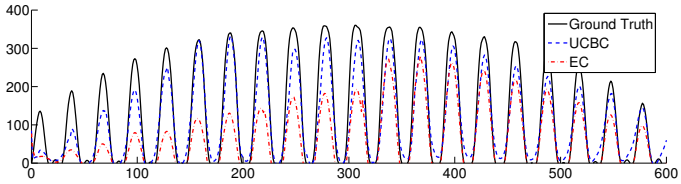


Fig. 9. Light intensity oscillation at location $(x, y) = (1.12, 2.39)$. Horizontal axis represents time in seconds and vertical axis represents intensity with no SI units. The mean of prediction for UCBC and EC is shown according to the legend.

TABLE II
HYPER-PARAMETERS FOR LUMINOSITY MONITORING

σ_n	η	σ_{f1}	l_1	σ_{f2}	l_2	σ_{f3}	γ	σ_{f4}	l_4
12.0	85.0	0.59	29	11.9	8.2	11.2	298	7.6	2

every one second. Ground truth is obtained by placing static sensor boards in five static locations, shown in Figure 7c. Figure 8 shows an interpolation of the measurements in these locations over space for five time stamps within a period. Figure 9 shows the interpolation over time for one source of light located at $(x, y) = (1.12, 2.39)$. Variations in amplitude are noticeable over time and similar to many natural phenomena. Even though the light sources are easily distinguishable for a human observer, the problem is much more complex for a robot that gathers noisy samples from the unknown time-changing phenomenon. The problem becomes even more interesting when the robot needs to decide where to take next samples based on past experience and future reward.

The same path-planning strategies in section V-A are compared in experimental trials that last for ten minutes. The robot used the following covariance function for building the GP model of the phenomenon:

$$k_{sep}\left((\mathbf{s};t),(\mathbf{s};t)'|\boldsymbol{\theta}\right) = k1_{mat3}(\mathbf{s},\mathbf{s}'|\boldsymbol{\theta_1})\cdot \left[(k2_{mat3}(t,t'|\boldsymbol{\theta_2}))\cdot k3_p(t,t'|\boldsymbol{\theta_3})) + k4_{mat3}(t,t'|\boldsymbol{\theta_4})\right], \quad (25)$$

where the estimated hyper-parameters $\boldsymbol{\theta}$ are shown in Table II.

Figure 10 shows the paths travelled by the robot using each technique. Results are similar to the experiment in the previous section. While random sampling strategies, RD and RC, derive paths mostly concentrated at the centre of the studied region, paths for entropy based strategies, ED and EC, are distributed more homogeneously over space. In constrast, UCB paths focus on areas with high luminosity while at the same time exploring the environment for unknown sources of light.

Table III shows numerical results for the evaluation of the performance indicators described in section V-A. Random policies perform close to entropy techniques for this obstacle-free environment. In a case of extremely low cpu availability it can be considered as a viable alternative; however, in real complex environments it is not a promising candidate. UCBC delivers the lowest error and weighted error for all the indicators. It is also shown that sampling over continuous domains results in smaller error for the case of UCB acquisition function. UCB strategies have the smallest error, demonstrating a central advantage in monitoring dynamic phenomena: monitoring areas of higher pollution more intensively results on lower overall error.

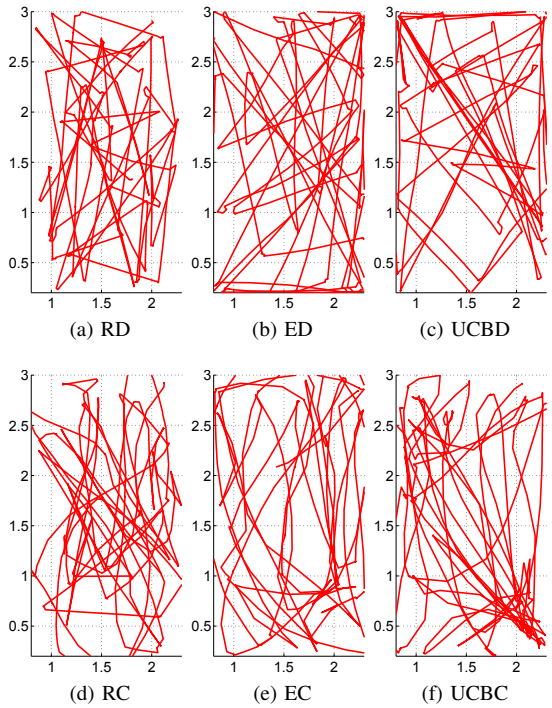


Fig. 10. Resulting paths for six different path planning techniques, axis in metres.

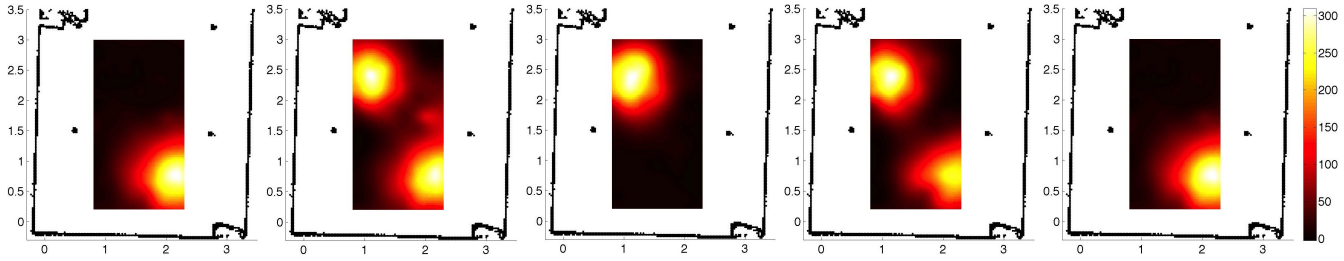
TABLE III
RESULTS FOR LUMINOSITY MONITORING

Method	RMSE	WRMSE	LogLoss	WLogLoss
RD	39.598	28.061	8.949	2.327
ED	38.389	25.106	10.013	2.779
UCBD	38.210	23.715	9.685	2.672
RC	43.390	27.045	10.197	2.754
EC	48.873	38.433	11.980	3.443
UCBC	**30.121**	**21.422**	**8.098**	**2.145**

The developed method runs close to real-time in a standard cpu (i5 processor). Each iteration for finding the next optimal path takes $4.8s$ and can be computed before the execution of the current path is finished, avoiding unwanted pause between consecutive paths.

VI. CONCLUSIONS

This paper proposed a new technique for informative path planning over continuous paths for environmental monitoring. The main contribution is the derivation of a continuous action space strategy by integrating over an acquisition function in a principled Bayesian optimisation framework. We model space-time phenomena using Gaussian processes which enables a robot to learn periodic patterns while preserving spatial correlations between observations. A first layer of BO is used to predict regions of high concentration.



(a) $t = 0$ s (b) $t = 7.5$ s (c) $t = 15$ s (d) $t = 22.5$ s (e) $t = 30$ s

Fig. 8. Spatially distributed light intensity variations. Axis in metres.

Then, a second layer of BO is used to estimate the curve parameters defining the best path to collect new observations.

The proposed method was evaluated in two experimental settings: on a large-scale autonomous monitoring problem for ozone concentration in the US, and for real-time monitoring of changes in luminosity indoors. In both cases, the mobile robot was able to learn a space-time model of the dynamic phenomena. Comparisons were performed between existing techniques for informative path planning indicating that the proposed algorithm captures more accurately the dynamics of areas where the monitored quantity has higher concentration. This ultimately results in an overall more accurate model with lower weighted error. Additionally, our technique can even reduce the total RMSE if the areas of interest account for a significant proportion of the error, as is the case for the two studied situations.

We believe that optimising over curves for path planning can produce more informative decisions achieving longer term rewards. The method explained in this paper can significantly improve the decision making process for efficiently monitoring a wide variety of environmental phenomena. As future work we expect to address long-term or infinite horizon planning with continuous POMDPs and derive a joint procedure possessing the advantages of both Bayesian optimisation and POMDPs.

REFERENCES

[1] 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. arXiv:1012.2599, 2010.

[2] R. Marchant and F. Ramos, "Bayesian Optimisation for Intelligent Environmental Monitoring," in *IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2012.

[3] M. Dunbabin and L. Marques, "Robotics for environmental monitoring," *IEEE Robotics Automation Magazine*, vol. 19, no. 1, pp. 24–39, 2012.

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

[5] 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.

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

[7] S. Vasudevan, F. Ramos, E. Nettleton, and H. Durrant-Whyte, "Gaussian Process Modelling of Large-Scale Terrain," *Journal of Field Robotics*, 2009.

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

[9] S. O'Callaghan and F. Ramos, "Gaussian Process Occupancy Maps," *International Journal of Robotics Research*, 2012.

[10] Y. Shen, A. Ng, and M. Seeger, "Fast Gaussian Process Regression using KD-Trees," in *Neural Information Processing Systems (NIPS)*, 2006.

[11] J. Hensman, N. Fusi, and N. Lawrence, "Gaussian Process for Big Data," in *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2013.

[12] C. Guestrin, A. Krause, and A. P. Singh, "Near-optimal sensor placements in Gaussian processes," in *International Conference on Machine Learning (ICML)*, vol. 1, 2005.

[13] L. A. Bush, B. Williams, and N. Roy, "Computing Exploration Policies via Closed-form Least-Squares Value Iteration," in *International Conference on Automated Planning and Scheduling (ICAPS)*, 2008.

[14] J. J. Chung, N. R. Lawrence, and S. Sukkarieh, "Gaussian Process for Informative Exploration in Reinforcement Learning," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2013.

[15] G. A. Hollinger, B. Englot, F. Hover, U. Mitra, and G. S. Sukhatme, "Uncertainty-driven view planning for underwater inspection," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2012.

[16] S. M. LaValle, "Rapidly-Exploring Random Trees: A New Tool for Path Planning," Tech. Rep., 1998.

[17] X. Lan and M. Schwager, "Planning Periodic Persistent Monitoring Trajectories for Sensing Robots in Gaussian Random Fields," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2013.

[18] G. A. Hollinger and G. S. Sukhatme, "Sampling-based Motion Planning for Robotic Information Gathering," in *Robotics Science and Systems (RSS)*, 2013.

[19] J. Witt and M. Dunbabin, "Go with the Flow: Optimal AUV Path Planning in Coastal Environments," in *Australian Conference on Robotics and Automation (ACRA)*, 2008.

[20] J. Suh and S. Oh, "A Cost-Aware Path Planning Algorithm for Mobile Robots," in *IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2012.

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

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

[23] J. Binney and G. S. Sukhatme, "Branch and Bound for Informative Path Planning," in *IEEE International Conference on Robotics and Automation (ICRA)*, vol. (Submitted), 2011.

[24] N. Cressie and C. K. Wikle, *Statistics for Spatio-Temporal Data*. Wiley, 2011.

[25] D. R. Jones, "A taxonomy of global optimization methods based on response surfaces," *Journal of Global Optimization*, pp. 345–383, 2001.

[26] D. D. Cox and S. John, "A Statistical Method for Global Optimization," in *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1992.

[27] W. Gauschi, *Numerical Analysis*. Springer New York, 2012.