Journal of Statistical Planning and Inference 171 (2016) 115–129

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Bayes linear analysis for Bayesian optimal experimental design

Matthew Jones [a,b,*], Michael Goldstein [a], Philip Jonathan [b], David Randell [b]

[a] Department of Mathematical Sciences, Durham University, Durham, DH1 3LE, United Kingdom

[b] Shell Research Ltd., Concord Business Park, Manchester M22 0RR, United Kingdom

ARTICLE INFO

Article history:
Received 2 April 2015
Received in revised form 5 August 2015
Accepted 26 October 2015
Available online 3 November 2015

Keywords:
Optimal experimental design
Bayes linear analysis
Emulation
Inverse problem
Calibration problem
Probabilistic numerics

ABSTRACT

In many areas of science, models are used to describe attributes of complex systems. These models are generally themselves highly complex functions of their inputs, and can be computationally expensive to evaluate. Often, these models have parameters which must be estimated using data from the real system. In this paper, we address the problem of using prior information supplied by the model, in conjunction with prior beliefs about its parameters, to design the collection of data such that it is optimal for decisions which must be made using posterior beliefs about the model parameters. Optimal design calculations do not generally have a closed form solution, so we propose a Bayes linear analysis to find an approximately optimal design. We motivate the approach by considering optimal specification of measurement locations for remote sensing of airborne species.

1. Introduction and outline

In most quantitative fields of research, models are used to study systems; these models are as diverse in complexity as they are in purpose. For example, in the social sciences, simple linear regression models are a popular tool for exploring relationships between explanatory variables (for example, individual income, or a measure of national development) and response variables (for example, age at death, or national life expectancy). Alternatively, in climate science, complex systems of coupled differential equations with many input parameters are numerically solved to provide a representation of the real climate; see, for example, Williamson and Goldstein (2012).

Despite this diversity, modellers often share a common goal, namely, to use their model, combined with prior beliefs about inputs and data from the system, to infer certain, unobserved inputs to the real system. Where these unobserved inputs are of direct interest, in which case, we refer to an inverse problem (see, for example, Stuart, 2010). Where interest lies in obtaining good information about unobserved inputs only as a means to improve the model's predictions for the real system, we refer to a calibration problem (see, for example, Kennedy and O'Hagan, 2001). In both cases, we refer to the unobserved parameters to be estimated from the data as *model parameters*.

When collecting data from the system, it will typically be possible to control the values of some input parameters when making observations; for example, the climate modeller may choose spatial locations within the Earth's atmosphere at which he makes observations of the real climate. Such inputs are referred to as *design parameters*. In this situation, it is natural to consider the problem of optimal experimental design; given the model, and prior beliefs about its parameters, how should design parameters be set when collecting data, so as to provide the best information about the parameters to be inferred?

* Corresponding author at: Department of Mathematical Sciences, Durham University, Durham, DH1 3LE, United Kingdom. Tel.: +44 7753 997 447.
E-mail address: m.j.jones@durham.ac.uk (M. Jones).

http://dx.doi.org/10.1016/j.jspi.2015.10.011
0378-3758/

The situation is often complicated, however, by inputs to the model (and, therefore, to the data collection process) which may not be controlled, and are themselves not of direct interest, but which are measured when an observation is taken and will affect the inference problem. These inputs will not be known at the point at which design parameters are selected, but modellers will generally have prior beliefs about their possible values. These inputs are referred to as *external variables*. To use the climate example again, while the modeller may be able to select spatial locations for his observations, he will not have control over the state of the atmosphere (eg: pressure or temperature) at these measurement locations, but he will have beliefs about likely values of these quantities based on weather forecasts for the area. In the presence of such unknown inputs, choice of design inputs must take into account all possible external inputs.

Optimal experimental design has been considered by many different authors; Raiffa and Schlaifer (1961), Lindley (1972) and DeGroot (1970) all give comprehensive introductions to the Bayesian treatment of decision theory and optimal experimental design, while Chaloner and Verdinelli (1995) give a review, presenting a number of popular criteria used to find optimal designs. Experimental design calculations can be computationally demanding, and so many authors have also considered ways of approximating them; Huan and Marzouk (2013) evaluate complex models using polynomial chaos expansions to improve computational efficiency of the design calculations, and Muller (1998) gives an overview of simulation methods for optimal design problems. The present article makes two contributions; first, it considers design problems where inference is affected by external variables which cannot be controlled and are unknown at the point when a design is selected, and second, it introduces a novel approximation framework, using Bayes linear emulators, to overcome computational difficulties presented by the optimal design calculations.

The remainder of this article is structured as follows; in Section 2, we illustrate some elements of an optimal design problem using an example in which properties of sources of airborne species (eg: gases and small particulate matter) must be estimated from remotely-observed concentrations. In Section 3, we introduce a notation for general Bayesian optimal experimental design problems, and outline the steps necessary to select an optimal setting of the design parameters; simultaneously, we consider the computational expense of the calculations which must be performed and motivate the need for an approximating framework. As an illustration, we perform optimal design calculations for a linear model.

In Section 4 we propose a general analysis, using Bayes linear emulation, as a suitable approximating procedure. In Section 5, we return to the linear model and source estimation problems and present numerical results generated using our framework. Finally, in Section 6, we discuss our results, and consider further areas of research suggested by our work.

2. Motivating example

We consider a motivating example in which optimal design for data collection is of importance. Hirst et al. (2012) describe a problem in which methane concentration data obtained from an aircraft-mounted sensor is available for a known set of locations and times; the authors wish to use this data to infer the properties (spatial location, emission rate and half-width) of a set of gas sources at ground level, about which they have prior beliefs. This is a typical Bayesian inverse problem; in order to infer these source properties, a model must be chosen to describe the transport of methane from sources (with particular properties) to sensors at particular locations and times, as well as the natural background concentration of methane at the same points. There are many different possible model choices, of varying complexity; for an introduction to the field of atmospheric dispersion modelling, see Stockie (2011). It is clear, though, that under any sensible model, the rate at which gas will be transported in a particular direction will be heavily dependent on the atmospheric conditions, and in particular, on the direction and magnitude of the wind, which we denote collectively by $w = \{w_\theta, w_m\}$.

We consider the optimal design problem presented by such a source estimation problem; for simplicity and clarity of discussion, suppose we have access to a single point sensor which may be fixed anywhere within a certain region and, once placed, will make a number of concentration observations $z = (z_1, \ldots, z_{n_z})^{\mathrm{T}}$ at particular times. The optimal design problem is to find a spatial location $d = \{d_x, d_y, d_h\}$ for the sensor which maximises the value of the observations which it will make; therefore, d is the set of design parameters. The 'value of the observations' depends on the decisions which will be made using posterior beliefs, having collected the observations, and the consequences of these decisions; we postpone detailed consideration of this idea until Section 3, in favour of making some general comments about the different types of parameters in this problem and their roles in the optimal design problem.

Consider the situation in which the region of interest is a cube in three dimensional space, whose vertical sides face North, East, South and West. We seek to infer source parameters $q = (q_x, q_y, q_s)$ (spatial location and emission rate for a single source) at ground level within this cube, constraining our sensor to lie somewhere within it. We follow Hirst et al. in using a Gaussian plume model for gas transport; the plume is illustrated for particular source locations (labelled A and B) in Fig. 1, with some candidate sensor locations also marked (labelled C and D). It is clear from these plots that concentration measurements at a sensor resulting from a particular source depend heavily on wind conditions. For example, in Fig. 1(a), when the wind is blowing from the North-East, the sensor at C measures a large increase in concentration, the majority of which comes from source A, with a small additional contribution from source B; however, the sensor at D receives no contribution from either source. When the wind blows from the South-West, as in Fig. 1(b), D receives a large contribution, whereas C sees nothing. In Fig. 1(c), a wind blowing from the South-East causes no elevated concentrations at either sensor.

Design implications are obvious; consider the situation in which, a priori, we believe the source equally likely to be in any location. Before data collection, if our prior beliefs about the wind indicate that it will blow only from somewhere in the North-Eastern corner, then C is a much better design choice than D, as a sensor in location C will pick up concentration

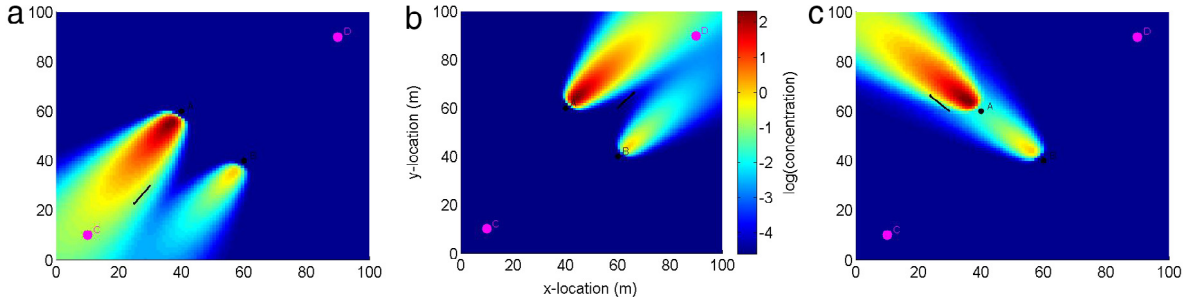


Fig. 1. Showing the log-concentrations for sensors at a height of 1 m generated by a Gaussian plume for a particular wind field. The black dots show the locations of two circular methane sources of radius 1m; the source labelled A is emitting at a rate of 100 kg/h, and the one labelled B at a rate of 20 kg/h. The magenta dots labelled C and D correspond to possible sensor locations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

measurements from sources in a greater range of locations. The opposite is true if our prior beliefs indicate the wind will blow from the South-Western corner, in which case D is a better design choice.

Unfortunately, we will not have wind information until we observe it; however, we can specify our prior beliefs about likely wind speeds and directions. The best that we can do is to construct a design optimal for our prior beliefs about the wind; this is the problem which we will outline, in its most general form, in Section 3.

3. Bayesian optimal design

The example presented in Section 2 is a particular case of a much more general problem in optimal experimental design. Models allow experimenters to specify beliefs about relationships between a set of input parameters and a set of response parameters. Generally, the objective is to use a specified model and set of prior beliefs to infer likely values of some of the inputs; whether these parameters are of interest themselves, or are simply to be used to generate future predictions from the model, the inference problem to be solved is the same, and so we make no further distinction between these two situations.

Once observations of the real system have been made, we wish to make decisions relating to the system using our updated beliefs about its behaviour. These decisions are referred to as *terminal decisions* (Raiffa and Schlaifer, 1961), and will have consequences which may be specified through an appropriate choice of loss function (negative utility function); in any situation we choose the decision with minimum risk (expected loss) under our posterior beliefs; see, for example, Smith (2010).

Considering the inference problem and the subsequent decision problem together, it is natural to design the collection of data to minimise the expected loss from the terminal decision. This is a problem in Bayesian optimal experimental design. First, we set out a notation for the general problem, present the relevant results, and consider the difficulties in performing the necessary calculations.

3.1. General model form and inference

We denote:

- the collection of model parameters to be inferred by $q = \{q_1, \ldots, q_{n_q}\}$.
- the set of design parameters to be selected to collect data by $d = \{d_1, \ldots, d_{n_d}\}$.
- the set of external (or environmental) variables affecting the inference by $w = \{w_1, \ldots, w_{n_w}\}$.
- the observations to be made on the real system by $z = (z_1, \ldots, z_{n_z})^{\mathrm{T}}$.

Additionally, in line with the considerations in Section 2, we make the following assumptions:

- external variables will become known (with no error) by the point at which the terminal decision is made. Before this point, we specify prior beliefs about their values through a conditional probability distribution $p(w|d)$, which may depend on settings of the design parameters
- design parameters are selected before data is collected and before values of the external parameters become known.

The treatment of the external variables in the general problem corresponds to the treatment of the wind speed and direction in the motivating gas sensing problem: this information is measured simultaneously with the observation process, but up until this point, we may only specify our beliefs over likely wind directions based on weather forecasts. The assumption that the external parameters will become known exactly is quite a strong one, and perhaps inappropriate for many applications; however, it simplifies the analysis and so is adopted for pragmatic reasons. We return to the possibility of relaxing this assumption in the discussion.

We assume that we have a model for the real system under study, and we specify a general form for the relationship between our model for the system at a particular set of inputs, $y(q, w, d)$ and the observations of the system as

$$z = y(q, w, d) + \delta(q, w, d)$$

where the values of $y(.)$ and $\delta(.)$ are assumed to be independent for given values of q, w and d, and $\delta(.)$ describes the mis-match between the model and the experimental data. We assume that our model for the system value at a particular parameter setting is a joint Gaussian distribution with mean and covariance matrices depending on the inputs

$$y(q, w, d) \sim \mathcal{N}(m(q, w, d), C(q, w, d)). \tag{1}$$

The specification (1) is sufficiently general to describe a wide class of typical model specifications. Perhaps the simplest which it encapsulates is the linear regression model, where $m(.)$ is linear in the model parameters q and has predictors which depend on the design and external parameters; in this case, for certain prior assumptions on q, this posterior is a Gaussian, significantly simplifying any calculations we wish to do. However, it could also describe a situation in which the full model is deterministic, highly nonlinear and computationally expensive to evaluate, motivating use of a Gaussian process surrogate (see, for example, the oil reservoir example of Craig et al., 1996 or the climate example of Williamson and Goldstein, 2012); in this case, $m(.)$ and $C(.)$ are generated from the Gaussian process posterior. If the model is deterministic, and can be evaluated with low computational expense, then we specify $C(.) = 0$ and use deterministic model predictions directly.

The model mis-match, $\delta(.)$ is assumed to have the form

$$\delta(q, w, d) = Hb + r(q, w, d)$$

where $H = H(q, w, d)$ is a matrix of basis function values, b is a vector of unknown parameters. Prior beliefs about b are specified by $p(b)$, and $r(.)$ is assumed to be a mean-zero Gaussian process with covariance function $k\left((q, w, d), (q', w', d')\right)$. The model mis-match term is used to capture both structured disagreement (for example, systematic errors due to processes which are not modelled) and unstructured disagreement (for example, measurement error) between predictions and observations. In general, the covariance function $k(.,.)$ will depend on a set of parameters (for example, smoothness parameters and marginal variance parameters); for a fully Bayesian treatment, we should specify prior beliefs about these before marginalising them out of the posterior. In practice, this is a computationally difficult calculation, so to simplify the problem, we assume that likely point values can be specified for these parameters, and perform all subsequent analyses conditional on these values.

Specifying prior beliefs about the model parameters through a prior distribution $p(q)$, the form of the posterior distribution can be written as (Kennedy and O'Hagan, 2001)

$$\begin{aligned} p(q, b|z, w, d) &\propto p(z|b, q, w, d)\, p(q)\, p(b) \\ &\propto p(q)\, p(b)\, |C(q, w, d) + K(q, w, d)|^{-1/2} \\ &\quad \exp\left[-\frac{1}{2}(z - m(q, w, d) - Hb)^{\mathrm{T}} (C(q, w, d) + K(q, w, d))^{-1} (z - m(q, w, d) - Hb)\right] \end{aligned}$$

where $K(q, w, d)$ is a matrix of covariance function values for the model mis-match. If, additionally, a Gaussian prior $b \sim \mathcal{N}(\mu_b, \Sigma_b)$ is assumed, integration over model parameter leaves a posterior for q

$$\begin{aligned} p(q|z, w, d) &\propto p(q)\, |C(q, w, d) + K(q, w, d)|^{-1/2}\, |W(q, w, d)|^{1/2} \\ &\exp\left[-\frac{1}{2}\left(z - m(q, w, d) - H\hat{b}(q, w, d)\right)^{\mathrm{T}} (C(q, w, d) + K(q, w, d))^{-1} \left(z - m(q, w, d) - H\hat{b}(q, w, d)\right)\right. \\ &\left. - \frac{1}{2}\left(\hat{b}(q, w, d) - \mu_b\right)^{\mathrm{T}} {\Sigma_b}^{-1} \left(\hat{b}(q, w, d) - \mu_b\right)\right] \end{aligned} \tag{2}$$

where

$$\begin{aligned} W(q, w, d) &= \left[H^{\mathrm{T}} (C(q, w, d) + K(q, w, d))^{-1} H + {\Sigma_b}^{-1}\right]^{-1} \\ \hat{b}(q, w, d) &= W(q, w, d)\left[H^{\mathrm{T}} (C(q, w, d) + K(q, w, d))^{-1} (z - m(q, w, d)) + {\Sigma_b}^{-1}\mu_b\right]. \end{aligned}$$

3.2. Making an optimal decision

We now consider the problem of choosing an optimal decision based on posterior beliefs about model parameters. In order to do so, we introduce additional notation. We denote

- the set of selected terminal decisions by $a = (a_1, \ldots, a_{n_a})^{\mathrm{T}}$, and specify that $a \in \mathcal{A}$ for some space of possible decisions
- by $L(a, q)$ the loss function which describes the consequences of selecting decisions a and then realising model parameter values q.

Then, our *risk* from taking measurements z at design parameter settings d, under external conditions w, and then making terminal decisions a is defined as the expectation of the loss under the posterior for the model parameters generated under these conditions (DeGroot, 1970)

$$\rho\left[a, z, w, d\right] = \int L\left(a, q\right) p\left(q|z, w, d\right) dq. \tag{3}$$

Our objective is to take the decision with minimum risk, based on current beliefs about q; we define the risk from an optimal terminal decision under these conditions as the minimum risk over possible $a \in \mathcal{A}$

$$\rho\left[z, w, d\right] = \min_{a \in \mathcal{A}} \rho\left[a, z, w, d\right]. \tag{4}$$

Using the expression (2) for the model parameter posterior, we see that, in general, no analytic solution exists for the integral in (3) (though for some particular model forms and loss function choices, closed-form solutions exist; see, for example, Chaloner and Verdinelli, 1995). We resort to numerical integration techniques, such as MCMC Robert and Casella (1999) or Bayesian quadrature (O'Hagan, 1991), to evaluate (3), introducing a potentially large computational expense, particularly for models with large numbers of parameters to be inferred. For general loss functions, the minimisation in (4) can also be a challenging numerical optimisation problem, particularly for continuous, high dimensional a (Chaloner and Verdinelli, 1995 give some special examples where this optimisation step is implicit or where the minimum can be found analytically).

3.3. Choosing an optimal design

We seek a design which minimises the risk from an optimal terminal decision over all possible external conditions and observed data sets. We must therefore find the expected behaviour of the risk (4) over our distributions for w and z

$$\rho\left[d\right] = \iiint \rho\left[z, w, d\right] p\left(z \mid q, w, d\right) p\left(q\right) p\left(w \mid d\right) dq dz dw. \tag{5}$$

The optimal design is that which minimises this risk

$$d^* = \arg\min_d \rho\left[d\right]. \tag{6}$$

For general problems, none of the expectations in (5) can be computed analytically; to perform the optimal design calculation numerically, we would need to evaluate the integral (3) over model parameters and the optimisation (4) over decisions for each evaluation of the risk $\rho\left[z, w, d\right]$ needed to perform the (potentially high-dimensional) integrals in (5). We conclude that the numerical solution of the optimal design problem is not possible in general.

3.4. Example: linear model

We now illustrate the optimal experimental design process using a simple example: the Bayesian linear model. In this instance, we wish to estimate q such that

$$z = y\left(q, w, d\right) + \delta\left(q, w, d\right) = X\left(w, d\right) q + \delta.$$

We assume that $\delta \sim \mathcal{N}\left(0, \sigma^2 I\right)$ (noise is independent for each observation, with constant variance), and also assign a Gaussian prior distribution $q \sim \mathcal{N}\left(m_t, V_t\right)$. Because of the linearity of the mean in the model parameters and the Gaussian prior and likelihood, the posterior (2) will also be a Gaussian distribution (Bishop, 2006)

$$p\left(q|z, w, d\right) = \mathcal{N}\left(\hat{m}_t\left(z, w, d\right), \hat{V}_t\left(w, d\right)\right)$$

with parameters

$$\hat{V}_t\left(w, d\right) = \left[V_t^{-1} + \frac{1}{\sigma^2} X\left(w, d\right)^{\mathrm{T}} X\left(w, d\right)\right]^{-1} \qquad \hat{m}_t\left(z, w, d\right) = \hat{V}_t\left(w, d\right)\left[V_t^{-1} m_t + \frac{1}{\sigma^2} X\left(w, d\right)^{\mathrm{T}} z\right].$$

We choose a quadratic loss function, $L\left(a, q\right) = \|a - q\|^2$, to describe the cost of failing to estimate the correct value. It is well known that the choice of estimator which solves the optimisation problem (4) in this case is $a = \hat{m}_t\left(z, w, d\right)$, and that with this choice, we may obtain a closed-form expression for (4) as the trace of the posterior variance

$$\rho\left[z, w, d\right] = \rho\left[w, d\right] = \mathrm{tr}\left[\hat{V}_t\left(w, d\right)\right]. \tag{7}$$

Note that, in this particular case, the posterior variance, and therefore the risk from an optimal terminal decision, do not depend on the value of the measurements z.

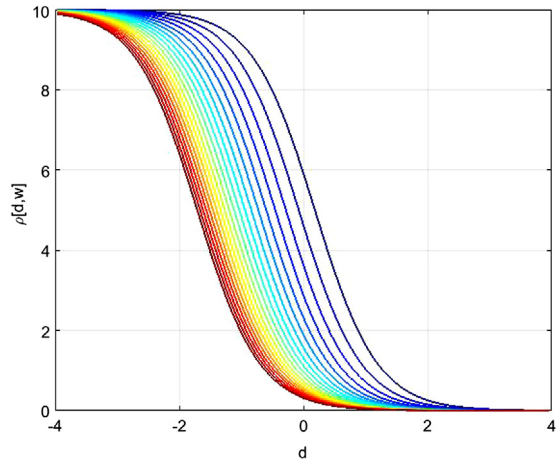


Fig. 2. Plot of the risk (7) for the linear model in Section 3 using design matrix (8), from a single observation z at design setting d. Colours correspond to different values of the environmental parameter w, spaced uniformly in the interval [0.5, 4]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Consider a simple example, in which there is only one model parameter to be estimated, and we make only one observation; we choose a design matrix with only one component:

$$X(w, d) = w \exp(d). \tag{8}$$

Our measurement error is assumed to have variance $\sigma^2 = 4$ and the prior parameters for q are set to $m_t = 0$, $V_t = 10$. Fig. 2 shows plots of the risk (7) as a function of d for a range of different values of w. Considering the form of (7), we expect that, whatever our beliefs about w and q, the optimal design would be to choose the greatest possible value for d, since this maximises the magnitude of the linear predictor, thus maximising the amount of information available about the model parameter; inspection of Fig. 2 demonstrates that our intuition is correct.

In order to find the optimal setting of the parameters d, we must now compute the expectation of this expression over the external parameters w (as in (5)), before minimising over d (as in (6)). This is the point at which analytic expressions for the solution cease to be available, and at which we must resort to approximations. We will return to this example in Section 5; the availability of an analytic expression for the risk from an optimal terminal decision, (7), makes it simple to evaluate the results of the approximation methods we will develop in Section 4 by comparison with risk values obtained by 'brute force' numerical integration.

4. Bayesian analysis of the optimal design calculations

A Bayes linear emulation framework is proposed to approximate the optimal design calculation; this preserves the structure of the design problem and makes some simplifying assumptions which make the calculation easier to handle, while leaving scope for further problem-specific simplifying assumptions that may be appropriate.

4.1. Bayes linear analysis

In a Bayes linear analysis, we describe our prior beliefs about a set of quantities through a full second-order specification (expectations, variances and covariances). On learning the values of some quantities, we adjust beliefs about the remainder by linear fitting (Goldstein and Wooff, 2007). Consider two collections of quantities, B and D, and suppose that we specify $\mathrm{E}[B]$, $\mathrm{E}[D]$, $\mathrm{Var}[B]$, $\mathrm{Var}[D]$ and $\mathrm{Cov}[B, D]$: then, if we learn the values of collection D, our adjusted expectation and variance for B can be obtained as

$$\mathrm{E}_D[B] = \mathrm{E}[B] + \mathrm{Cov}[B, D]\,\mathrm{Var}[D]^{-1}(D - \mathrm{E}[D])$$
$$\mathrm{Var}_D[B] = \mathrm{Var}[B] - \mathrm{Cov}[B, D]\,\mathrm{Var}[D]^{-1}\,\mathrm{Cov}[D, B].$$

Bayes linear analysis for functional data (referred to as Bayes linear emulation) has been performed by a number of authors (for example, Goldstein and Rougier, 2006 and Cumming and Goldstein, 2009); we choose a Bayes linear specification, rather than a fully probabilistic one, for two reasons:

- when emulating a risk function, it is unclear what (if any) distributional assumption may be appropriate for the function output at a particular input setting. A Bayes linear analysis avoids making a possibly inappropriate distributional assumption.

- to find a space of candidate optimal designs, we later use the emulator to rule out design inputs which are unlikely to produce a minimum of the risk. Under a fully Bayesian specification, such an analysis would introduce a series of highly complicated, non-linear constraints, which would be hard to deal with even through sampling methods; however, a Bayes linear specification allows us to handle these constraints easily. Bayes linear emulators have previously been successfully applied to the closely related problem of history matching in a wide variety of fields; see, for example, Vernon et al. (2010).

4.2. Emulating the risk

We fit an emulator to the risk from an optimal terminal decision, (4), of the following form

$$\rho_q[z, w, d] = f(q, w, d) + \xi_z.$$

Here, $\rho_q[z, w, d]$ denotes the conditional risk obtained by evaluating (4) using data z generated from $p(z|q, w, d)$ at a known value of model parameters q. Approximation of the risk by the sum of an emulator $f(q, w, d)$ and a zero-mean, uncorrelated process ξ_z assumes that any variability in the risk caused by variation in z for a given value of q is small in comparison to other sources of variation. When collecting risk data to update the emulator, multiple evaluations should be obtained at some fixed values of q, w and d to assess the size of this variability, and to test whether this is a reasonable assumption.

Where this is a reasonable assumption, it is a useful one, as it reduces the complexity of the integrals in (5); typically, we want to collect a large number of observations at a particular design setting to infer a relatively low-dimensional q. If we can assume that z only causes small variation in the risk for given q, we may emulate the risk as a function of q and eliminate the high-dimensional integral over data in (5), leaving only lower-dimensional integrals over w and q.

It may further be helpful to reduce the dimensionality of the set of external variables w used as inputs to the emulator $f(.)$ in a similar way, by using only a subset of all external variables as inputs and absorbing any variation caused by the rest into an additional residual term. The effect of any such simplification must be also checked to ensure that the majority of variation in the risk is still captured by the emulator.

We further assume that $f(.)$ is composed of a mean process, with basis functions $g() = (g_1(.), \ldots, g_M(.))^{\mathrm{T}}$ and parameters $\beta_{\cdot} = (\beta_1, \ldots, \beta_M)^{\mathrm{T}}$ which we estimate, and a weakly stationary residual process $u(.)$, so that

$$\rho_q[z, w, d] = g(q, w, d)^{\mathrm{T}} \beta + u(q, w, d) + \xi_z.$$

We complete our prior by specifying $\mathrm{E}[\beta]$ and $\mathrm{Var}[\beta]$, and by specifying an appropriate positive definite covariance function for $u(.)$, so that

$$\mathrm{Cov}\left[u(q, w, d), u\left(q', w', d'\right)\right] = c\left((q, w, d), (q', w', d')\right).$$

We now adjust beliefs using data obtained by evaluating the risk (4) using data sets z generated from $p(z|q, w, d)$ at some known design of N input points in (q, w, d). We denote the chosen set of design inputs by $\{q^{(j)}, z^{(j)}, w^{(j)}, d^{(j)}\}$, and denote the value of the risk obtained at this input by

$$F_j = \rho\left[z^{(j)}, w^{(j)}, d^{(j)}\right].$$

Risk values are obtained by evaluating (4) using an appropriate numerical integration technique. We do not consider the selection of a set of appropriate design points here, except to say that points should be selected to give good coverage of the input space; see Santner et al. (2003) for a discussion of some common design choices.

Having obtained the risk values $F = (F_1, \ldots, F_N)$, we select appropriate forms for the basis and covariance functions, select any necessary covariance function parameters by cross-validation and compute posterior beliefs as updated means and covariances $\mathrm{E}_F[f(q, w, d)]$ and $\mathrm{Cov}_F\left[f(q, w, d), f\left(q', w', d'\right)\right]$; details of this update are deferred to Appendix A.

In much of the previous work in this area, attention focuses on generating optimal designs for data collection to fit emulators for complex computer models (see, for example, Bates et al., 1996); in this work, we instead perform an additional layer of modelling by fitting an emulator to data from the risk calculation. This approach is advantageous in that it allows approximation of optimal design calculations for a much broader class of models, and in that it allows us to reduce the complexity of some of the necessary calculations by making simplifying assumptions for our risk emulators. Emulation of decision calculations has been reported by Williamson and Goldstein (2012), who examine a sequential policy-making problem in climate science; here, we examine the related, but non-sequential, problem of optimal design, integrating the risk emulator directly to handle environmental parameters which we may not control, and to handle the variability in the experimental data. Future work will involve extension to sequential problems: see Section 6 for details.

4.3. Approximation of expected risk

Once a suitable emulator for the risk from an optimal terminal decision has been constructed, we use this as a surrogate for the true risk function and approximate the expected risk (5) by computing the expectation of the emulator. This strategy was referred to as Bayesian quadrature by O'Hagan (1991), and as Bayesian Monte Carlo by Rasmussen and Ghahramani (2003).

We must only compute the expectation of our emulator over model parameters and environmental parameters, since it no longer depends on the observed data values

$$f(d) = \iint f(q, w, d)\, p(q)\, p(w|d)\, dq dw.$$

By integrating the adjusted mean and covariance functions appropriately, we obtain adjusted mean and covariance functions $\mathrm{E}_F[f(d)]$ and $\mathrm{Cov}_F\left[f(d), f\left(d'\right)\right]$ for the integrated quantity; details of the calculation are presented in Appendix A. Care must be taken when constructing the original emulator to ensure that relevant components of the adjusted mean and covariance can be integrated easily, without the need for computationally expensive numerical integration; in practice this is usually achieved using a squared exponential covariance function which is separable in its arguments and a set of polynomial or Fourier basis functions; for a detailed discussion, see Williamson (2010).

4.4. Identifying potentially optimal designs

Having approximated the expected risk from a particular design, we use this to approximately solve the optimisation problem (6). Finding the global minimum of a complex function is challenging, and has itself generated an extensive literature; emulation has previously been used to find approximate solutions. Hennig and Schuler (2012) demonstrate the difficulty of the problem by writing the probability distribution for the minimising argument of a Gaussian process, comprising an integral over an infinite product, before proposing an algorithm to generate approximate solutions, and Jones et al. (1998) use standard optimisation techniques to minimise Gaussian process surrogates for expensive functions and propose some criteria for identifying when a minimum has been found.

We propose an alternative strategy of sequentially eliminating parts of the design input space unlikely to produce a minimum of the risk. This allows us to report a space of 'non-implausible' design inputs which, after this analysis, have still not been ruled out as unlikely to be minima. This form of analysis is similar to history matching (Vernon et al., 2010; Craig et al., 1996), where emulators for computer models of complex physical systems are compared to observations of the real system in order to rule out parts of the input space which could not have produced acceptable matches to the data.

We first minimise the mean function of our emulator using standard numerical optimisation techniques, to obtain an estimate $\hat{d}$ for the location of the minimum of the risk. We then obtain a more careful estimate of the risk by re-emulating (or sampling) the risk with $d = \hat{d}$ fixed for all inputs, and then computing the expectation of this emulator as described in the previous section (or computing the sample mean and variance), to obtain an estimate $\hat{f}$, with variance $\hat{V}$. These more accurate risk estimates are used to define an implausibility measure against which we screen design input settings.

Our choice of implausibility measure is given by Eq. (9); for each candidate set of design parameters d, we compare the predicted risk from the emulator at d with the more accurate evaluation of the risk at the minimum, $\hat{f}$ in terms of mean and standard deviation predictions. If sets of three-standard deviation error bars at the minimum and the candidate point do not overlap, and the candidate has a greater mean risk than the existing minimum estimate, then the candidate is ruled out, since it is unlikely to generate a minimum of the risk.

As in history matching, we rule out parts of the design parameter space in waves. If the 'non-implausible' space defined by the first implausibility measure is not judged to be small enough, or is not judged to have small enough risk, then we perform another wave of space reduction: we re-emulate the risk within the non-implausible space defined by the first implausibility measure, using newly generated data and prior beliefs derived from the emulator fit at the previous wave, calculate its expectation, minimise, emulate the risk at the minimum more carefully, and then compare input points with the implausibility measures defined at both stages to find a further reduced non-implausible space. Further waves may be performed in the same way, until no further reduction in the size of the space is observed, or until we are satisfied that the space has been reduced sufficiently.

We iterate as follows for waves $j = 1, \ldots$:

- make a second-order prior specification, obtain data F_j and adjust beliefs as described in Section 4.2 and Appendix A.
- calculate the expectation of the emulator over all relevant inputs, as described in Section 4.3 and Appendix A.
- find design input $\hat{d}_j$ minimising $\mathrm{E}_{F_j}[f(d)]$ using numerical optimisation techniques.
- re-emulate or sample the risk from an optimal terminal decision at $\hat{d}_j$, to obtain a more careful estimate $\hat{f}_j$, with variance $\hat{V}_j$.
- define the implausibility measure at wave j by

$$I_j(d) = \frac{\hat{f}_j - \mathrm{E}_{F_j}[f(d)]}{\hat{V}_j^{1/2} + \mathrm{Var}_{F_j}\left[f\left(\hat{d}\right)\right]^{1/2}} \tag{9}$$

- define the remaining non-implausible space after wave j to be those d for which $I_k(d) \leq 3$ for $k = 1, \ldots, j$.

Elimination of points with implausibility measure greater than 3 is motivated by the work of Pukelsheim (1994), who proved that for any random variable with unimodal density function, 95% of the probability mass is concentrated within 3 standard deviations of the mean. Therefore, if the standardised distance (9) exceeds 3, we can be confident that this input d is unlikely to have produced a minimum of the risk.

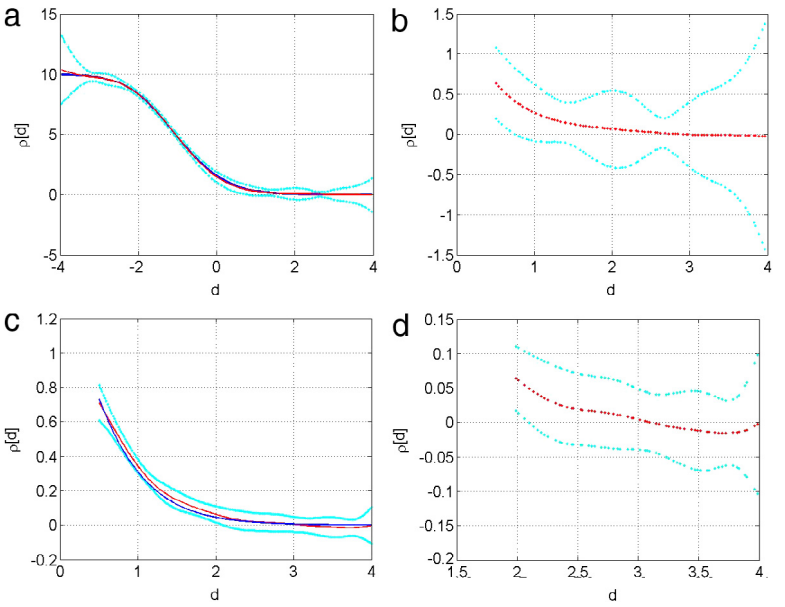


Fig. 3. Plots of the design space reduction process for the linear model example discussed in 5; red (mean) and cyan (2 std. dev. error bars) indicate predictions from the Bayes linear emulator, whereas blue indicates results obtained by 'brute force' numerical integration of the risk (7) over w. Fig. 3(a) shows the emulator at wave 1, fitted to samples drawn from risks such as those in Fig. 2 and integrated over w; 3(b) shows a set of d values sampled from the non-implausible space defined by this emulator; Fig. 3(c) then shows the wave 2 emulator fitted to this non-implausible space using data sampled at the design points in 3(b); Fig. 3(d) shows the risk at a set of values of d sampled from the non-implausible space defined by both the wave 1 and wave 2 emulators (3(a) and (c)). Note that the size of the remaining design space and the upper bound for the minimum risk both decrease as we perform each wave. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5. Numerical results

5.1. Linear model

We now return to the linear model example of Section 3 and apply the approximation framework developed in Section 4 to generate approximately optimal designs. Earlier, we saw that, in the case of the linear model, a closed-form expression, (7), was available for the risk from an optimal terminal decision which did not depend on the values of the observations. This motivates the following, simplified form for the Bayes linear emulator for the risk

$$\rho\left[z, w, d\right] = \rho\left[w, d\right] = f\left(w, d\right).$$

We specify that the design and external parameters are restricted to lie in the ranges $d \in [-4, 4]$ and $w \in [0.5, 4]$, and assume that the external parameter is independent of the design input, with $p\left(w\right) = \mathcal{U}\left([0.5, 4]\right)$.

We now proceed as described in Section 4, generating data from (7) using a Latin hypercube design of 100 points in d and w, using this to update beliefs, and calculating the expectation of the emulator over the distribution of w, $p\left(w\right)$. The chosen emulator consists only of a residual process with a squared exponential covariance function, with marginal variance and correlation lengths determined by cross validation; the fitted emulator was also tested against an additional validation data set, the members of which were within 3 standard deviations of the fitted mean. We then optimise the mean function of the emulator and sample a set of design input values which lie in the non-implausible space defined by the emulator and its minimum. These two steps constitute one wave of space reduction, and are depicted in Fig. 3(a) and (b).

Having considered the expected risk for the non-implausible space remaining after the first wave, we decide that further improvement in the risk is desirable, and so perform another wave of space reduction; we fit a second emulator to the remaining non-implausible space (shown in Fig. 3(c)), and calculate its expectation over $p\left(w\right)$. Optimising this emulator, and sampling points not ruled out by the implausibility measure from either emulator, we obtain the 100 samples of d shown in Fig. 3(d). On consideration of the possible risks for these designs, we decide that we have reduced the risk enough, and that no further waves are necessary; the design space was reduced by about 75% overall. An emulator fitted to the remaining non-implausible space is shown in Fig. 4; optimising this emulator indicates that we should take the single available sample at $d = 4$, agreeing with the intuition developed in Section 3.

5.2. Gas sensing problem

We return to the gas sensing problem introduced in Section 2, describing the forward model used therein in more detail, and using the methods introduced in Section 4 to find spaces of candidate designs. Suppose we seek to select a three

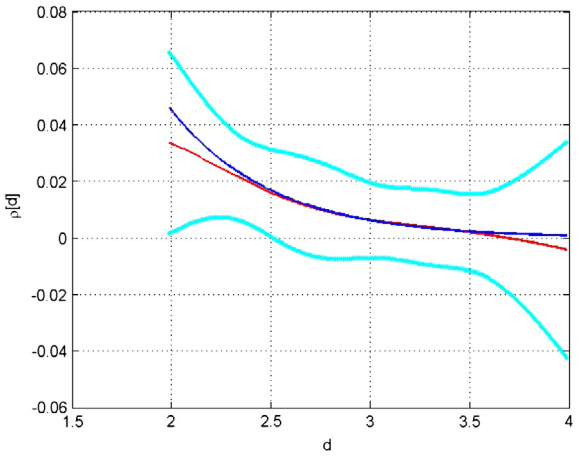


Fig. 4. Emulator fitted to the non implausible space remaining after the second wave of history matching, using data generated at design parameter values sampled in 3(d). Optimising this emulator leads to the conclusion that we should take the single available sample at $d = 4$, which agrees with intuition developed in Section 3 about the location of the optimal design input setting for this problem.

dimensional location for a single point-sensor within a particular space, to be used to collect a set of n_z observations with which to infer the location of a single gas source; we therefore write our design parameters as $d = \{d_x, d_y, d_h\}$, lying in a box $\mathcal{B} = \{d : d_x \in [0, \mathcal{L}_x], d_y \in [0, \mathcal{L}_y], d_z, \in [0, \mathcal{L}_h]\}$. We assume that the wind direction for each observation varies randomly around a prevailing direction, itself a random variable with specified prior distribution. We parametrise the model in terms of this single prevailing direction (we assume that the wind is the same everywhere in the box and fix the wind speed to be the same for all observations). The model parameters q are the two-dimensional location (q_x, q_y) and emission rate q_s of a single source; $q = \{q_x, q_y, q_s\}$.

We first consider an example in which we must choose a single point sensor location under a narrow range of different possible wind directions, to show that our method can generate results consistent with the intuition developed in Section 2; we then present an example in which three point sensors must be chosen to collect data under a wide range of wind directions, to demonstrate good performance in higher dimensional problems.

5.3. One-sensor example

Forward model. We follow Hirst et al. (2012) in using the following forward model

$$z = A\,(d, w, q)\,q_s + b + \epsilon_z$$

where $A\,(d, w, q)$ is a vector of coupling constants calculated under the plume model depicted in Fig. 1, and b is a vector of background concentration levels, to which we assign a prior multivariate Gaussian distribution. The model used by Hirst et al. contains other atmospheric parameters, such as wind speeds and plume opening angles; for simplicity, we fix these to a set of likely values. As the background levels are not of direct interest to us, we treat them as the set of model mis-match parameters in Eq. (2) and integrate them out of the posterior. The errors ϵ_z are assumed to be independent zero-mean Gaussian variables, with standard deviation $\sigma_z = 1$.

We assume a truncated-Gaussian prior ($\mathcal{TN}\,(\mu, \sigma, a, b)$: mean μ, standard deviation σ, lower truncation a and upper truncation b) for each model parameter, with the source lying anywhere at ground level within the box ($q_x \sim \mathcal{TN}\,(\mu_x, \sigma_x, 0, \mathcal{L}_x)$ and $q_y \sim \mathcal{TN}\left(\mu_y, \sigma_y, 0, \mathcal{L}_y\right)$ independently), and source strengths $q_s \sim \mathcal{TN}\,(\mu_s, \sigma_s, 0, \mathcal{L}_s)$. Wind directions are assumed to be independently uniformly distributed around the North-Eastern quadrant for each observation (i.e.: $w_i \sim \mathcal{U}\,([0, \pi/2])$, adopting mathematical convention for directional specification). Interest lies in obtaining accurate estimates of the parameters, and so the loss function used is the weighted quadratic

$$L\,(a, q) = \alpha_x(q_x - a_x)^2 + \alpha_y(q_y - a_y)^2 + \alpha_s(q_s - a_s)^2$$

where the $\alpha_.$ terms regulate the relative importance of obtaining accurate estimates of location components and emission rate. For the remainder of this example, we set $\alpha_x = \alpha_y = 1$ and $\alpha_s = 0.1$, so we are more sensitive to the quality of the estimate for spatial source location than its emission rate.

Emulation. We design for a problem in which $n_z = 100$ observations are to be made, and in which $\mathcal{L}_x = 100$ m, $\mathcal{L}_y = 100$ m, $\mathcal{L}_h = 3$ m and $\mathcal{L}_s = 0.05$ m^3/s. First, a 1000-point Latin hypercube in $\{q, d\}$ was generated, a random prevailing wind and corresponding set of wind directions was sampled for each, and a random set of observations was sampled from $p\,(z|q, w, d)$. $\rho_q\,[z, w, d]$ was then approximated for each of these design points using a Metropolis–Hastings sampler, and the following emulator was fitted to the samples

$$\rho_q\,[z, w, d] = f\,(q, w, d) + \xi_{w'} + \xi_z.$$

The parameter $\xi_{w'}$ is an additional zero-mean, uncorrelated process introduced to capture additional variability in $\rho_q[.]$ caused by variation in wind direction for each observation around the prevailing direction w. This assumes that any variation in the risk induced by variation in the wind around its prevailing direction has zero mean and relatively small variance; this assumption was checked by sampling the risk at multiple different realisations of the full wind field and found to be reasonable. We do not attempt to distinguish between in ξ_z and ξ_w, choosing to specify a level of variability which captures both sources of variability.

Emulators fitted to the risk data were chosen to have a zero mean function and squared exponential covariance functions; correlation lengths and marginal variances were selected based on discrepancy measures calculated through cross-validation. The emulator is fitted with 20 points excluded from the data, and then is used to predict the left out-points; the standardised distances from the mean function to the points is computed, and the covariance function parameters are adjusted manually to ensure that all points are within 3 standard deviations of the mean. Emulators are then additionally tested through prediction of a validation set of 50 new points. Details of the mean and covariance functions used are given in Appendix B, and details of the parameters chosen and descriptions of the quality of the emulator fit are provided below for each wave. Fig. 5 shows the emulator fits, and the sequential reduction of the space of designs.

Wave 1. The mean of the first-wave emulator for a fixed sensor location is shown in Fig. 5(a), with prevailing wind blowing from the North-East. For the covariance function, the marginal variance was chosen to be $\theta = 200^2$, and correlation lengths were fixed at $r_x = r_y = 110$, $r_h = 1.2$, $r_w = 0.1$ and $r_s = 0.01$; the set of validation points were all within 3 standard deviations of the emulator mean, so this emulator was deemed a good fit. The expectation of this emulator over q and w was then computed; the mean function of the integrated emulator is shown in Fig. 5(b). The minimising argument $\hat{d}_1 = (6.06, 0, 0.80)$ of the mean function was found using numerical optimisation; the resulting mean function is plotted in Fig. 5(b). The risk was then sampled for multiple values of q and w, with $d = \hat{d}_1$ fixed, to give a more careful estimate of the minimum of the risk, $\hat{f}_1 = 4.77$. This was then used to define an implausibility measure, which was used to rule out space as shown in Fig. 5(c).

Wave 2. An emulator was then fitted to a design of points generated to lie within the remaining non-implausible space; its mean function is shown in Fig. 5(d). This time, the covariance function was chosen to have parameters $\theta = 100^2$, $r_x = r_y = 100$, $r_h = 1$, $r_w = 0.1$ and $r_s = 0.01$; all of the validation points were within 3 standard deviations of the mean. After computing its expectation (the mean function of which is shown in Fig. 5(e)), minimising (to obtain $\hat{d}_2 = (0, 0, 0.82)$) and sampling the minimum risk value more carefully (to obtain $\hat{f}_2 = 3.88$), the remaining non-implausible space was as shown in Fig. 5(f). No further waves were performed, since the space has already been significantly reduced, and little reduction in the size of the space was achieved from wave 1 to wave 2; we therefore place the sensor at $\hat{d}_2$. Note that this result agrees with the intuition developed in Section 2; that, for winds likely to come from a direction somewhere in the North-Eastern quadrant, most information is obtained with a sensor in the South-Western corner.

5.4. Three-sensor example

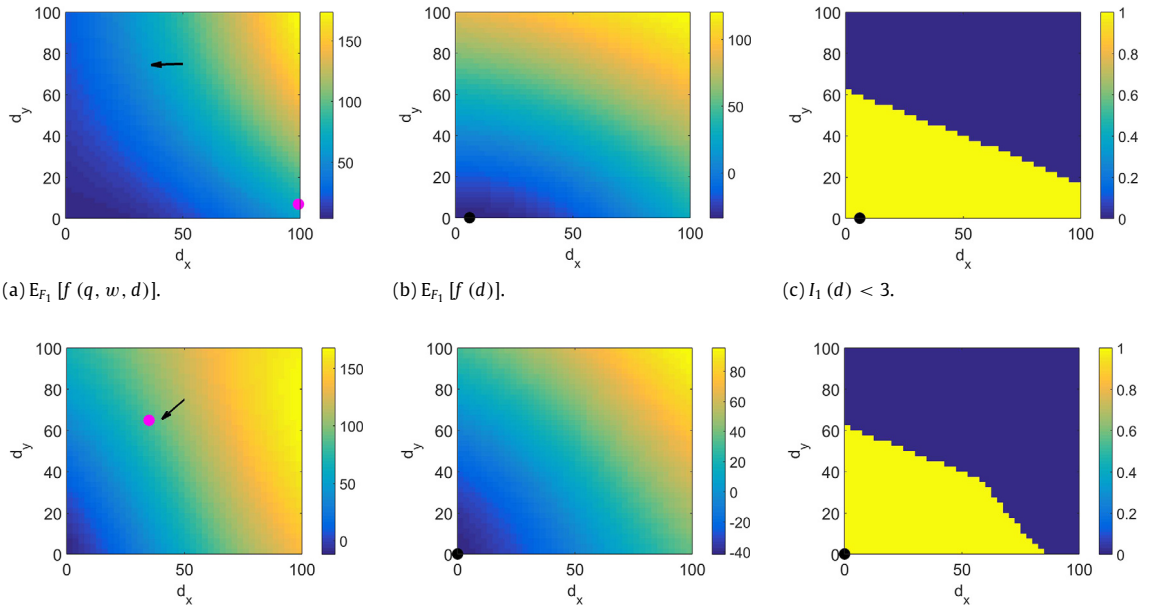
We now design for a problem in which we must optimally select locations for three point sensors; we assume that each of the sensors will collect $n_z = 200$ observations, and that each can be placed anywhere inside the same $100 \times 100 \times 3$ m box. The forward model to be used remains exactly as in the one-sensor example. We now assume, however, that the source location can be outside of the box constraining the beams, specifying that $q_x \in [-50, 150]$, $q_y \in [-50, 150]$, and assume a smaller maximum emission rate, with $q_s \in [0, 0.01]$. We also reduce the precision of the sensors, specifying that $\sigma_z = 10$ for each.

Additionally, we modify the range of possible wind conditions for each observation, specifying that $w_i \sim \mathcal{U}([0, 2\pi])$ independently for each observation. The same quadratic loss is used, but the weightings of the components of the loss function are set to $\alpha_x = \alpha_y = 1$ and $\alpha_s = 10^8$, so that in this example, our risk depends as heavily on the posterior variance of the emission rate as it does on the location components. All other parameters remain fixed as in the one-sensor case.

Emulation. The emulation procedure for the three-sensor problem was similar to that for the one-sensor problem; a 500-point Latin hypercube design was chosen in the design variables and model parameters, and data from the risk function was obtained through sampling the posterior using a Metropolis–Hastings sampler. Note that in this instance, we do not emulate the risk as a function of the environmental parameter, as the large variability in the wind direction, coupled with the large number of observations, ensures that we will receive concentration information from all directions under every possible sensor configuration, meaning that the wind direction causes only small variation in the risk.

The covariance function chosen for the one point example is extended to take account of the use of multiple sensors; further detail is given in Appendix B. Again, a zero mean function is used. The emulator parameters were chosen by manual validation against 20 freshly-generated risk values, and the performance against these diagnostics was good (none lying outside three-standard-deviation error bars).

The results of the first wave of emulation are shown in Fig. 6. Fig. 6(a) shows a sample from the emulator fitted to the risk data, with the locations of the source and two of the sensors fixed. For a given source location, the risk appears to be minimised when the sensor is close to the source, but separated from the other sensors. Fig. 6(b) shows the expected risk for the same fixed sensor locations, which is again smallest when the third sensor is separate from the others. Fig. 6(c)



(a) $\mathrm{E}_{F_1}[f(q, w, d)]$. (b) $\mathrm{E}_{F_1}[f(d)]$. (c) $I_1(d) < 3$.

(d) $\mathrm{E}_{F_2}[f(q, w, d)]$. (e) $\mathrm{E}_{F_2}[f(q, w, d)]$. (f) $(I_1(d) < 3)$ and $I_2(d) < 3$.

Fig. 5. Waves of space reduction for gas source sensing example in Section 5. 5(a) shows the mean of the original emulator, fitted to 1000 samples of the risk from an optimal terminal decision, as a function of sensor location, with source fixed at the magenta dot location (and $q_s = 0.02\ \mathrm{kg/m^3}$), with wind blowing from the North-East; Fig. 5(b) shows the expectation of this mean function over source parameters and winds, with (x, y) location of the numerical minimum marked with a black dot; 5(c) shows the non-implausible space defined by the integrated emulator (with yellow corresponding to parts of the space not yet ruled out); Fig. 5(d)–(f) show the same quantities for the emulator fitted at the second wave. All plots generated with $d_h = 1$ m. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

shows the non-implausible space induced by this emulator, for the same fixed sensor locations; under these placements, for only a small number of locations for the third sensor does the configuration lie in the non-implausible space. Finally, Fig. 6(d) shows the non-implausible space as a function of one of the sensor locations, with the other two sensors fixed at their numerically-obtained minima; with two of the sources fixed to their optimal locations, the risk is sufficiently flat that all possible placements for the third sensor lie within the non-implausible space. Qualitatively, good designs are obtained by separating the sensors as far as possible within the box.

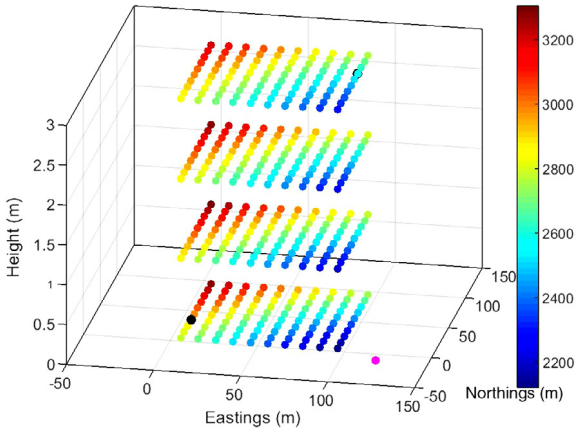
After numerical minimisation, and re-emulation of the risk value at the minimum, sampling using a 1000 point Latin hypercube indicates that the remaining non-implausible space is approximately 1.4% of the size of the original space; the value of the risk at the minimum was 2472 and the standard deviation 40.2. For the purposes of illustration, we cease emulation at this point, as we have substantially reduced the space of possible designs; however, if it was deemed that there was too much uncertainty around the value of the risk within the non-implausible space to make a decision, then re-emulation inside the remaining space may produce further reductions.

6. Discussion

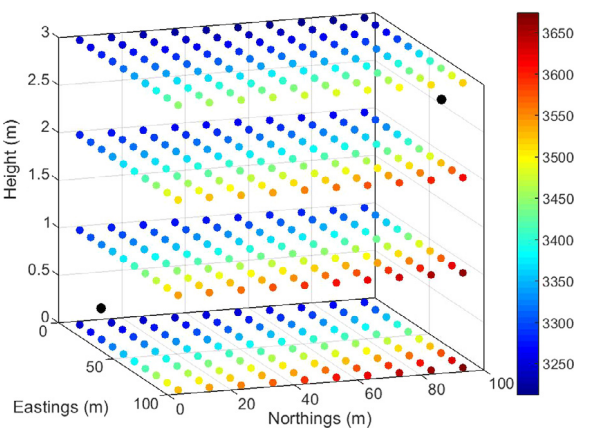
In this article, we demonstrate the need for an optimal design framework accounting for the effect of parameters which cannot be controlled, but which become known when data is collected, and affect the inference. Source estimation in remote sensing is used as a motivating example, in which wind largely determines the information available at a particular sensor location, but wind direction is not known until after a design is selected and observations are made.

A general framework for such problems is proposed for models with Gaussian likelihoods and computational challenges are discussed. A Bayes linear emulation framework is proposed to identify space of approximately optimal designs, greatly reducing the computer time needed to find spaces of good designs, while preserving problem structure. The performance of the Bayes linear framework is evaluated in application to a simple linear model and to a source estimation problem; the framework is effective in reproducing intuitive results, and greatly reduces the size of the non-implausible design space in only a small number of waves.

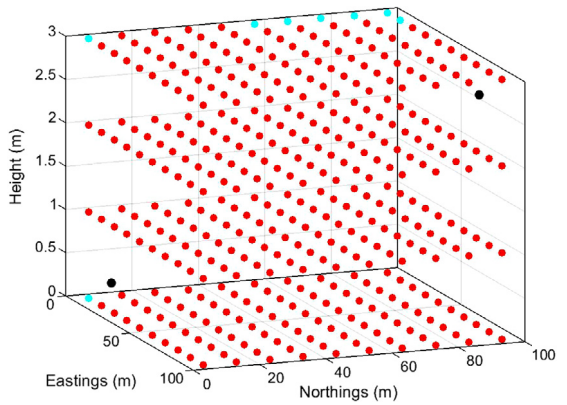
Two main areas for future research are identified. First, extension to incorporate sequential optimal design problems, in which we are free to select a new optimal design after a number of observations have already been collected. Consider, for example, a similar source estimation problem in which, after an initial batch of observations using a sensor at an a priori optimal location, updated beliefs indicate that a small area of the space is most likely to contain a source; if we are now permitted to choose a new optimal design, we may use this information in conjunction with beliefs about external conditions to choose a new location, close to, but downwind of the high-probability location.



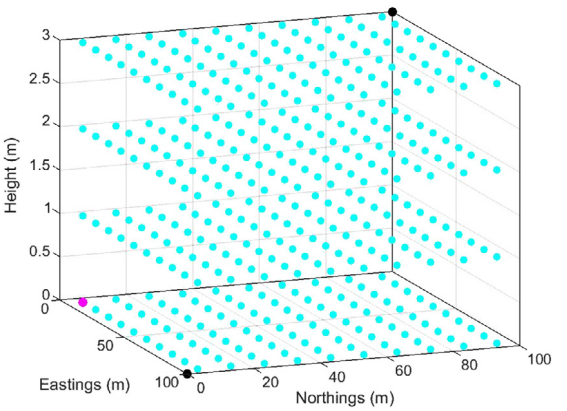
(a) $\mathrm{E}_{F_1}[f(q, w, d)]$.



(b) $\mathrm{E}_{F_1}[f(q, w, d)]$.



(c) $I_1(d) < 3$.



(d) $I_1(d) < 3$.

Fig. 6. Multi-sensor example: Fig. 6(a) shows a sample from the emulator fitted to the risk data, with two of the sensors located at the black dots ((10, 10, 0.2) and (90, 90, 2.8)) and the source located at the magenta dot ((120, −10), with emission rate 0.01 m^3/s), for different placements of the third sensor; Fig. 6(b) shows the mean of the integrated emulator, when two of the sensors are fixed in the same location (black dots), for different placements of the third sensor; Fig. 6(c) shows the non-implausible space for different locations of the third sensor; Fig. 6(d) shows the non-implausible space for different locations with the other two sensors fixed to their optimal locations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Secondly, extension to incorporate uncertainty in the external parameters. To illustrate, we might consider source estimation where we only have access to uncertain information about the true wind direction, again after a design has been chosen and measurements made. We still design according to our prior beliefs, but we must account for the additional uncertainty in the model parameter estimation and terminal decision-making stages appropriately.

Acknowledgements

Matthew Jones was supported by an EPSRC CASE Studentship in partnership with Shell Projects and Technology.

Appendix A. Computational details

A.1. Adjusting beliefs

Having obtained risk data F as described in Section 4, we adjust beliefs about the expected value of the risk at a particular input using the standard Bayes linear update equation (Goldstein and Wooff, 2007; Williamson and Goldstein, 2012)

$$\begin{aligned}\mathrm{E}_F[f(q, w, d)] &= \mathrm{E}[f(q, w, d)] + \mathrm{Cov}[f(q, w, d), F]\,\mathrm{Var}[F]^{-1}(F - \mathrm{E}[F]) \\ &= g(q, w, d)^{\mathrm{T}}\,\mathrm{E}[\beta] + W(q, w, d)^{\mathrm{T}}\,\mathrm{Var}[F]^{-1}(F - \mathrm{E}[F])\end{aligned}$$

where

$$\mathrm{Var}\left[F\right] = G\mathrm{Var}\left[\beta\right]G^{\mathrm{T}} + V \qquad W\left(q,w,d\right) = G\mathrm{Var}\left[\beta\right]g\left(q,w,d\right) + k\left(q,w,d\right)$$

$$G = \begin{pmatrix} g\left(q^{(1)},w^{(1)},d^{(1)}\right)^{\mathrm{T}} \\ \vdots \\ g\left(q^{(N)},w^{(N)},d^{(N)}\right)^{\mathrm{T}} \end{pmatrix} \qquad k\left(q,w,d\right) = \begin{pmatrix} c\left(\left(q,w,d\right),\left(q^{(1)},w^{(1)},d^{(1)}\right)\right) \\ \vdots \\ c\left(\left(q,w,d\right),\left(q^{(N)},w^{(N)},d^{(N)}\right)\right) \end{pmatrix}$$

and V is the matrix of covariance function values at all pairs of design points. Similarly, we obtain the adjusted covariance between function values at a new pair of input points as

$$\mathrm{Cov}_F\left[f\left(q,w,d\right),f\left(q',w',d'\right)\right] = g\left(q,w,d\right)^{\mathrm{T}}\mathrm{Var}\left[\beta\right]g\left(q,w,d\right) + c\left(\left(q,w,d\right),\left(q',w',d'\right)\right) - W\left(q,w,d\right)^{\mathrm{T}}\mathrm{Var}\left[F\right]^{-1}W\left(q,w,d\right).$$

A.2. Approximating the expected risk

Having fitted an emulator to the risk from an optimal terminal decision, we approximate the expectations required for Eq. (5) by computing the expectation of the emulator (O'Hagan, 1991; Rasmussen and Ghahramani, 2003). Since $f\left(.\right)$ is our approximation to the risk, the integral that we wish to perform is

$$f\left(d\right) = \iint f\left(q,w,d\right)p\left(q\right)p\left(w|d\right)dqdw.$$

Since this integral is a linear functional, we can compute our beliefs about the integrated quantity as

$$\mathrm{E}_F\left[f\left(d\right)\right] = g\left(d\right)^{\mathrm{T}}\mathrm{E}\left[\beta\right] + W\left(d\right)^{\mathrm{T}}\mathrm{Var}\left[F\right]^{-1}\left(F - \mathrm{E}\left[F\right]\right)$$

$$\mathrm{Cov}_F\left[f\left(d\right),f\left(d'\right)\right] = g\left(d\right)^{\mathrm{T}}\mathrm{Var}\left[\beta\right]g\left(d'\right) + c\left(d,d'\right) - W\left(d\right)^{\mathrm{T}}\mathrm{Var}\left[F\right]^{-1}W\left(d'\right)$$

where we can obtain the required integrated quantities as

$$W\left(d\right) = G\mathrm{Var}\left[\beta\right]g\left(d\right) + k\left(d\right)$$

$$g\left(d\right) = \iint g\left(q,w,d\right)p\left(q\right)p\left(w|d\right)dqdw \qquad k\left(d\right) = \iint k\left(q,w,d\right)p\left(q\right)p\left(w|d\right)dqdw$$

$$c\left(d,d'\right) = \iiiint c\left(\left(q,w,d\right),\left(q',w',d'\right)\right)p\left(q\right)p\left(w|d\right)p\left(q'\right)p\left(w'|d'\right)dqdwdq'dw'.$$

Appendix B. Covariance functions for Gaussian plume example

We present details of the emulator fitted to the risk in the Gaussian plume example for the single point-sensor case. A constant basis function, $g\left(q,w,d\right)$ was fitted, and the following covariance function chosen for the residual process

$$c\left(\left(q,w,d\right),\left(q',w',d'\right)\right) = \theta\exp\left[-\frac{1}{2}\left(\frac{\left(\left(d_x-q_x\right)-\left(d'_x-q'_x\right)\right)^2}{r_x^2} + \frac{\left(\left(d_y-q_y\right)-\left(d'_y-q'_y\right)\right)^2}{r_y^2} + \frac{\left(d_h-d'_h\right)^2}{r_h^2} + \frac{\left(q_s-q'_s\right)^2}{r_s^2} + \frac{\left(w-w'\right)^2}{r_w^2}\right)\right].$$

This covariance function imposes that pairs of points with similar horizontal and vertical distances between source and sensor locations, similar prevailing wind directions, similar sensor heights and similar source emission rates have similar risk values.

To obtain the required expected risks, this function must be integrated as in Appendix A. Using the prior distributions described in Section 5, we obtain the following expression for the once-integrated covariance function

$$\begin{aligned} c\left(d,\left(q',w',d'\right)\right) = {} & \frac{\hat{\sigma}_x}{\sigma_x D_x}\left(\Phi\left(\frac{\mathcal{L}_x-\hat{\mu}_x}{\hat{\sigma}_x}\right)-\Phi\left(\frac{-\hat{\mu}_x}{\hat{\sigma}_x}\right)\right)\frac{\hat{\sigma}_y}{\sigma_y D_y}\left(\Phi\left(\frac{\mathcal{L}_y-\hat{\mu}_y}{\hat{\sigma}_y}\right)-\Phi\left(\frac{-\hat{\mu}_y}{\hat{\sigma}_y}\right)\right) \\ & \times\frac{\hat{\sigma}_s}{\sigma_s D_s}\left(\Phi\left(\frac{\mathcal{L}_s-\hat{\mu}_s}{\hat{\sigma}_s}\right)-\Phi\left(\frac{-\hat{\mu}_s}{\hat{\sigma}_s}\right)\right)\frac{\sqrt{2\pi}r_w}{\pi/2}\left(\Phi\left(\frac{\pi/2-w'}{r_w}\right)-\Phi\left(\frac{-w'}{r_w}\right)\right) \\ & \times\theta\exp\left[-\frac{1}{2}\left(\frac{\left(\left(d_x-\mu_x\right)-\left(d'_x-q'_x\right)\right)^2}{r_x^2} + \frac{\left(\left(d_y-\mu_y\right)-\left(d'_y-q'_y\right)\right)^2}{r_y^2} + \frac{\left(d_h-d'_h\right)^2}{r_h^2}\right.\right. \\ & \left.\left. + \frac{\left(\mu_s-q'_s\right)^2}{r_s^2}\right)\right] \end{aligned}$$

where $\Phi(.)$ is the univariate Gaussian CDF,

$$D_v = \Phi\left(\frac{\mathcal{L}_v - \mu_v}{\sigma_v}\right) - \Phi\left(\frac{-\mu_v}{\sigma_v}\right) \qquad \hat{\sigma}_v = \left[\frac{1}{r_v^2} + \frac{1}{\sigma_v^2}\right]^{-1}$$

for any variable v, and

$$\hat{\mu}_x = \hat{\sigma}_x^2\left[\frac{d_x - (d'_x - q'_x)}{r_x^2} + \frac{\mu_x}{\sigma_x^2}\right]$$

$$\hat{\mu}_y = \hat{\sigma}_y^2\left[\frac{d_y - (d'_y - q'_y)}{r_x^2} + \frac{\mu_y}{\sigma_y^2}\right]$$

$$\hat{\mu}_s = \hat{\sigma}_s^2\left[\frac{q'_s}{r_s^2} + \frac{\mu_s}{\sigma_s^2}\right].$$

This expression is used to compute the elements of $k(d)$. For the double integrated covariance function $c\left(d, d'\right)$, we numerically integrate the above expression in each of the relevant variables. Note that since we have chosen a separable covariance function, each of the required numerical integrals is univariate, and therefore cheap to evaluate.

For the multi-sensor case, the covariance function is simply extended by including terms in the exponent corresponding to the scaled squared distances between each of the sensors and the source location. The once-integrated covariances can still be computed in closed form, and since we only seek to estimate a single source, the doubly-integrated covariance can still be computed by performing only univariate numerical integrals.

References

Bates, R.A., Buck, R.J., Riccomagno, E., Wynn, H.P., 1996. Experimental design and observation for large systems. J. R. Stat. Soc. Ser. B Stat. Methodol. 58, 77–94.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer.

Chaloner, K., Verdinelli, I., 1995. Bayesian experimental design: A review. Statist. Sci. 10, 273–304.

Craig, P.S., Goldstein, M., Seheult, A.H., Smith, J.A., 1996. Bayes linear strategies for history matching of hydrocarbon reservoirs. In: Bayesian Statistics. Vol. 5. Oxford University Press, pp. 69–98.

Cumming, J., Goldstein, M., 2009. Small sample designs for complex, high-dimensional models based on fast approximations. Technometrics 51, 377–388. http://dx.doi.org/10.1198/TECH.2009.08015.

DeGroot, M.H., 1970. Optimal Statistical Decisions. Wiley.

Goldstein, M., Rougier, J., 2006. Bayes linear calibrated prediction for complex systems. J. Amer. Statist. Assoc. 475, 1132–1143. http://dx.doi.org/10.1198/016214506000000203.

Goldstein, M., Wooff, D., 2007. Bayes Linear Statistics: Theory and Methods. Wiley.

Hennig, P., Schuler, C.J., 2012. Entropy search for information-efficient global optimization. J. Mach. Learn. Res. 13, 1809–1837.

Hirst, B., Jonathan, P., Gonzalez del Cueto, F., Randell, D., Kosut, O., 2012. Locating and quantifying gas emission sources using remotely obtained concentration data. Atmos. Environ. 45, 141–158.

Huan, X., Marzouk, Y.M., 2013. Simulation-based optimal Bayesian experimental design for nonlinear systems. J. Comput. Phys. 232, 288–317.

Jones, D.R., Schonlau, M., Welch, W.J., 1998. Efficient global optimization of expensive black-box functions. J. Global Optim. 13, 455–492.

Kennedy, M.C., O'Hagan, A., 2001. Bayesian calibration of computer models. J. R. Stat. Soc. Ser. B Stat. Methodol. 63, 425–464.

Lindley, D.V., 1972. Bayesian Statistics, A Review. Society for Industrial and Applied Mathematics.

Muller, P., 1998. Simulation based optimal design. In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian Statistics. Vol. 6. Oxford University Press, pp. 459–474.

O'Hagan, A., 1991. Bayes–Hermite quadrature. J. Statist. Plann. Inference 29, 245–260.

Pukelsheim, F., 1994. The three sigma rule. Amer. Statist. 48, 88–91.

Raiffa, H., Schlaifer, R., 1961. Applied Statistical Decision Theory. Harvard Business School.

Rasmussen, C.E., Ghahramani, Z., 2003. Bayesian Monte Carlo. In: Becker, S., Thrun, S., Obermayer, K. (Eds.), Advances in Neural Information Processing Systems, Vol. 15. MIT Press, pp. 505–512. URL: http://papers.nips.cc/paper/2150-Bayesian-monte-carlo.pdf.

Robert, C.P., Casella, G., 1999. Monte Carlo Statistical Methods. Springer.

Santner, T.J., Williams, B.J., Notz, W.I., 2003. The Design and Analysis of Computer Experiments. Springer-Verlag.

Smith, J.Q., 2010. Bayesian Decision Analysis: Principles and Practice. Cambridge University Press.

Stockie, J.M., 2011. The mathematics of atmospheric dispersion modelling. SIAM Rev. 53, 349–372.

Stuart, A.M., 2010. Inverse problems; a Bayesian perspective. Acta Numer. 19, 451–559. http://dx.doi.org/10.1017/S0962492910000061.

Vernon, I., Goldstein, M., Bower, R.G., 2010. Galaxy formation: a Bayesian uncertainty analysis. Bayesian Anal. 5, 619–669. http://dx.doi.org/10.1214/10-BA524.

Williamson, D., 2010. Policy making using computer simulators for complex physical systems: Bayesian decision support for the development of adaptive strategies. URL: http://etheses.dur.ac.uk/348/1/thesis.pdf.

Williamson, D., Goldstein, M., 2012. Bayesian policy support for adaptive strategies using computer models for complex physical systems. J. Oper. Res. Soc. 63, 1021–1033. http://dx.doi.org/10.1057/jors.2011.110.