



UNIVERSITY OF LEEDS

This is a repository copy of *An efficient screening method for computer experiments*.

White Rose Research Online URL for this paper:

<http://eprints.whiterose.ac.uk/84123/>

Version: Accepted Version

Article:

Boukouvalas, A, Gosling, JP and Maruri-Aguilar, H (2014) An efficient screening method for computer experiments. *Technometrics*, 56 (4). 422 - 431. ISSN 0040-1706

<https://doi.org/10.1080/00401706.2013.866599>

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

An efficient screening method for computer experiments

Alexis Boukouvalas
Aston University, United Kingdom
and
John Paul Gosling
University of Leeds, United Kingdom
and
Hugo Maruri-Aguilar
Queen Mary, University of London, United Kingdom

October 30, 2013

Abstract

Computer simulators of real world processes are often computationally expensive and require many inputs. The problem of the computational expense can be handled using emulation technology; however, highly-multidimensional input spaces may require more simulator runs to train and validate the emulator. We aim to reduce the dimensionality of the problem by screening the simulator's inputs for non-linear effects on the output rather than distinguishing between negligible and active effects. Our proposed method is built upon the elementary effects method for screening (Morris, 1991) and utilises a threshold value to separate the inputs with linear and non-linear effects. The technique is simple to implement and acts in a sequential way in order to keep the number of simulator runs down to a minimum, whilst identifying the inputs that have non-linear effects. The algorithm is applied on a set of simulated examples and a rabies disease simulator where we observe run savings ranging between 28% and 63% compared to the batch elementary effects method. Supplementary materials for this paper are available online.

Keywords: Morris design, Sensitivity analysis, Variable selection.

22 1 Analysis of complex computer simulators

23 Formal design and analysis of computer experiments stems from the seminal paper by Sacks
24 et al. (1989). In recent years, computer experiments have been used to make climate pro-
25 jections (Hargreaves et al., 2004), investigate biological processes (Wedge et al., 2009) and
26 estimate national carbon balances (Kennedy et al., 2008). In many applications, emula-
27 tors are constructed as efficient surrogate alternatives to expensive computer simulators
28 (O’Hagan, 2006). Further analyses are then carried out on the emulator; for example, sen-
29 sitivity analyses or prediction of response at new points. However, even when using the
30 emulator technology, the analysis of complex simulators can be restricted by the number
31 of input dimensions. In the present paper, we introduce a sequential screening procedure
32 to identify inputs that have a non-linear effect on the output. Our proposal stems from
33 the elementary effects (EE) method introduced in Morris (1991), and consists of sequential
34 estimation coupled with a space-filling criterion. Estimation of model behaviour is a com-
35 mon practice in sensitivity analysis (Morris, 1991) and use of space filling design ensures
36 reasonable coverage of the input space (Campolongo and Braddock, 1999). There exist a
37 wide range of methods in the literature to estimate functional effects. Screening methods
38 allow for efficient estimation of effects using a minimal number of simulator evaluations.
39 Variance-based sensitivity analyses methods (Saltelli et al., 2000) offer more precise results
40 in terms of the percentage of variance explained by each input factor and their interactions.
41 However, these methods require many more model runs and are employed at a later stage
42 of simulator analysis than screening algorithms.

43 Our aim in this paper is to develop a method that identifies inputs with non-linear
44 effects on the output. Morris (1991) considers screening for factors with linear, non-linear
45 or interacting effects, defined as active. Identification of inputs with active effects is the
46 usual aim of screening, also known as variable selection. Our focus is to distinguish factors
47 with non-linear or interaction effects from factors with linear or negligible effects. It is of
48 interest to identify the latter since they can be dealt with more simply in subsequent stages
49 of simulator analyses as discussed in Section 4. Screening methods are further discussed in
50 Section 2. We review the EE method in Section 3 and introduce our sequential method and
51 its properties in Section 4. We compare the method to other popular screening methods

52 through a simulated example in Section 5 and finish by demonstrating our method on a
53 computer model of rabies disease spread in Section 6.

54 **2 Screening approaches**

55 In classical experimentation on physical processes, screening usually incorporates some of
56 factorial designs, fractional factorial designs, nonregular orthogonal arrays, or optimal de-
57 signs. However, without further model assumptions, sequential use of such designs is not
58 straightforward and cannot be utilised to identify a nonlinear effect. When the number of
59 inputs dimensions is very large, but it is expected a priori that only a small subset of them
60 will have an effect on the output, the iterated fractional factorial design (IFFD) of Andres
61 (1997) has been shown to be highly efficient in practice. The IFFD method is most effi-
62 cient when there are very few highly influential factors (Saltelli et al., 2000). The IFFD
63 is an example of a supersaturated design. These designs use fewer model runs than input
64 dimensions, but typically make assumptions on the number of active inputs or the type of
65 effects on the response (Saltelli et al., 2000, Section 4.8). Kleijnen (2009) proposes sequential
66 bifurcation that can be effective, provided monotonicity of the model output with respect to
67 the inputs is satisfied.

68 If the simulator is computationally cheap, traditional sensitivity methods can be used such
69 as the Fourier amplitude sensitivity test (Cukier et al., 1973) or the method of Sobol' (Sobol,
70 1993). These approaches utilise a simulator-based functional ANOVA. The robust applica-
71 tion of such methods requires typically many more runs of the simulator than most screening
72 methods.

73 Response surface methods have been proposed for screening in high-dimensional problems
74 where a surrogate model is utilised to approximate the simulator response. In Linkletter et al.
75 (2006) and Savitsky et al. (2011), the surrogate is a Gaussian-process-based model and a
76 prior encapsulates the assumptions of effect sparsity. Reich et al. (2009) proposes a functional
77 ANOVA decomposition where a spline surrogate model is used to estimate each term. When
78 using any approach that relies on a surrogate model, careful validation is needed to ensure
79 the adequacy of the surrogate model for the purposes of screening.

80 The elementary effects (EE) method (Morris, 1991) is a popular methodology for sensi-

81 tivity analysis of computer simulators. The EE method requires no simplifying assumptions
 82 to be made on the ratio of active to total number of factors or their effect on the response
 83 (Saltelli et al., 2000). Furthermore, the method is easy to implement and computationally
 84 efficient. In the next section, we describe the EE method in detail, and, in Section 4, we
 85 extend the EE method to a more efficient, sequential approach.

86 **3 The elementary effects method**

87 Consider a deterministic simulator $Y(\cdot)$ with k input variables and design region $[0, 1]^k$.
 88 The simulator is assumed to be a smooth real valued function with a domain containing
 89 the design region. Computation of elementary effects starts from a point \mathbf{x} from which a
 90 trajectory is constructed with k random moves of size Δ . One-at-a-time (OAT) moves are
 91 performed along each single coordinate axis in turn to end at point $\mathbf{x} + \Delta(\mathbf{e}_1 + \dots + \mathbf{e}_k)$.
 92 The elementary effect for the i -th input variable for the trajectory starting at $\mathbf{x} \in [0, 1]^k$ is

$$EE_i(\mathbf{x}) = \frac{\mathbf{Y}\left(\mathbf{x} + \Delta \sum_{j=1}^i \mathbf{e}_j\right) - \mathbf{Y}\left(\mathbf{x} + \Delta \sum_{j=0}^{i-1} \mathbf{e}_j\right)}{\Delta}, \quad (1)$$

93 where $\Delta > 0$ is fixed. Here $i = 1, \dots, k$ indexes input factors and \mathbf{e}_i is the unit vector in
 94 the direction of the i -th axis where \mathbf{e}_0 is defined as $\mathbf{0}$. A total of $k + 1$ evaluations of $Y(\cdot)$
 95 are performed, ending with effects $EE_1(\mathbf{x}), \dots, EE_k(\mathbf{x})$. Each $EE_i(\mathbf{x})$ is a measure of the
 96 variation in the output with respect to a change in input i at point \mathbf{x} .

97 Consider R starting points \mathbf{x}_r , $r = 1, \dots, R$. From each point \mathbf{x}_r , we perform k OAT
 98 moves and compute elementary effects $EE_i(\mathbf{x}_r)$ for every input factor so that the total
 99 number of runs used in the EE method is $(k + 1) \times R$. The following sample moments are
 100 computed for each input factor:

$$\mu_i = \frac{1}{R} \sum_{r=1}^R EE_i(\mathbf{x}_r), \quad \mu_i^* = \frac{1}{R} \sum_{r=1}^R |EE_i(\mathbf{x}_r)| \quad \text{and} \quad \sigma_i = \sqrt{\sum_{r=1}^R \frac{(EE_i(\mathbf{x}_r) - \mu_i)^2}{R - 1}}. \quad (2)$$

101 The moment μ_i is an average effect measure, and high values suggest dominant contribution
 102 of the i -th input factor. The moment μ_i^* is a main effect measure, proposed in Campolongo
 103 et al. (2004) since μ_i may prove misleading due to cancellation of effects. An input with

104 a negligible or linear effect will have constant EE_i values and the corresponding σ_i will be
105 zero. Non-linear and interaction effects are identified with relatively large σ_i . An effects plot
106 can be constructed by plotting μ_i or μ_i^* against σ_i . This plot is a visual tool to detect and
107 rank effects. Factor effects close to the origin on this plot are the least influential.

108 There is interest in doing input screening with as few runs as possible and as the number
109 of input factors k is fixed, the size of the experiment is controlled by R . Usually small values
110 of R are used; for instance, Morris (1991) used $R = 3$ and $R = 4$ in his examples. A value
111 of R between 10 and 50 is mentioned in recent literature (Campolongo et al., 2004, 2007).
112 A larger value of R will improve the quality of the estimations, but at the price of extra
113 simulator runs.

114 The step size Δ is selected in such a way that all the simulator runs lie in the input space
115 and the elementary effects are computed within reasonable precision. The choice of Δ in
116 the literature is determined by the input space considered for experimentation, which is a k
117 dimensional grid constructed with p uniformly spaced values for each input. The number p
118 is recommended to be even and $\Delta = p/(2(p - 1))$, ensuring the elementary effects for each
119 input have an equal probability of selection during the trajectory design generation (Morris,
120 1991). To ensure that the trajectories remain in the design region, the implementation of the
121 EE method will use $-\Delta$ as necessary in place of Δ . The step Δ is usually kept at the same
122 value for all the inputs, but the method can be generalised to instead use different values of
123 Δ and p for every input. In the original proposal Morris (1991), the points $\mathbf{x}_1, \dots, \mathbf{x}_R$ were
124 taken at random from the input space grid. Campolongo et al. (2007) proposed spreading
125 runs over the design space by generating a large number of trajectory designs and selecting
126 a subset by maximising the minimum distance between them.

127 A potential drawback of OAT designs used in the EE method is that design points may
128 fall on top of each other when projected into lower dimensions. This disadvantage becomes
129 more apparent when the design runs are to be used in further modelling after discarding
130 unimportant factors. An alternative is to screen with randomly rotated simplices located
131 at points x_r , see Pujol (2009). The computation of distribution moments μ_i, μ_i^*, σ_i and
132 further analysis is similar to the EE method, with the advantage that projections of the
133 resulting design do not fall on top of existing points, and all observations can be reused in
134 a later stage. A disadvantage of this approach is the loss of efficiency in the computation

135 of elementary effects; that is, computing effects from a rotated simplex is suboptimal when
136 compared with (1) which is optimal for computing elementary effects, see Pujol (2009).

137 4 Sequential elementary effects method

138 We propose a sequential screening method that has the potential to reduce computation time
139 significantly. The methodology aims to distinguish between factors with linear or no effect
140 and factors with non-linear effect. In the EE method, all $(k + 1)$ evaluations are performed
141 for each trajectory irrespective of the type of factor effect. Whereas, in our sequential
142 strategy, when effects are identified as non-linear, their corresponding one-step evaluations
143 are removed from subsequent trajectories. This results in fewer model runs. The rationale
144 is that if σ_i is small for a given factor, then we should investigate whether σ_i remains small
145 when adding extra trajectories. At the end of experimentation, those input factors for which
146 σ_i remained small are considered to have linear or no effect, and factors for which σ_i was
147 bigger than a threshold have a non-linear effect on the output. A method of eliciting the
148 choice of threshold is presented in Section 4.1.

149 The justification of thresholding solely on the variance of the elementary effects σ_i is
150 that independent linear effects of factors may be removed from the simulator output at a
151 preprocessing stage or during the emulation phase. As an example of the latter approach,
152 factors with linear effects may be incorporated in the mean function of a Gaussian Process
153 (GP) emulator while omitted from the covariance specification. If we denote by A the
154 subset of $\{1, \dots, k\}$ that indexes factors with linear effects, the GP prior may be written
155 as $Y(x) = \beta + \sum_{i=1}^k a_i X_i + Z^*$ where Z^* is a stochastic process whose covariance structure
156 depends only on the variables with non-linear effects; that is, the x_i with $i \in \{1, \dots, k\} \setminus A$.
157 The residual process Z^* is therefore placed in a lower dimensional space simplifying the
158 design and inference tasks.

159 For our algorithm to run, a space filling design with M points is created. This design
160 provides the sequence of points at which the Morris OAT runs will be tested. Initially, we
161 select a good space filling design, such as a maximin Latin hypercube (LH) (Morris and
162 Mitchell, 1995). The generation of maximin designs is discussed further in (Ravi et al.,
163 1994). The value of M is selected such that $(k + 1)M$ is the maximum number of runs that

164 can be performed during the whole screening process.

165 A preprocessing stage orders the design points according to the biggest distance between
166 points. The first two points are those whose Euclidean distance is largest; then the third
167 point maximises the minimum distance between itself and the first two points, then a fourth
168 point is ordered in the same way, and so on. This procedure of ordering points mirrors
169 nearest neighbour clustering, but acts in an opposite manner as points are ordered from
170 those farthest apart to those that are closest.

171 **Example 1.** For $k = 5$ input factors and $M = 6$ runs, consider a maximin LH design in
172 $[0, 1]^5$ with point coordinates $\mathbf{x}_1 = (0, 4, 1, 4, 4)/5$, $\mathbf{x}_2 = (1, 5, 5, 1, 3)/5$, $\mathbf{x}_3 = (2, 1, 3, 0, 1)/5$,
173 $\mathbf{x}_4 = (3, 0, 4, 5, 2)/5$, $\mathbf{x}_5 = (4, 3, 0, 3, 0)/5$ and lastly $\mathbf{x}_6 = (5, 2, 2, 2, 5)/5$. The preprocessing
174 stage first selects the points \mathbf{x}_2 and \mathbf{x}_5 , which are furthest apart. The next point, \mathbf{x}_4 ,
175 maximises the distance between those remaining points and the first two points chosen.
176 The procedure continues by selecting \mathbf{x}_6 , then \mathbf{x}_1 and finishes with \mathbf{x}_3 . This preprocessing
177 stage produces the ordered sequence of points $\mathbf{x}_2, \mathbf{x}_5, \mathbf{x}_4, \mathbf{x}_6, \mathbf{x}_1, \mathbf{x}_3$, which are relabelled as
178 $\mathbf{X}_{(1)}, \dots, \mathbf{X}_{(6)}$.

179 The screening algorithm starts with the computation of elementary effects for all input
180 factors at the first two design points of the ordered maximin LH design. OAT runs are
181 created at those two points and elementary effects are computed. With this initial data,
182 a first estimation of the moments μ_i, μ_i^* and σ_i is available. If, for a given input factor,
183 its sample moment σ_i is larger than a specified threshold σ_0 , then we say that the output
184 is responding non-linearly to this input. This leads us to declare that input as active and
185 remove it from the list of current input factors. The technique continues by adding OAT runs
186 at the next point, only for factors suspected to be linear. Elementary effects are computed
187 and moments are updated for each added point, removing a factor if the condition for σ_i is
188 met. The methodology ends when all input factors have been removed or after computing
189 elementary effects for all M points. On ending, the input factors are separated into two
190 groups: those having non-linear effect and those with linear or no effect on the output.
191 Algorithm 1 sets out the procedure in pseudo-code form.

192 **Example 2.** To show how the proposed sequential algorithm works, consider $Y(x_1, x_2, x_3, x_4, x_5) =$
193 $\cos(x_3/5)(x_2 + 1/2)^4/(x_1 + 1/2)^2 + x_5$ on the design region $[0, 1]^5$. Note that x_4 has negligible

Algorithm 1 The procedure for completing our screening technique.

Screening algorithm

Input: Simulator $Y(\cdot)$ with k inputs; total number of one-at-a-time experiments M ; step size Δ ; threshold σ_0 .

Output: Moments μ_i, σ_i, μ_i^* ; lists of factors with linear (\mathcal{C}) and with non-linear effect (\mathcal{A}).

A. Preprocessing stage

1. Set design region to $[0, 1]^k$ and create space filling design with M points $\mathbf{x}_1, \dots, \mathbf{x}_M$.
2. Order the design points using maximum distance between points. Label the ordered points as $\mathbf{x}_{(1)}, \dots, \mathbf{x}_{(M)}$.

B. Calculating the elementary effects

1. Set $R := 2$ and the initial design to be $D := \{\mathbf{x}_{(1)}, \mathbf{x}_{(2)}\}$. Set list of current factors to $\mathcal{C} := \{1, \dots, k\}$ and list of active effects $\mathcal{A} := \emptyset$.
2. For every point in D , create one-at-a-time runs only for those input factors indexed by \mathcal{C} . Run the simulator at those points. This totals $|\mathcal{C}| + 1$ experiments for every point in D .
3. Using simulator runs from step B2 and (1), compute elementary effects $\{EE_i(\mathbf{x}) : \mathbf{x} \in D, i \in \mathcal{C}\}$.
4. If $R = 2$, compute moments μ_i, μ_i^* and σ_i using elementary effects for all factors. If $R > 2$, only update moments for the current list of input factors, indexed by \mathcal{C} .
5. For $i \in \mathcal{C}$, if $\sigma_i > \sigma_0$ then update $\mathcal{C} := \mathcal{C} \setminus \{i\}$ and $\mathcal{A} := \mathcal{A} \cup \{i\}$.
6. If $\mathcal{C} = \emptyset$, then all the inputs were identified non-linear. Algorithm ends.
7. If $R = M$, then all the design points available are exhausted. Algorithm ends.

C. Producing the next design point

1. Update $R := R + 1$; set $D = \{\mathbf{x}_{(R)}\}$.
2. Goto B2.

194 effect here. The function Y is treated as a simulator, from which the only information we
 195 require are its values at design points. We use the same pre-ordered LH design of Example
 196 1; set $p = 10$ for step size $\Delta = 5/9$ and threshold $\sigma_0 = 0.15$. See 4.1 for details on the con-
 197 struction of the threshold σ_0 . Random trajectories are constructed with the first two ordered
 198 points, giving the following moment estimates $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5) = (-7.87, 4.66, -0.047, 0, 1)$
 199 and $(\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5) = (6.62, 0.27, 0.06, 0, 0)$. Note that the method yields exact zeroes for
 200 the fourth variable. The estimates σ_1, σ_2 are greater than the threshold σ_0 and thus x_1
 201 and x_2 are separated as having non-linear effects. As all σ_3, σ_4 and σ_5 are smaller than
 202 σ_0 , further investigation is required for x_3, x_4 and x_5 . At the third design point, OAT ex-
 203 perimentation only for those factors produces updated moments $(\mu_3, \mu_4, \mu_5) = (-0.06, 0, 1)$
 204 and $(\sigma_3, \sigma_4, \sigma_5) = (0.05, 0, 0)$, i.e. the factors remain under investigation. The sequential
 205 methodology continues for all of them until finishing with all the design points. At this final
 206 step, updated moments are $(\mu_3, \mu_4, \mu_5) = (-0.039, 0, 1)$ and $(\sigma_3, \sigma_4, \sigma_5) = (0.04, 0, 0)$, where
 207 again for the fourth factor the results are exactly zero. We conclude that linearity of the
 208 response in terms of x_3, x_4 and x_5 over the design region could not be rejected. Near zero
 209 variances for x_4 and x_5 suggest strong conclusions for those factors.

210 The total experimental effort was 28 runs, from which the first 12 runs involved tra-
 211 jectories for all factors, while further 16 runs were required for the linear factors under
 212 investigation. This is a 22% reduction from the $(5 + 1) * 6 = 36$ runs needed to perform the
 213 batch EE method. The moment estimates obtained for the non-linear factors are only rough
 214 approximations of the true moment values, and the moment estimates for the linear factor
 215 were computed with more information. This asymmetry is apparent when comparing with
 216 exact analytic sensitivity results $\mu = (-5.34, 6.62, -0.04, 0, 1)$ and $\sigma = (8.88, 7.42, 0.06, 0, 0)$,
 217 where we use $\mu_i = \int_{[0,1]^5} \frac{\partial}{\partial x_i} Y(\mathbf{x}) d\mathbf{x}$ and $\sigma_i^2 = \int_{[0,1]^5} \left(\mu_i - \frac{\partial}{\partial x_i} Y(\mathbf{x}) \right)^2 d\mathbf{x}$.

218 4.1 Heuristic selection of variance threshold

219 In Algorithm 1, the elementary effect variance threshold σ_0 is an input. In this section, we
 220 present a heuristic to choose σ_0 indirectly by eliciting the expected divergence from linearity
 221 of the factor effect as the variance of an auxiliary random variable. A linear (or near-linear)

222 effect of the variable x_i is represented by an additive noise model:

$$Y(x_i) = ax_i + b + \varepsilon_i, \quad (3)$$

223 where ε_i is an independent and identically distributed normal random variable with zero
224 mean and known variance γ , and a, b are constants. In other words, the marginal effect due
225 to the factor x_i is modelled with a simple regression line. The auxiliary random variable ε_i
226 captures interaction with other factors and non-linearities.

227 In practice, the variance γ has to be elicited prior to the screening experiment and there
228 are two alternatives that we have considered in the examples in the present paper.

- 229 1. We might believe that the factor x_i has a linear effect, but the simulator runs contain a
230 numerical error. In this case, we expect γ to be set to a small value, such as a multiple
231 of machine precision.
- 232 2. If we believe that small non-linear effects will not have an appreciable impact on the
233 model output, then γ should be chosen to reflect the level of variation from a straight
234 line to be tolerated.

235 The elicitation of the variance parameter is in contrast to Kadane et al. (1980) and
236 Garthwaite and Dickey (1988), where full probability distributions are elicited that reflect
237 beliefs about the parameters of the linear model. In the present application, we do not wish
238 to prejudge the behaviour of the model: we want a number specifying how far we can tolerate
239 from being linear.

240 Given the variance γ , the sampling distribution of the variance of the elementary effects
241 can be calculated according to the following lemma, whose proof is given in Appendix A.
242 For simplicity we omit the subscript i from the quantities σ^2 , σ_Φ , γ and Δ , all of which could
243 potentially take different values for each input. In practice we choose a common γ and Δ
244 for all inputs.

245 **Lemma 1.** *Let x_1, \dots, x_R be univariate design points, at each of which trajectories are*
246 *constructed. Assume that observations taken at design points and trajectories follow the*
247 *model given in (3). Let elementary effects and moments be defined as in (1) and (2) and let*

248 $\sigma_{\Phi}^2 = \frac{2\gamma}{\Delta^2}$. Then

$$\sigma^2 \sim \frac{\sigma_{\Phi}^2}{R-1} \chi_{R-1}^2. \quad (4)$$

249 where χ_{R-1}^2 denotes a chi-square random variable with $R-1$ degrees of freedom.

We propose to use the 99% quantile of the cumulative distribution function of the chi-square distribution to derive the EE variance threshold σ_0 . We have found this choice of quantile sufficiently conservative for the examples we have investigated. The following equation

$$P(\sigma^2 \leq \sigma_0^2) = P\left(\frac{\sigma_{\Phi}^2}{R-1} \chi_{R-1}^2 \leq \sigma_0^2\right) = 0.99,$$

250 when inverted yields the threshold

$$\sigma_0 = \sqrt{\chi_{0.99, R-1}^2 \sigma_{\Phi}^2 / (R-1)}, \quad (5)$$

251 where $\chi_{0.99, R-1}^2$ is the 99% quantile of a chi-squared distribution with $R-1$ degrees of freedom.
 252 In other words, σ_0 defines a threshold over which the effect is considered non-linear; that is, if
 253 $\sigma_i < \sigma_0$, then the input variable x_i is retained for further testing. Note that Lemma 1 applies
 254 directly in a multivariate setting; in which case, the comparison is performed separately for
 255 each input variable.

256 In Example 2, we used a single threshold σ_0 for all variables. In order to obtain $\sigma_0 = 0.385$
 257 using the method described in this section, the values $R = 6$, $\Delta = 5/9$, $\sqrt{\gamma} = 8.7 \times 10^{-2}$ and
 258 quantile $\chi_{0.99, 5}^2 = 15.08$ could be used.

259 To simplify the algorithm, the threshold σ_0 may be kept fixed for all computations rather
 260 than adapting σ_0 to the actual number of trajectories involved. The main difference is in
 261 the degrees of freedom for the scaled chi-square distribution in (5). The adaptive approach,
 262 which was utilised in the simulation experiments presented, involves recomputing σ_0 with
 263 updated degrees of freedom prior to step B5 in Algorithm 1. In the simplified approach using
 264 a single value σ_0 , the method is more conservative than when varying degrees of freedom;
 265 i.e. the rejection rate is higher with fixed threshold than otherwise.

266 The detection of nonlinearity under the heuristics in this section is further explored in
 267 Appendix B. The robustness of the proposed selection of σ_0 in our sequential algorithm is
 268 explored under different departures from linear model (3).

5 Simulated high dimensional example

We illustrate the sequential screening method on the synthetic test function introduced in Morris (1991). The function is defined on 20 inputs $\mathbf{x} \in [0, 1]^{20}$ as follows:

$$y = \beta_0 + \sum_{i=1}^{20} \beta_i w_i + \sum_{i<j}^{20} \beta_{ij} w_i w_j + \sum_{i<j<l}^{20} \beta_{ijl} w_i w_j w_l + \sum_{i<j<l<s}^{20} \beta_{ijls} w_i w_j w_l w_s, \quad (6)$$

where $w_i = 2(x_i - \frac{1}{2})$ except for $i = 3, 5, 7$ where $w_i = 2(1.1x_i/(x_i + 0.1) - \frac{1}{2})$. The coefficients are set to $\beta_i = 20$ for $i = 1, \dots, 10$, $\beta_{ij} = -15$ for $i, j = 1, \dots, 6$, $\beta_{ijl} = -10$ for $i, j, l = 1, \dots, 5$ and $\beta_{ijls} = 5$ for $i, j, l, s = 1, \dots, 4$. The remaining first and second order coefficients are generated independently from a zero mean unit variance normal distribution and the remaining third and fourth order coefficients are set to zero. Factors x_1, \dots, x_7 have a non-linear effect on the function output while factors x_8, x_9, x_{10} have a linear effect and factors x_{11}, \dots, x_{20} have negligible effect (Morris, 1991; Pujol, 2009).

First, we compare the performance of the batch and sequential EE methods in terms of required model runs. The screening experiment was performed under the configuration used in Pujol (2009) for 100 realisations. As in Pujol (2009) the discretisation level has been set to $p = 20$ so $\Delta = 0.5263$ and the number of trajectories to $R = 10$. For the sequential procedure, a threshold value of $\gamma = 2.6$ was used. This corresponds to around 0.005% of the range of the response y in (6), which varies between -225 and 139 .

A total of 210 function evaluations are required for the batch EE procedure. For the sequential EE procedure on average 150 function evaluations are required with a standard deviation of 13, that is an average savings of 28% compared to the batch approach. Factors x_1, \dots, x_7 are correctly identified as having non-linear effect in 99% of the realisations because the corresponding σ_i is larger than the threshold σ_0 . Factors x_8, \dots, x_{10} are found to have linear effects in 92% of the realisations because μ_i is large and σ_i is small. Factors x_{11}, \dots, x_{20} are found to have negligible in all realisations due to small σ_i values. The full batch EE screening results and the first iteration of one realisation of the sequential algorithm are shown in Figure 1. This first iteration is equivalent to running the algorithm up to step B.4 in Algorithm 1. We note that six of the seven factors with non-linear effects are identified at this early stage.

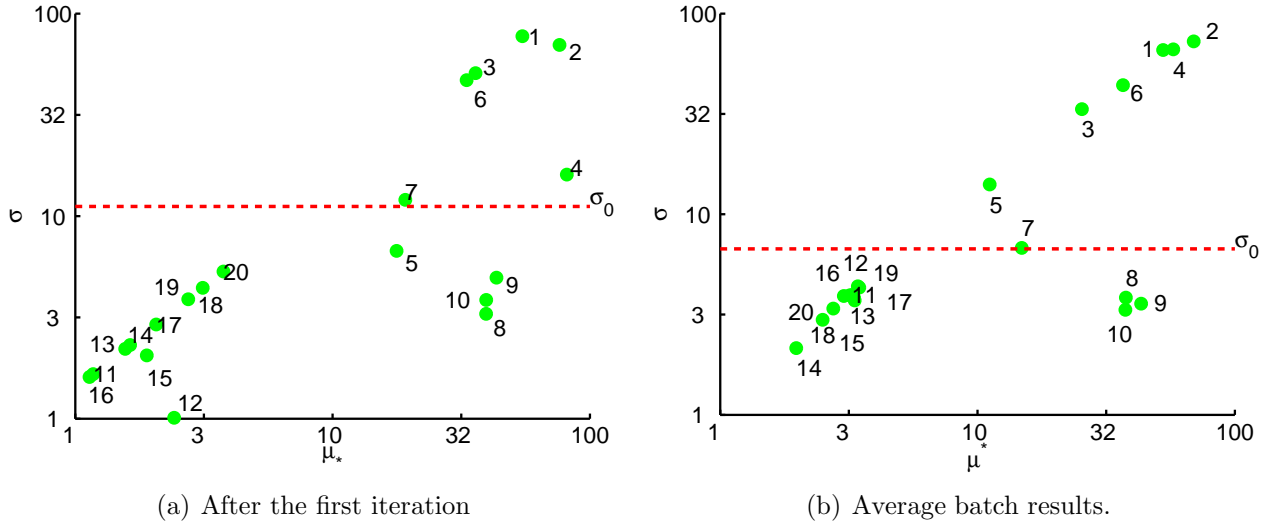


Figure 1: Applying the batch and sequential EE screening method on the 20 input factor Morris test function. Horizontal dashed line denotes the σ_0 threshold value for the given iteration.

296 Finally, we illustrate the efficiency of the EE method by contrasting it to a traditional
 297 sensitivity analysis method. Applying the Sobol' sensitivity analysis method (Sobol, 1993) to
 298 compute first-order and total indices using a random design of 220 runs, results in large 95%
 299 confidence intervals indicating that more model runs are required before any conclusions can
 300 be drawn from the examination of the indices (Figure 2). We also show how the uncertainties
 301 dramatically reduce when a larger design is used. These results were obtained by using the
 302 sensitivity R package (Pujol et al., 2013). This illustration confirms our expectation that
 303 screening methods such as the EE method, can be utilised prior to a more detailed sensitivity
 304 analysis in order to minimise the number of model runs.

305 We conclude that the sequential approach results in significant computational savings
 306 compared to the batch EE method as factors with clear non-linear effects can be eliminated
 307 in the early screening stages with high confidence. Furthermore, we have demonstrated the
 308 efficiency of the EE screening methods compared to classical sensitivity analysis as illustrated
 309 by the number of runs required to effectively calculate the Sobol's first-order and total indices
 310 for this example.

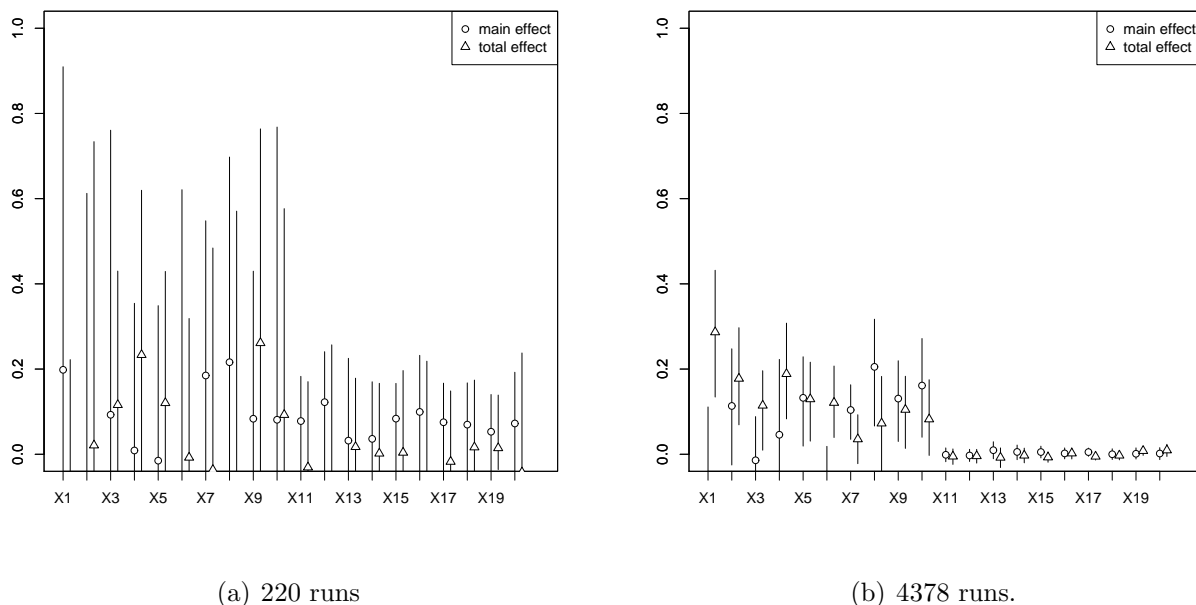


Figure 2: Applying Sobol's sensitivity analysis method on data generated using (6). Random designs of 220 and 4378 have been used. The error bars indicate 95% confidence intervals for the first-order and total effect indices.

6 Rabies Model

In this section, we discuss the application of the Morris sequential screening method described in Section 4 on a stochastic model provided by the Food and Environment Research Agency (FERA) (Singer et al., 2008, 2009). An overview of the stochastic simulator is given in Section 6.1, followed by a description of the screening methodology in Section 6.2) and a discussion of the results in Section 6.3.

6.1 Model Description

Although rabies in wild animal populations has been eradicated from large parts of Europe, there is a remaining risk of disease re-introduction. The situation is aggravated by an invasive species, the raccoon dog (*Nyctereutes procyonoides*) that can act as a second rabies vector in addition to the red fox (*Vulpes vulpes*). The purpose of the rabies model is to analyse the risk of rabies spread in this new type of vector community (Singer et al., 2008). The individual-based, non-spatial, discrete-time model incorporates population and disease

324 dynamical processes such as host reproduction and mortality as well as disease transmission.

325 The model includes two vector species: raccoon dogs and foxes. The model is non-spatial
 326 and disease propagation is calculated solely with respect to population dynamics.

	<i>Description</i>	min	max
1	Number of replications	200	300
2	Fox population winter density (ind./km ²)	0.1	0.5
3	Raccoon dog population winter density (ind./km ²)	0.1	1
4	Shape parameter for the probability distribution of raccoon dog infection	0.39	0.47
5	Dummy variable with no influence	0.9	1.1
6	Fox population mortality	0.9	1.1
7	Raccoon dog population mortality	0.9	1.1
8	Winter hunting proportion	0.9	1.1
9	Fox population birth rate	0.9	1.1
10	Raccoon dog population birth rate	0.9	1.1
11	Fox population infection rate	0.9	1.1
12	Fox population rabies incubation rate	0.95	1.05
13	Raccoon dog rabies incubation rate	0.95	1.05

Table 1: Input parameters for the rabies model and their ranges.

327 The model has 13 free parameters, shown in Table 1. In addition, three simulator config-
 328 uration parameters were kept fixed: the maximum number of time steps was set to 400; the
 329 cross infection input was 0.002 and the environment area size was kept fixed at 5400 km².

330 The response used is the probability that the rabies disease becomes extinct in both
 331 species within 5 years. This output is scaled to lie in the range [0, 100] and is important
 332 in deciding on the response to a potential rabies outbreak since it indicates the risk of long
 333 term rabies disease persistence (Singer et al., 2008, 2009).

334 6.2 Screening Methodology

335 For screening, we utilise a Morris design with the number of trajectories set to $R = 20$
 336 resulting in $(k + 1) \times R = 14 \times 20 = 280$ simulator evaluations, where $k = 13$. The mean
 337 μ^* of the absolute values of the Elementary Effects is used to rank the input factors. We set
 338 the number of levels to $p = 6$ and $\Delta = p/(2(p - 1)) = 0.6$.

339 The sequential Morris method discussed in Section 4 allows the specification of a variance
 340 γ from which the threshold on the Elementary Effect deviation σ_0 is derived. We set $\gamma = 3.5$,

341 which reflects a prior belief that individual factor effects on the output are considered near-
342 linear if the effect on the output is within three standard deviations of purely linear, i.e.
343 $\pm 3\sqrt{\gamma} = 5.6$. Since the output is bounded in the range $[0, 100]$ a factor has near-linear effect
344 if the output varies no more than 5.6% from linear. This variability encapsulates both the
345 internal variability of the stochastic model and our prior definition of a near-linear effect.

346 **6.3 Screening Results**

347 Singer and Kennedy (2008) performed sensitivity analysis on this model using the standard
348 Morris method with the same setup as here as well as the Sobol' method. They noted the
349 most important parameters are species winter densities (inputs (2) and (3)) and mortalities
350 (inputs (6) and (7)). They also noted the least influential factors are the dummy variable (5)
351 that has by definition no influence on the model output and parameter 4, a shape parameter
352 for the probability distribution of raccoon dog infection. It is also noted that the Sobol'
353 method is prohibitively expensive and offers low accuracy with a sample size of 300. They
354 suggested increasing the sample size and reducing the dimensionality of the problem by
355 fixing some of the factors to their nominal values. For expensive simulators this motivates
356 the usage of the Morris method.

357 The standard Morris method variable ranking with $R = 20$ trajectories is presented in
358 Figure 3(a) where the four dominant factors were found to be the winter densities for both
359 species (2,3) and the associated mortality rates (6,7). Singer and Kennedy (2008) have noted
360 that the dominant factors have strong non-linear and interaction effects which is reflected in
361 the high σ value observed in the Morris method.

362 The sequential Morris method is initialised with $R = 2$ trajectories on all 13 factors
363 requiring $(k + 1)R = 28$ simulator runs. The same step size Δ and the computed threshold
364 σ_0 is the same for all factors. The Morris plot with the associated threshold value is shown
365 in Figure 3(b). The elementary effect variance is over the threshold for four factors, the
366 raccoon dog winter density (3), raccoon dog rabies incubation probability (13), number of
367 replicated runs (1) and raccoon dog population birth rate (10) are eliminated from further
368 consideration since they have strong non-linear effects on the simulator output.

369 Another trajectory design for the remaining 9 factors is evaluated and requires 10 further
370 simulator evaluations (Figure 3(c)). The parameters fox winter density (2) and mortality

371 (6) factors are found to have non-linear effects and are removed from further consideration.
372 As evidenced by the Morris plot, the σ value for parameter 2 changed significantly from the
373 previous iteration where the effect was considerably below the threshold and very close to
374 linear.

375 For the third iteration, the seven factor trajectory requires 8 more simulator evaluations
376 (Figure 3(d)). Three further parameters are eliminated, the raccoon dog (7) mortality rate,
377 the fox birth rate (9) and the fox population rabies incubation probability (12) where large
378 changes in the moments of the elementary effects are again observed due to the increased
379 accuracy from the increased Morris design size. No more factors are eliminated until iteration
380 7 requiring a further $(4 + 1) \times 4 = 20$ simulator evaluations (Figure 3 (e)-(h)). At iteration
381 7, the winter hunting proportion (8) and fox population infection rate (11) parameters are
382 removed from further consideration.

383 The remaining two factors, the shape parameter for the probability distribution (4) and
384 the dummy variable (5), are found to be below the σ threshold for all subsequent twelve
385 iterations requiring $(2 + 1)12 = 36$ simulator evaluations. The total number of simulator
386 evaluations for the sequential procedure is 102 compared to the 280 evaluations required by
387 the standard batch Morris method with $R = 20$ trajectories, yielding a 27% reduction in
388 required code runs.

389 We have also performed the threshold calculation on the full Morris set with $R = 20$
390 trajectories and the same factors as with the sequential version are identified as near-linear.
391 In summary, the sequential Morris method for the rabies model has been successfully used
392 to identify factors with no or near-linear effects on the simulator response at a significant
393 saving to the standard Morris method.

394 7 Discussion

395 We have presented a novel sequential screening method that extends the batch elementary ef-
396 fects method originally proposed by Morris (1991). The method aims to identify inputs with
397 non-linear effects with a minimum number of trajectories. We have empirically demonstrated
398 the computational savings achieved compared to the batch approach on both synthetic data
399 and a real-world simulator. A critical aspect of the method is the specification of the thresh-

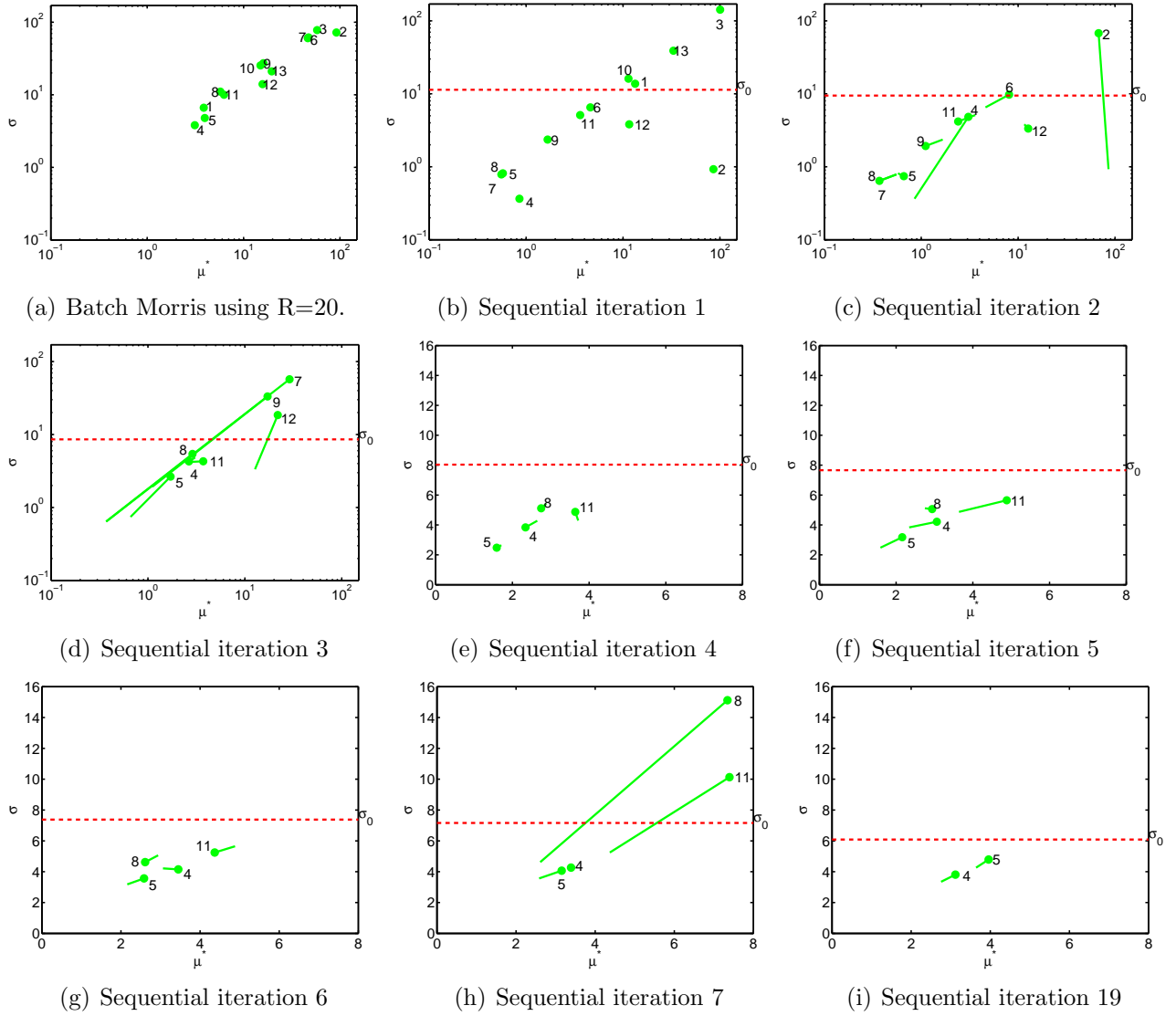


Figure 3: Batch and Sequential EE screening on the rabies simulator. Solid line denotes path from previous value of (μ^*, σ) for each factor. Horizontal dashed line denotes the σ_0 threshold value for the given iteration.

old σ_0 which is used to determine whether an input has a non-linear effect. The elicitation of this value directly can be challenging and we have presented an indirect approach which utilises an easily interpretable variance value γ specified on the simulator output space.

In order to apply the screening method of the present paper, the analyst must make a number of choices. To create the ordered design of OAT-experiment start-points, M must be specified. We recommend that M is chosen with respect to the computational effort needed to run the simulator; at worst, the simulator will run $(k+1)M$ times. The threshold value σ_0 can be set to zero so that only true linear- and no-effect inputs are investigated. We suggest to use a threshold $\sigma_0 > 0$ as in practice some small non-linearities might be tolerated and replaced instead by a linear function.

In cases where direct elicitation of the EE variance threshold σ_0 is not feasible or straightforward, elicitation of the variance γ may be preferable. The variance γ may be interpreted as the expected divergence of a factor from strict linearity under which the factor effect may still be considered linear for modelling purposes. In the case of stochastic simulators, this parameter includes the internal simulator variability whereas for deterministic simulators near-linear definitions only include errors due to machine precision and degree of departure from truly linear effects on the output. In the future, alternative elicitation methods may be constructed to allow for other non-Gaussian deviations from linear effect. An area of future interest is double classification of factors using the joint distribution of moments μ_i^* and σ_i . This not only would classify a factor as linear or non-linear, but would discriminate between active and non-active factors. A starting point for this would be the t test for μ in Morris (1991).

A variation of our algorithm is to select OAT runs to improve space filling properties of the whole design. The simplest approach is to use random OAT runs. However, OAT runs can be selected to maximise distance between them, as in Campolongo et al. (2007). An adaptation of the OAT runs suggested in Pujol (2009) can also be used where we create a simplex in the current set of input factors. Another criterion is to select OAT runs to improve uniformity of unidimensional projections of the design thus minimising the unidimensional discrepancy of the whole design, conditioned on the initial hypercube. The use of discrepancy to measure uniformity stems from quasi-Monte-Carlo methods (Niederreiter, 1992). We have used Sobol's *low discrepancy* space filling sequence (Niederreiter, 1992). The only

431 change required in the pseudo-code of Algorithm 1 is to remove step A2. Sampling from low
432 discrepancy sequences has the advantage of sequential generation of points. However, for
433 small sample sizes the spread of points of a low discrepancy sequence may not be as good as
434 that of a maximin space filling design with fixed size.

435 In this paper, we have not addressed the question of how multiple simulator outputs could
436 be handled in our method. The simplest approach, of generating separate OAT designs for
437 each output, is inefficient. In a sequential setting, the initial design could be shared for all
438 outputs. In our approach, factors are excluded from subsequent screening stages when non-
439 linear effects are detected. Therefore, subsequent stages need only include factors that are
440 under the EE variance threshold across all outputs. As the number of outputs grows, it is
441 more likely that a factor will have a non-linear effect on at least one of the outputs. Therefore
442 fewer simulator runs will be required for screening but it is also more likely fewer factors with
443 only linear or no effects across all outputs are detected. An alternative approach would be to
444 use a functional summary of all simulator outputs as the response for the screening analysis.
445 However the EE variance threshold will need to be elicited for the functional summary rather
446 than for each individual simulator output.

447 **Acknowledgements**

448 This work was carried out within the Managing Uncertainty in Complex Models (EPSRC
449 grant EP/D048893/1). We thank Dan Cornford for many useful discussions. We also thank
450 Graham Smith and Alexander Singer at the Food and Environment Research Agency for
451 providing the rabies model and explaining the simulator structure.

452 **A Proof of Lemma 1.**

453 *Proof.* All computations in this lemma are defined for a single factor. The elementary effect
454 at point x_i follows a normal distribution $EE(x_i) \sim N(a, \frac{2\gamma}{\Delta^2})$. Independence of elementary
455 effects $EE(x_1), \dots, EE(x_R)$ follows from independence of observations of model.

456 The mean of R elementary effects is distributed $\mu \sim N(a, \frac{2\gamma}{R\Delta^2})$. To find the distribution

457 of $\sigma^2 = \frac{1}{R-1} \sum_{i=1}^R (EE(x_i) - \mu)^2$, the following sum of squares is used

$$\sum_{i=1}^R \left(\frac{EE(x_i) - a}{\sqrt{\frac{2\gamma}{\Delta^2}}} \right)^2 = (R-1) \frac{\sigma^2}{\frac{2\gamma}{\Delta^2}} + R \frac{(\mu - a)^2}{\frac{2\gamma}{\Delta^2}}.$$

458 The left hand side above is a sum of squared independent standard normal variables and
 459 thus it has a chi-squared distribution with R degrees of freedom. By independence of μ
 460 and σ^2 , the first summand on the right has a chi-squared distribution with $R-1$ degrees of
 461 freedom and the second summand has a chi-squared distribution with one degree of freedom.
 462 Therefore $\sigma^2 \sim \frac{2\gamma}{(R-1)\Delta^2} \chi_{R-1}^2$. □

463 B Power computations

464 The input x_i is declared non-linear the first time that σ_i satisfies $\sigma_i > \sigma_0$ in step B5 of the
 465 algorithm. If the total number of experiments satisfies $M > 2$, then detection of nonlinearity
 466 for a given input can occur before the total number of allowed experiments M is reached.
 467 Power computations are possible for different departures from the linear model used to
 468 determine σ_0 , here we explore two.

469 The first departure we consider is a change in variance, with variance γ being multiplied
 470 by a factor δ^2 to become $\gamma' = \delta^2\gamma$. The factor δ quantifies change in variance from the original
 471 model. The probability of detecting the change at the step R is $\Pr(X_{R-1}^2 > \chi_{0.99, R-1}^2 / \delta^2)$.
 472 A geometric walk argument with probabilities changing as more data is included yields the
 473 power: $1 - \prod_{R=2}^M (1 - \Pr(X_{R-1}^2 > \chi_{0.99, R-1}^2 / \delta^2))$.

474 A different departure from linear model is when data follows

$$Y(x_i) = cx_i^2 + ax_i + b + \varepsilon_i. \tag{7}$$

475 Here ε_i still satisfies the same assumptions as for the linear model and a, b and c are constants.
 476 The following scenarios were used for power computations through simulations under the
 477 (wrong) model (7): $M = 2, 3, 4, 5$, $a = \frac{1}{10}, \frac{1}{5}, \frac{1}{2}, 1$; and the quadratic coefficient depended
 478 on the slope through the following relation $c = \delta a$ with δ ranging from 10^{-2} to 10^3 . Here
 479 the quantity δ is used to quantify departures from the original model through an increasing

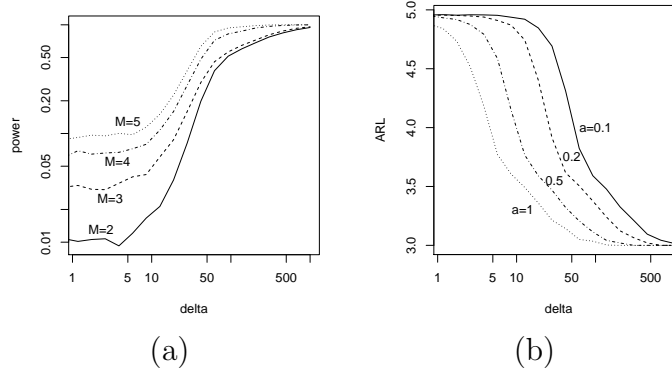


Figure 4: Simulation results: (a) power and (b) ARL against scaling factor δ .

quadratic coefficient. The errors in the models simulated were standard normal, i.e. $\gamma = 1$ and in each scenario used a uniform design of M points over $[0, 1]$. The results of power against different values of M with $a = \frac{1}{10}$ are shown in Figure 4 (a). As would be expected, the test reacts quickly to increasing quadratic trend through δ , but also as M increases and even for very small changes there is already a higher rejection rate. For instance, when $M = 5$ the rate is already larger than 10%.

In each simulation experiment, the run length was recorded. Recall that run length is the number of steps required for a factor to be declared nonlinear and it only takes integer values. A number of 12,000 simulations were carried out and run lengths obtained were averaged to produce the average run length (ARL). This computation was performed for $M = 5$ and same scenarios for parameters a and c as above. The graph of ARL against δ is shown in Figure 4 (b), where the top curve ($a = \frac{1}{10}$) required more exploration and produced higher ARL than the rest of the cases for a , with the lowest ARL in the simulations achieved for $a = 1$ (bottom curve), where the departure from linear is highest and the method detects it very quickly. Simulations for both Figures 4 (a) and (b) used the same range of δ but in the right figure the results are plotted only for $M = 5$.

Supplementary Materials

Technical Report: Supplementary information on the scalability of the sequential elementary effects method. (pdf)

499 **Matlab code for the sequential screening method:** Matlab code that demonstrates the
500 sequential screening method on a synthetic dataset. The example can be started by
501 running seqMorrisExample.m. (GNU zipped tar file)

502 **References**

- 503 Andres, T. (1997). Sampling methods and sensitivity analysis for large parameter sets.
504 *Journal of Statistical Computation and Simulation*, 57(1-4):77–110.
- 505 Campolongo, F. and Braddock, R. (1999). The use of graph theory in the sensitivity analysis
506 of the model output: a second order screening method. *Reliability Engineering & System
507 Safety*, 64(1):1–12.
- 508 Campolongo, F., Cariboni, J., and Saltelli, A. (2007). An effective screening design for
509 sensitivity analysis of large models. *Environmental Modelling Software*, 22(10):1509–18.
- 510 Campolongo, F., Cariboni, J., Saltelli, A., and Schoutens, W. (2004). Enhancing the Morris
511 Method. In Hanson, K. M. and Hemez, F. M., editors, *Proceedings of the 4th International
512 Conference on Sensitivity Analysis of Model Output*, pages 369–79, Santa Fe, New Mexico.
- 513 Cukier, R., Fortuin, C., Shuler, K., Petschek, A., and Schaibly, J. (1973). Study of the
514 sensitivity of coupled reaction systems to uncertainties in ratecoefficients. *Journal of
515 Chemical Physics*, 59:38733878.
- 516 Garthwaite, P. H. and Dickey, J. M. (1988). Quantifying expert opinion in linear regression
517 problems. *Journal of the Royal Statistical Society. Series B. Methodological*, 50(3):462–474.
- 518 Hargreaves, J. C., Annan, J. D., Edwards, N. R., and Marsh, R. (2004). An efficient cli-
519 mate forecasting method using an intermediate complexity Earth System Model and the
520 ensemble Kalman filter. *Climate Dynamics*, 23:745–60.
- 521 Kadane, J. B., Dickey, J. M., Winkler, R. L., Smith, W. S., and Peters, S. C. (1980).
522 Interactive elicitation of opinion for a normal linear model. *Journal of the American
523 Statistical Association*, 75(372):845–854.

- 524 Kennedy, M., Anderson, C., O'Hagan, A., Lomas, M., Woodward, F., Gosling, J., and
525 Heinemeyer, A. (2008). Quantifying uncertainty in the biospheric carbon flux for England
526 and Wales. *Journal of the Royal Statistical Society - Series A*, 171:109–135.
- 527 Kleijnen, J. P. C. (2009). Factor screening in simulation experiments: Review of sequential
528 bifurcation. *Advancing the Frontiers of Simulation: A Festschrift in Honor of George S.*
529 *Fishman*, pages 147–173.
- 530 Linkletter, Crystal, Bingham, Derek, Hengartner, Nicholas, Higdon, David, Ye, and Kenny,
531 Q. (2006). Variable selection for gaussian process models in computer experiments. *Tech-*
532 *nometrics*, 48(4):478–490.
- 533 Morris, M. and Mitchell, T. (1995). Exploratory designs for computer experiments. *Journal*
534 *of Statistical Planning and Inference*, 43:381–402.
- 535 Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments.
536 *Technometrics*, 33:161–74.
- 537 Niederreiter, H. (1992). *Random number generation and quasi-Monte Carlo methods*. Society
538 for Industrial and Applied Mathematics, Philadelphia, PA, USA.
- 539 O'Hagan, A. (2006). Bayesian analysis of computer code outputs: a tutorial. *Reliability*
540 *Engineering & System Safety*, 91:1290–1300.
- 541 Pujol, G. (2009). Simplex-based screening designs for estimating metamodels. *Reliability*
542 *Engineering & System Safety*, 94:1156–60.
- 543 Pujol, G., Iooss, B., and Janon, A. (2013). *sensitivity: Sensitivity Analysis*. R package
544 version 1.7.
- 545 Ravi, S. S., Rosenkrantz, D. J., and Tayi, G. K. (1994). Heuristic and special case algorithms
546 for dispersion problems. *Operations Research*, 42(2):pp. 299–310.
- 547 Reich, B. J., Storlie, C. B., and Bondell, H. D. (2009). Variable selection in bayesian smooth-
548 ing spline anova models: Application to deterministic computer codes. *Technometrics*,
549 51(2):110–120.

- 550 Sacks, J., Welch, W., Mitchell, T., and Wynn, H. (1989). Design and analysis of computer
551 experiments. *Statistical Science*, 4:409–23.
- 552 Saltelli, A., Chan, K., and Scott, E., editors (2000). *Sensitivity Analysis*. New York: Wiley.
- 553 Savitsky, T., Vannucci, M., and Sha, N. (2011). Variable Selection for Nonparametric Gaus-
554 sian Process Priors: Models and Computational Strategies. *ArXiv e-prints*.
- 555 Singer, A., Kauhala, F., Holmala, K., and Smith, G. (2008). Towards the elimination of
556 rabies in eurasia. *Developments in Biologicals*, 131:213–222.
- 557 Singer, A., Kauhala, F., Holmala, K., and Smith, G. (2009). Rabies in northeastern europe
558 - the threat from invasive raccoon dogs. *Journal of Wildlife Diseases.*, 45(4):1121–1137.
- 559 Singer, A. and Kennedy, M. (2008). Sensitivity analysis on the rabies model. Personal
560 Communication.
- 561 Sobol, I. M. (1993). Sensitivity analysis for non-linear mathematical models. *Math. Modelling*
562 *Comput. Exp.*, 1:407–414.
- 563 Wedge, D. C., Rowe, W., Kell, D. B., and Knowles, J. (2009). In silico modelling of di-
564 rected evolution: Implications for experimental design and stepwise evolution. *Journal of*
565 *Theoretical Biology*, 257:131–41.