Durham Research Online

Deposited in DRO:

03 January 2018

Version of attached file:

Published Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Wilson, A.L. and Dent, C.J. and Goldstein, M. (2018) 'Quantifying uncertainty in wholesale electricity price projections using Bayesian emulation of a generation investment model.', Sustainable energy, grids and networks., 13 . pp. 42-55.

Further information on publisher's website:

Publisher's copyright statement:

Additional information:

Quantifying uncertainty in wholesale electricity price projections using Bayesian emulation of a generation investment model

A.L. Wilson [a,*], C.J. Dent [a], M. Goldstein [b]

[a] *School of Mathematics, University of Edinburgh, Edinburgh, EH9 3FD, United Kingdom*
[b] *Department of Mathematical Sciences, Durham University, Durham, DH1 3LE, United Kingdom*

ARTICLE INFO

ABSTRACT

Policy-makers need to be confident that decisions based on the outputs of energy system models will be robust in the real-world. To make robust decisions it is critical that the consequences of uncertainty in model outputs are assessed. This paper presents statistical methodology for quantifying uncertainty associated with the output of a computer model of the long-term GB electricity supply. The output of the computer model studied is the projection of wholesale electricity prices from 2016 to 2030. The effect on wholesale prices of both uncertainty in input parameters and structural discrepancy is modelled. A probability distribution is used to model uncertainty over four inputs of the model: gas price, demand, EU ETS price and future offshore deployment. Estimates of the structural discrepancy introduced by the use of smoothed gas price projections and assuming that coal prices out to 2030 are known are obtained from experimentation with the computer model. A statistical model, known as an emulator, is fitted to a set of computer model evaluations and used to model uncertainty in the output of the computer model at inputs that have not been tested. The emulator is combined with the probability distribution over the inputs and the estimate of structural discrepancy to make an assessment of the overall uncertainty in the wholesale electricity price projections. A sensitivity analysis is also performed to investigate the effect of each of the four inputs on the trajectory of wholesale electricity prices.

© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Computer models are widely used to help study the behaviour of energy systems and to make decisions about these systems (e.g. see [1] and [2]). Computer models generally combine a set of uncertain input assumptions with an approximation of some part of an energy system to give some output of interest. Decision-makers need confidence that decisions made on the basis of this model output will be robust in the real-world, rather than within the computer model. This distinction makes it critical to study the uncertainty in how the output of the computer model relates to the real energy system. Without assessing uncertainty in model output it is not possible to draw conclusions based on model output that relate to the real-world. Uncertainties that should be considered to make this link include: parametric uncertainty, which stems from lack of knowledge about which input parameters to use, and structural discrepancy, which relates to the imperfect approximation of the system. This paper presents methodology for quantifying

uncertainty so that real-world decisions that are reliant on computer model output can be based on a more complete picture of the uncertainties associated with this output.

The computer model studied in this paper is a long-term model of the future Great Britain (GB) electricity supply. This model takes inputs such as projected future demand, future fossil fuel prices and the costs of future technology and uses these assumptions to model investment in generation and electricity supply in GB. Outputs from the model are wide ranging and include future wholesale electricity prices, future generation mix and emissions. In this paper one particular aspect of the modelling is studied, namely obtaining projections of wholesale electricity prices out to 2030. Uncertainty in these projections is of interest to policy-makers because of the impact that electricity prices have on other aspects of an energy system, such as on demand for electricity and on future investment (e.g. see [3]). In GB, wholesale electricity prices are particularly of interest because of political concerns about the effect of high electricity prices on domestic and industrial consumers [4].

The computer model estimates wholesale electricity prices by comparing daily load curves on sample days (net of wind generation, interconnection, storage and reserve requirements) to the

* Corresponding author.
E-mail address: Amy.L.Wilson@ed.ac.uk (A.L. Wilson).

generation merit order, accounting for increased prices above the marginal cost at times of short supply. This merit order is estimated using assumptions about the short term costs for each plant in the system. The future projections depend on a large number of uncertain model inputs. Four of these are focused on: gas prices, demand, European Union Emissions Trading System (EU ETS) price and offshore wind deployment. These inputs were selected after discussion with model developers as being inputs which may have a large impact on wholesale electricity prices, as well as being associated with a high level of uncertainty. Confidentiality issues prevent further details about the model being discussed here but this does not affect the validity of this study, which focuses on the assessment of uncertainty in model outputs, and not on technical modelling details.

When assessing uncertainty for wholesale electricity price projections, one consideration is the extent to which uncertainty in the four selected input assumptions affects the wholesale price projections. An assessment of the overall uncertainty in wholesale price projections due to input uncertainty as well as the effect that individual inputs have on the wholesale price projections will be made. However, even if the true values of the four input assumptions were known and the model was run at these values, there would still likely be some error, or structural discrepancy, in the price projections. A further aspect of this study therefore considers the quantification of this structural discrepancy. In particular, two possible sources of structural discrepancy are considered. The first is the use of a smoothed time-series of future gas prices as an input to the model, when in reality gas prices are volatile. The second is the effect that uncertainty associated with future coal prices (which are held as a fixed time series in this study) has on the wholesale price projections.

A major practical difficulty associated with the study of uncertainty in computer models is that these models can take a long time to run. The model studied in this paper takes around one hour to perform one model evaluation. Traditional Monte Carlo (MC) simulation involves drawing many inputs from a probability distribution (used to represent uncertainty in these inputs) and running the computer model at every input drawn. When models are slow to evaluate, this process is infeasible because it is not possible to run the model at enough inputs to get a complete picture of uncertainty over the input space. To resolve this difficulty, a statistical model (known as an emulator) is fitted to a small number of model evaluations. A computer model can be thought of as a function $\mathbf{f}(\cdot)$, taking some vector of inputs $\mathbf{x}$ and returning some vector of outputs $\mathbf{f}(\mathbf{x})$. An emulator specifies the uncertainty in $\mathbf{f}(\mathbf{x})$ for any input $\mathbf{x}$. Thus, even for inputs that have not been run, an approximation of the model output at that input (given by the mean of the emulator) and the error in this approximation (given by the standard deviation of the emulator) can be obtained. Emulation makes it possible to assess uncertainty when the number of model evaluations is limited, whilst also quantifying the additional uncertainty arising from this sparse coverage of the input space.

This paper uses emulation and a model for structural discrepancy to assess uncertainty in projections of the wholesale electricity price obtained using a computer model. The steps taken are applicable to a wide range of applied problems in energy systems modelling. Any modelling study which uses a slow-to-run computer model to obtain results and has uncertain inputs and/or structural discrepancy could benefit from the methodology described here. This paper focuses on a model that is used to obtain long-term projections. Examples of similar computer modelling studies that could benefit from the methodology include those associated with the TIMES/MARKAL family, PRIMES and ESME (see [2] for a summary of the major UK energy models).

Emulation has previously been used to assess uncertainty in the output of energy system models in [5] and [6]. This paper extends [5] and [6] with three key contributions. The first is that Bayes linear methods are used to fit the emulator. As with a conventional Bayesian analysis, Bayes linear methods have the advantage that expert judgement can be incorporated when fitting the emulator, but Bayes linear fitting uses only linear algebra rather than time-consuming Markov Chain Monte Carlo (MCMC). The second contribution is that computer model experimentation is used to assess structural discrepancy. Using experimentation as demonstrated in this paper is particularly useful in the absence of historical data, or where historical data are not thought to be representative of the future behaviour of a system. More details on these two points are given in Section 2. The third contribution is that emulation is demonstrated on a real-world policy problem, that of assessing uncertainty in projections of wholesale electricity prices.

The structure of the rest of this paper is as follows. In Section 2 existing methods for assessing uncertainty in energy models are reviewed. In Section 3, the process for the selection of model evaluations to run is described. In Section 4, an emulator is fitted to these model evaluations. In Section 5, results from the study on structural discrepancy are given. Finally, in Section 6, the emulator in combination with uncertainty specifications for structural discrepancy and parametric uncertainty is used to study uncertainty in wholesale electricity price projections.

2. Literature review

In energy systems modelling, the standard technique for assessing the consequences of input uncertainty is to run a scenario analysis, considering model outputs for a small selection of different future scenarios, without considering the relative probabilities of these scenarios [7–9]. There are several disadvantages when using this approach to assess uncertainty in model output. One disadvantage is that only a small set of scenarios are tested, and so the full range of possible inputs is not explored. It is unlikely that any single scenario will occur in practice, and so testing only a small number of scenarios gives an unrealistic view of the possible range of outputs. A study in [10] highlights the importance of this issue by comparing historic demand scenarios to observed demand in the UK, concluding that the demand scenarios considered did not capture the observed behaviour of UK demand. Another disadvantage is that a scenario-based analysis cannot be used to estimate statistics such as the mean and variance of the model output. To estimate statistics such as these, an assessment of uncertainty in model inputs must be given and combined with computer model runs (using e.g. Monte Carlo simulation). Without a clear and quantitative specification of uncertainty in model output, it is very difficult to make good decisions using computer models because the relative probabilities of different possible model outputs are not known.

Using a scenario analysis to investigate the effect on model output of varying a single input of interest or a set of inputs of interest (i.e. a sensitivity analysis) can also be problematic. Firstly, it is common for scenario analyses to focus only on sensitivity to a single input and the effect of interactions between inputs is not considered [11]. This is problematic when the effect of varying an input is different as other inputs vary. Secondly, a sensitivity analysis is usually carried out by performing several model runs, varying some input of interest in each one whilst keeping all other inputs fixed at some estimated value. The variation in output as this input of interest is varied is used to assess the sensitivity of the model output to the input of interest. In the real-world, the inputs held fixed are usually uncertain and this uncertainty should be accounted for to get an accurate view of the effect on model output of varying any given parameter. To account for this uncertainty fully it is necessary to probabilistically specify the uncertainty over

the inputs held fixed and to compute the impact of this uncertainty on the model output when performing a sensitivity analysis (e.g. by using Monte Carlo sampling combined with runs of the computer model).

In [11], probabilities are attached to scenarios in a stochastic optimisation to investigate the effect of mid-term uncertainty on short-term decisions in a two-stage stochastic version of UK-MARKAL. In [12] a method is described for eliciting these probabilities. Probabilistic scenario-based analyses have also been used to study uncertainty in multi-stage problems (e.g. [13] and [14]). Whilst these methods incorporate probabilities alongside scenarios, they still restrict consideration to a small number of scenarios and so do not give a full picture of output uncertainty. Traditional Monte Carlo simulation can be used in combination with a probability distribution over the input parameters of a computer model to give a full picture of uncertainties in the outputs of a computer model (e.g. [15,16]), but for models which take even a few minutes to run it is usually not possible to perform enough computer model runs to obtain a robust and accurate assessment of uncertainty in model outputs. Monte Carlo simulation can be combined with scenario-reduction techniques (e.g. [17]) to overcome this problem, but the effectiveness of this approach is reliant on an assumption that the reduced scenario space will adequately capture the output uncertainty. The extra uncertainty arising from this assumption is rarely considered.

The methods described above are all used to assess the impact of input uncertainty on the outputs of a computer model. The impact of structural discrepancy is a key part in linking the model and the real-world but it is rarely considered in the energy system planning literature. One example where the importance of this source of uncertainty was highlighted is in [18], where model output is compared across a variety of different energy planning models. Another example is in [19], where the impact of structural discrepancy is assessed by testing the stability of the solution of an optimisation routine to modelling assumptions.

Rather than a scenario-based approach, this paper uses emulation in combination with Monte Carlo simulation and a statistical model for structural discrepancy to assess uncertainty in computer model outputs. The quantification of uncertainty in computer models using emulation for a variety of different non-energy applications is described in [20–23] and [24]. There are few examples of the use of emulation in energy systems modelling. In [6] an emulator of a transmission costs model is used to make decisions about transmission network expansion. Uncertainty in three input parameters is accounted for in this decision making process but there is no consideration of structural discrepancy. In [5] emulation is used to calibrate a generation investment model. Both input uncertainty and structural discrepancy are modelled. Structural discrepancy is estimated by training a statistical model using historical data. In real studies, there may be no historical data available to fit a model for structural discrepancy, or any available historical data may not be relevant to future projections. For the example in this paper, the computer model was only set up to run for future years, so although historic wholesale electricity prices were available, it was not possible to compare this historic data to computer model runs to estimate structural discrepancy. We instead used experimentation to investigate the possible impact of structural discrepancy on future projections, a methodology that can be used in the absence of relevant historical data, or in combination with any available data.

Both [6] and [5] use Gaussian Processes to emulate the computer models studied. In [5], the Gaussian Process is fitted using Bayesian updating, with Markov Chain Monte Carlo (MCMC) used to update the prior distributions of any unknown parameters with training data. These prior distributions are used to describe uncertainty in the parameters of the emulator before any computer model runs have been performed. This Bayesian approach of updating prior distributions has the benefit that expert knowledge about the computer model incorporated in the prior distributions can be combined with data from computer model runs. There are two disadvantages to the Gaussian Process approach described in [5]. The first is that full probability distributions must be specified as prior distributions for all unknown parameters. This is not an easy task for complicated statistical models with many parameters and so many studies are forced to resort to using standard prior distributions without clear justification. As prior distributions can impact results (especially if few computer model runs are possible), it is important to have prior distributions that properly reflect our uncertainty about the unknown parameters. For more details on the difficulties of prior selection, see [25]. The second disadvantage is that MCMC routines can be difficult to implement and may take time to converge. This paper uses Bayes linear methods [26] to fit the emulator to avoid these problems. Bayes linear methods are still Bayesian in nature, and hence retain the ability to update prior information with computer model runs, but require only prior means and variances to be specified for unknown quantities. By using only means and variances, the burden of specifying full probability distributions for all unknown quantities is avoided. Another benefit is that the computations for fitting the emulator are reduced to simple linear algebra and there is no need to use MCMC. Without the need for MCMC, there is no need to assess convergence of the MCMC algorithm and the fitting of the emulator can be done much more quickly.

This paper demonstrates the use of emulation combined with a statistical model for structural discrepancy for overcoming the deficiencies of a scenario based analysis as described above. By using an emulator in combination with Monte Carlo simulation and a probabilistic specification of uncertainty over the computer model inputs a full picture of the impact of input uncertainty can be assessed. As an emulator is quick to evaluate, scenario-reduction techniques are not required. The uncertainty that arises because limited computer model runs are available is quantified using the emulator and incorporated into the uncertainty assessment. The emulator is combined with a statistical model for structural discrepancy, the parameters of which are estimated by experimenting with computer model parameters that are usually set at fixed values. Without a full picture of the uncertainties associated with an energy model it is not possible to make good decisions based on the model output. This paper describes methodology for quantifying these uncertainties in models which are slow to run, and demonstrates this methodology on a widely used model for projecting energy prices.

3. Experimental design

3.1. Parametrisation of input parameters

The model inputs studied were gas price, demand, EU ETS price and offshore wind deployment. All other inputs to the computer model were set to a central scenario chosen by the model developers. We chose to focus on uncertainty arising from these four input assumptions as they were deemed by the model developers to be of most interest to policy-makers, to have a large impact on wholesale electricity prices and to be associated with substantial uncertainty. The impact of these inputs on wholesale electricity prices in GB is discussed in [27]. Uncertainty in the EU ETS price and offshore wind deployment largely arises from uncertainty about future policy, whereas uncertainty in gas price and demand can arise from economic conditions, policy, global political stability and weather conditions. The methods discussed here focus on four model inputs but can be extended in principle to a larger input space. With a larger input space more model evaluations will be needed to fit the
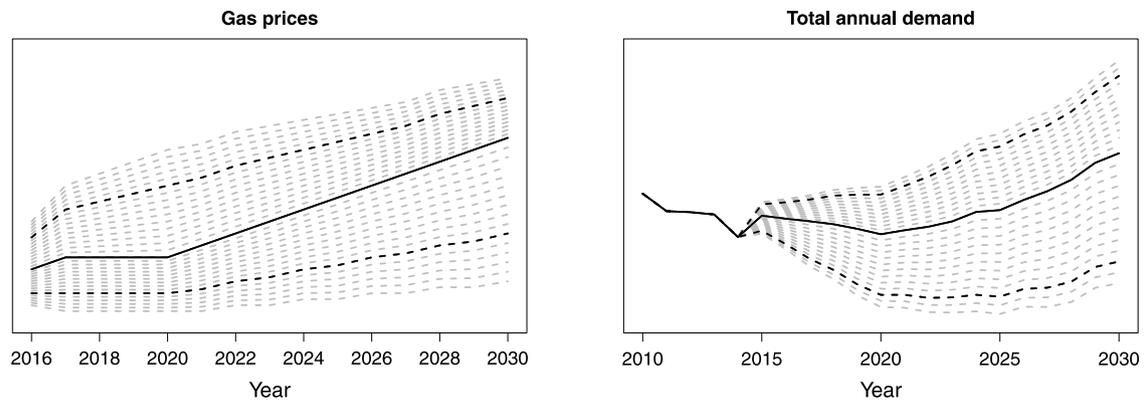
Fig. 1. Central, high (dashed) and low (dashed) demand scenarios in black, parametrisation (possible model inputs) in grey, for gas price (left hand side) and demand (right hand side).
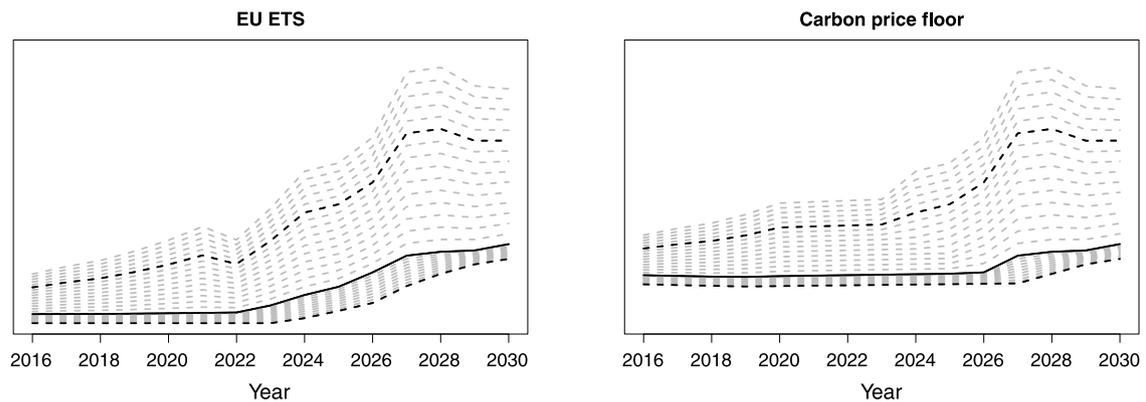

Fig. 2. LHS—central (solid), low (dashed) and high (dashed) scenarios for EU ETS price in black, parametrisation (possible model inputs) in grey. RHS—carbon price floor corresponding to these EU ETS price assumptions.

emulator but the extent of the increase in requirement depends on the sensitivity of the outputs of the computer model to each of the inputs. It is likely that after an initial analysis, some inputs will be found to have little effect on the outputs and it will therefore be possible to exclude them from further consideration. In [28], emulation for input spaces of dimension 50 are discussed and there are examples in the literature of emulators designed for around twenty input parameters (e.g. [24,29]). For very large input spaces, uncertainty arising from any inputs not included in the emulator can be included as an extra error term in the uncertainty analysis (as we demonstrate for coal price in the following section). The size of this extra error term will determine whether any extra investigation into the inputs not included in the emulator is required.

Gas price, demand and EU ETS price take the form of an annual average time series from 2010 to 2030. For offshore wind deployment the model inputs consist of a list of the plants to be deployed in each year from 2016, along with the capacity of each plant (where this deployment relates to future contract for difference auctions [30]). For each input, low, central and high assumptions were available from the model developers. For gas price and EU ETS price, the inputs were known up to and including 2015. For demand, inputs were known up to and including 2014.

As described above, each of the four model inputs is multivariate, so there were a large number of individual parameters to consider when forming the input assumptions. Given the time constraints of the project, it was necessary to reduce the number of parameters studied. Each model input (gas price, demand, EU ETS price and offshore wind deployment) was therefore represented by one parameter: a shift away from the central projection for each input. For all inputs the extent of the shift from the central

projection in each year was determined using the high and low scenarios. A shift of $+1$ from the central projection corresponds to the high scenario. A shift of -1 corresponds to the low scenario. Positive shifts interpolate between the high and central scenarios whereas negative shifts interpolate between the central and low scenarios.

For offshore wind, the central, high and low scenarios consist of the total capacity of offshore wind that will be deployed in future contract for difference auctions. Any offshore wind plants that had already been contracted at the time of analysis were not treated as uncertain. For a given shift parameter, it was assumed that the same number of offshore wind plants would be deployed over the period of interest, but the size of each deployed plant was proportionately adjusted according to the shift parameter. This assumption was made because the aim was to study the effect of variation in the total future capacity of offshore wind rather than the effect of changes in the total number of deployed plants.

The carbon price floor is a minimum price for carbon emissions, currently set as the EU ETS price plus some government determined carbon price support which is known to 2020/21. Dependence of the carbon price floor on the EU ETS price was incorporated into the input assumptions by assuming that the carbon price floor in 2020/21 will increase in line with an assumed Retail Price Inflation (RPI) for all years post 2020/21, unless the carbon price floor is less than the EU ETS price in which case the carbon price floor will be set to the EU ETS price.

Figs. 1–3 show graphically the effect of the parametrisations described above. The *y*-axis labels on these plots and on other plots in this paper have been removed to protect the potential sensitivity of the data. For each plot in Figs. 1–3, the grey lines
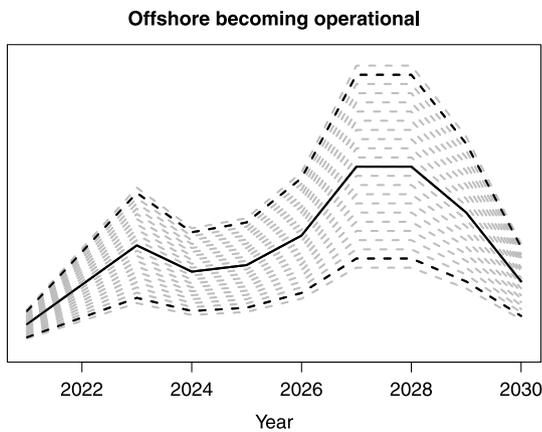
Offshore becoming operational

Fig. 3. Total deployment of offshore wind generation in each year. Central (solid), low (dashed) and high (dashed) scenarios for EU ETS price in black, parametrisation (possible model inputs) in grey.

show possible parametrisations of the model input assumptions. In Fig. 3, the deployed wind plants have been aggregated by the year in which it is assumed they will become operational. The ranges of shift parameters tested were: gas price, $[-1.5, 1.5]$; EU ETS price, $[-1, 1.2]$; demand $[-1.2, 1.2]$ and offshore deployment $[-1.1, 1.1]$. The increased range for gas prices was chosen to reflect that model developers thought gas price was the most influential input assumption. The EU ETS price was capped below at a shift parameter of -1 as shift parameters smaller than this led to negative prices. All other input assumptions in the model were taken to be fixed at the current central estimates used by the model developers.

The effect of parametrising the inputs in this way is that it will not be possible to obtain an emulator prediction for any demand, gas price, EU ETS price or offshore wind deployment projection that cannot be represented as a shift parameter as described above. If it is possible to approximate the input with some shift parameter, then this approximation can be used with the emulator, but an additional error (that of uncertainty in the parametrisation) will be introduced. In Section 5, the effect on wholesale prices of deviations from the gas price parametrisation will be considered.

3.2. Data collection

Three designs (or sets of model evaluations) were set up over the four shift parameters. The first two designs were used for fitting and validating the emulator. The third design was used to assess structural discrepancy. In each design, the values to run for the four input parameters were selected using a maximin Latin hypercube design (see [31] for details and [32] for the R package used to select the design). A Latin hypercube design aims to select model evaluations that are space-filling, rather than focusing the model evaluations on regions of particular interest. This design was chosen so that the emulator could be combined with a wide range of assumptions relating to the probability distribution chosen to represent uncertainty in the inputs.

Brief descriptions of the three designs are given below:

- **Design 1:** 150 model evaluations to use for fitting the emulator.
- **Design 2:** 60 model evaluations to use as a test set to validate the emulator.
- **Design 3:** four sets of 20 model evaluations, used for assessing structural discrepancy. These four sets of model evaluations consist of:

  – Design 3a: 20 model evaluations with the standard parametrisations used for the first and second designs.
  – Design 3b: the 20 model evaluations used in design 3a, but with volatility of $2 \sin(\pi/2 + (\text{year} - 2015)\pi/2)$ added to each year of the gas price assumption. Further discussion of this volatility is given in Section 5.
  – Design 3c: the 20 model evaluations used in design 3a, but with volatility of $8 \sin(\pi/2 + (\text{year} - 2015)\pi/2)$ added to each year of the gas price assumption.
  – Design 3d: the 20 model evaluations used in design 3c, but with a randomly chosen shift factor for coal away from the central assumption. The shift was chosen from a Normal distribution with mean zero and variance set so that high and low coal price scenarios provided by the model developers corresponded to a 95% probability interval around the central assumption.

The model was run for each design point (or model evaluation) described above.

4. Emulation

In this section details of the statistical emulator used to model the wholesale electricity price output as a function of gas price, demand, EU ETS price and offshore wind deployment are given.

The output of the model considered here is an annual time-series from 2010 to 2030 of wholesale electricity prices. To fit an emulator to each of these years independently would be time-consuming and would require a large number of models runs, so principal components analysis (PCA) was used to reduce the dimension of the output space (see [33] for an introduction to PCA, [34] for an example with emulation and [35] for a discussion of multi-output emulation). The aim was to select a small number of principal components that explain as much of the variance in the outputs as possible. Independent emulators were then used to model each principal component.

Before applying PCA, the output for each year was scaled by the mean and standard deviation of the completed runs. The first six principal components were found to explain 98.7% of the total variance, with 78.3% of the total variance arising from the first principal component. Thus, the first six principal components were selected for analysis. The error introduced by discarding the remaining 15 principal components will later be incorporated into the emulator.

Let $f_i(\mathbf{x})$ be the ith principal component of the model output at some vector of inputs $\mathbf{x}$. The value of $f_i(\mathbf{x})$ is unknown at untested $\mathbf{x}$ so we represent our uncertainty in $f_i(\mathbf{x})$ as

$$f_i(\mathbf{x}) = \sum_{j=0}^{p_i} \beta_{ij} h_{ij}(\mathbf{x}) + \epsilon_i(\mathbf{x}), \tag{1}$$

for $i \in \{1, \dots, 6\}$, where $h_{ij}(\mathbf{x})$ are a set of known and deterministic basis functions with $h_{i0}(\mathbf{x}) = 1$, $B = \{\beta_{ij}\}$ are unknown constants and $\epsilon_i(\mathbf{x})$ is a stochastic process.

By including a mean function, given by $\sum_{j=0}^{p_i} \beta_{ij} h_{ij}(\mathbf{x})$, it is possible to incorporate prior judgements about the effect that individual inputs have on the output into the emulator. The basis functions $h_{ij}(\mathbf{x})$ that were chosen for each principal component are given in Table 1. These basis functions were chosen by comparing the coefficients of determination (R^2) of linear regression models fitted to different sets of polynomial and interaction terms formed from $\mathbf{x}$. That is, different possible sets of basis functions were tested by fitting linear regression models to the computer model inputs and outputs. Those basis functions that resulted in a better fit when included in the regression model (as assessed using the coefficient of determination) were retained in the mean function of the emulator. The emulator given by (1) was fitted to the model

Table 1
Basis functions for each principal component, where x_1 = gas price, x_2 = EU ETS price, x_3 = demand and x_4 = offshore wind deployment.

Principal component	Basis functions
1	$x_1, x_2, x_3, x_4, x_2^2, x_3^2, x_1 x_2$
2	$x_1, x_2, x_3, x_4, x_1^2, x_1^3, x_1^4, x_1^5, x_2^2$ $x_3^2, x_3^3, x_3^4, x_1 x_2, x_1 x_3$
3	$x_1, x_2, x_3, x_4, x_1^2, x_1^3, x_2^2, x_3^2, x_2 x_3$
4	$x_1, x_3, x_4, x_1^2, x_1^3, x_1^4, x_1^5, x_2^2, x_3^3, x_1 x_4$
5	$x_1, x_2, x_3, x_4, x_1^2, x_1^3, x_1^4, x_1^5, x_1^6, x_3^2$ $x_3^3, x_1 x_3, x_1 x_4$
6	$x_1, x_2, x_4, x_1^2, x_1^3, x_1^4, x_1^5, x_1^6$

evaluations using Bayesian techniques, requiring specification of prior judgements about all uncertain quantities. For prior beliefs for the coefficients β_{ij}, we set $\mathbb{E}[\beta_{ij}] = 0$ for all i, j, all β_{ij} were assumed independent, and $\text{Var}[\beta_{ij}]$ was set to 1.

The prior mean of $\epsilon_i(\mathbf{x})$ was set to zero and the prior covariance was set to $\sigma_i^2 c_i(\mathbf{x}, \mathbf{x}')$, where $c_i(\mathbf{x}, \mathbf{x}')$ is given by the Gaussian correlation function,

$$c_i(\mathbf{x}, \mathbf{x}') = \exp\left(-\sum_k \frac{(x_k - x_k')^2}{\delta_i^2}\right), \qquad (2)$$

and the δ_i are constants to be specified. Prior independence between $\epsilon_i(\mathbf{x})$ and β_{ij} was assumed. To complete the prior specification, we set $\delta_1 = 0.3$, $\delta_i = 0.1$ for $i > 1$, and σ_i^2 for each i to the residual variance from a standard linear regression fit. The prior beliefs for δ_i were chosen because a reasonable degree of correlation was expected for the emulator for the first principal component, as based on initial fits to training sets it was clear that the mean $\sum_{j=0}^{p_i} \beta_{ij} h_{ij}(\mathbf{x})$ explained much of the variance. For later principal components, the selected basis functions explained a smaller proportion of the variance, so prior expectations were that the local variation around the mean of the emulator would have a lower correlation (i.e. the stochastic process $\epsilon_i(\mathbf{x})$ will be less smooth).

Before fitting the model, the inputs and outputs were scaled to lie between -1 and 1. Bayes linear methods [26] were used to fit (1) to the model evaluations detailed in the previous section. A Bayes linear analysis specifies prior means and covariances for uncertain quantities (β_{ij}, δ_i and σ_i^2), rather than full probability distributions. Given that there is often limited ability to specify prior probability distributions with great detail, using only prior means and covariances can be more natural and results in no meaningful loss of detail. Bayes linear updating equations are used to update these prior specifications with the model evaluations. Using Bayes linear methods to update beliefs in this way can speed up emulator computations in comparison to a Bayesian analysis using full distributions as numerical linear algebra can be used in place of MCMC.

In [5], the author states that the fitting of the emulator (which uses MCMC) takes six minutes for 25 model evaluations and that one emulator evaluation takes 10^{-4} s. Using the Bayes linear updating equations, we were able to estimate the mean and variance of the emulator of the first principal component for the 60 sets of inputs in the test set in approximately 0.3 s for 150 model evaluations (this estimation includes the fitting of the emulator). These two tests are not exactly comparable owing to the different number of model evaluations, a slight difference in the dimension of the input space and different coding platforms but it is clear that even with small variation in the timings, it would be substantially quicker to fit an emulator to the test set using Bayes linear than MCMC. There are many situations where repeated fitting of an emulator is useful. For example, to check the fit of an emulator in the

absence of a test set a leave-one-out analysis might be performed. Each model evaluation would be left out in turn, and the emulator fitted to the remaining model evaluations and used to predict the output for the evaluation left out. For 150 model evaluations, this would require the mean and variance to be computed for one set of inputs using 150 emulators. In these situations, reducing the time taken to fit each emulator from 6 min to less than a second makes a substantial impact on the total required computing time.

To fit an emulator using Bayes linear methods, the Bayes linear updating equations are used to update prior beliefs about the emulator, given a set of model evaluations. As a different emulator is fitted for each principal component, let $D_i = \{(\mathbf{x}_n, y_{in}),$ for $n \in \{1, \ldots N\}\}$ be the data from the N model evaluations used to fit the emulator of the ith principal component. Here, y_{in} is the ith principal component of the output of the nth model evaluation, and $\mathbf{x}_n$ is the corresponding input. The Bayes linear equation for the updated expectation of the ith principal component of the computer model output at untested input $\mathbf{x}$ is given by

$$\mathbb{E}_{D_i}[f_i(\mathbf{x})] = \mathbb{E}[f_i(\mathbf{x})] + \text{Cov}[f_i(\mathbf{x}), D_i](\text{Var}[D_i])^{-1}(Y_i - \mathbb{E}[D_i]), \qquad (3)$$

where $Y_i = (y_{i1}, \ldots, y_{iN})$ is a vector of the ith principal component of the N observed model evaluations. Expectations and variances in (3) are taken with respect to the prior judgements. Similarly, the Bayes linear equation for the updated covariance between two outputs $f_i(\mathbf{x})$ and $f_i(\mathbf{x}')$ is

$$\text{Cov}_{D_i}[f_i(\mathbf{x}), f_i(\mathbf{x}')]$$
$$= k(\mathbf{x}, \mathbf{x}') - \text{Cov}[f_i(\mathbf{x}), D_i](\text{Var}[D_i])^{-1}\text{Cov}[D_i, f_i(\mathbf{x}')], \qquad (4)$$

where $k(\mathbf{x}, \mathbf{x}')$ is the prior covariance between $f_i(\mathbf{x})$ and $f_i(\mathbf{x}')$. All the terms on the right hand sides of Eqs. (3) and (4) can be calculated using the emulator (1) in combination with the prior judgements described above and the data from the model evaluations.

By using Eqs. (3) and (4) the mean and variance of $f_i(\mathbf{x})$ can be estimated for any $\mathbf{x}$. In other words, for any untested input $\mathbf{x}$, a mean and variance for the ith principal component of the computer model at this input can be given. From this, the mean and variance of the wholesale price projection at any untested input can be constructed by reversing the principal component transformation. The variance of this projection describes the uncertainty arising because a limited number of model evaluations have been used to fit the emulator. For more details on fitting emulators of the form given in (1) using Bayes linear methods, see [36,22,21] and [24].

A test set was used to validate the emulator. The emulator (1) was fitted to the 150 model evaluations in the first design, and the ability of the emulator to predict the output for the 60 model evaluations in the second design was tested. Fig. 5 shows the results of this study for each principal component. Almost all of the model outputs lie within the predicted interval for all of the principal components. Out of 60 model evaluations tested, 4, 1, 2, 1, 1 and 2 of the model outputs were outside the prediction interval for each of the six principal components, indicating a good fit of the emulator to the model. In Fig. 5, the size of the error bars increases for each principal component, reflecting the difficulty of finding a good set of basis functions to explain increasingly smaller variations in the output with a fixed number of model evaluations. This is also shown by the sums of squared errors for each of the plots in Fig. 5, which were 0.06, 0.79, 1.11, 1.62, 2.09 and 4.31 for the first, second, third, fourth, fifth and sixth principal components respectively. As described above, the vast majority of the variance was explained by the first principal component. This means that a large emulator error for later principal components will have a small effect on the emulator error of the final wholesale electricity price predictions (as later principal components only contribute a small amount to the overall variance in these projections).

Fig. 4 summarises the steps set out in this section for fitting an emulator to a set of model evaluations. The arrow from the final
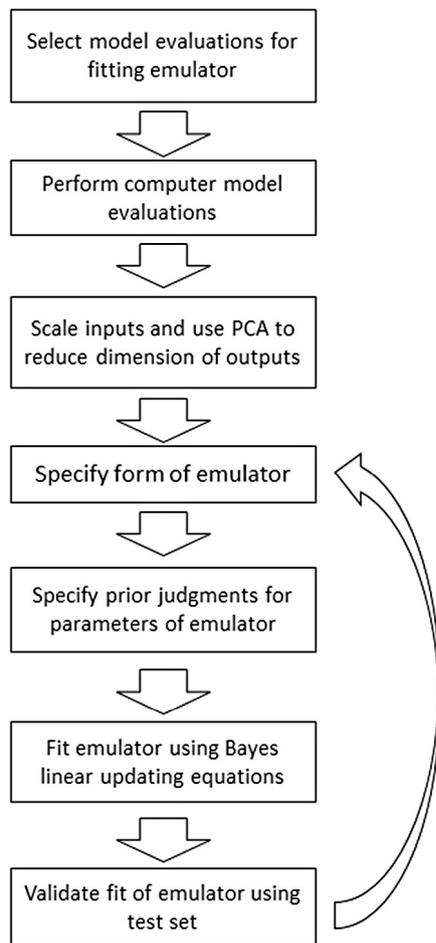
Fig. 4. Flow chart describing steps used to build emulator of computer model.

step to the fourth step demonstrates that fitting the emulator is an iterative process. The form of the emulator should be chosen to get the best fit to the model evaluations possible. If after validating the emulator the fit needs improvement the emulator in (1) can be altered to better fit the model evaluations (for example by investigating different basis functions). Some of the steps in this flow chart will vary for different applications. For example, it may be possible to run extra sets of model evaluations after an initial emulator has been fitted. By doing this, it is possible to use the emulator to investigate which extra model evaluations are likely to be beneficial for improving the emulator. An example detailing this process for a computer model of the Milky Way galaxy can be seen in [24].

The computer code used to fit the emulator described above has been made available as supplementary material. This computer code uses the R programming language in combination with the Rcpp and RcppArmadillo packages to enable integration between R and C++ [37,38]. For confidentiality reasons it was not possible to make available the computer model runs used to fit the emulator for wholesale electricity prices in this paper but code for generating a dataset to test the statistical methodology described here has been included within the supplementary material.

5. Structural discrepancy

The third design in Section 3 was used to study two different sources of structural discrepancy. Designs 3b and 3c were used to investigate the effect of adding volatility to smoothed gas price

assumptions. Design 3d was used to study the effect of varying the fixed annual time series of coal prices. By experimenting with the model using these designs it was possible to investigate the effect of particular modelling assumptions on the wholesale prices. An alternative method for assessing structural discrepancy is to compare model output to real-world data. The model studied here was set-up to be used only for future projections so it was not possible to use the computer model to produce predictions for historic years to compare with any real-world data. Even if this were possible, use of historic structural discrepancy assessments with future wholesale price projections would require an assumption that the past is representative of the future.

As described in Section 3, model inputs were parametrised to be some shift away from the central projection. This parametrisation introduces uncertainty, because in reality the inputs are unlikely to follow the smooth parametrisation. The left hand side of Fig. 6 shows historic annual gas prices from 1998 to 2015. Comparing this to Fig. 1 it is clear that the historic prices are considerably more volatile year to year than the model inputs. To assess the effect of this smoothing, a deterministic effect given by $A \sin(\pi/2 + (\text{year} - 2015)\pi/2)$ for $A = 2$ (design 3b) and $A = 8$ (design 3c) was added to each year of the model input. The gas price inputs tested for $A = 8$ can be seen on the right hand side of Fig. 6. Although the plot shown on the right hand side of Fig. 6 is still not fully representative of the volatility seen on the left hand side of Fig. 6 (because the variation is too regular), these results should still allow for an approximation of the effect on wholesale electricity prices if large jumps are introduced in gas prices year-on-year.

In design 3d, the effect of varying the fixed time series of coal price assumptions was tested. A random shift was added to the time-series of coal price inputs. This shift was defined as for the gas, demand and EU ETS price shifts described in Section 3.

Fig. 7 shows the effect of varying the coal price assumption by plotting the wholesale price projection for 40 possible sets of input assumptions. Lines of the same colour have all input assumptions equal except for coal price. There are 20 colours in the plot and two lines associated with each colour: one of these two lines gives the wholesale price projection when varying the coal price and one gives the wholesale price projection when the coal price is kept at the fixed central scenario. Fig. 8 does the same, but for gas price volatility. Again, the only difference between lines of the same colour is that extra volatility has been added to the gas price. Both plots have been shown with the same scale so that they can be compared. Adding volatility to the gas price clearly has a large effect on wholesale prices in comparison to the effect of varying coal price. The mean absolute difference from 2020 to 2030 was calculated for each set of inputs to compare the effect of varying coal price to the effect of varying volatility in gas price. For coal price, the mean of the mean absolute difference over the twenty sets of inputs was £0.50/MWh, ranging from £0.02/MWh to £3.32/MWh. For gas price volatility, the mean of these mean absolute differences over the twenty sets of inputs was £2.60/MWh, ranging from £2.09/MWh to £4.54/MWh. For coal price variation, the relative size of the range is larger because the variation in coal price is dependent on a randomly drawn shift parameter (as the future coal price is uncertain, whereas we can be reasonably sure that given historical data volatility in gas price is larger than that assumed). Any difference between different sets of input assumptions (i.e. lines of different colour in Fig. 7) can arise both because of this randomly drawn shift parameter and because of interaction between coal price and the other input assumptions. For gas price volatility, it is assumed that future volatility will be approximately equal to historic volatility, so any difference in effect between different sets of inputs can be attributed only to interactions between gas price volatility and the other input assumptions.
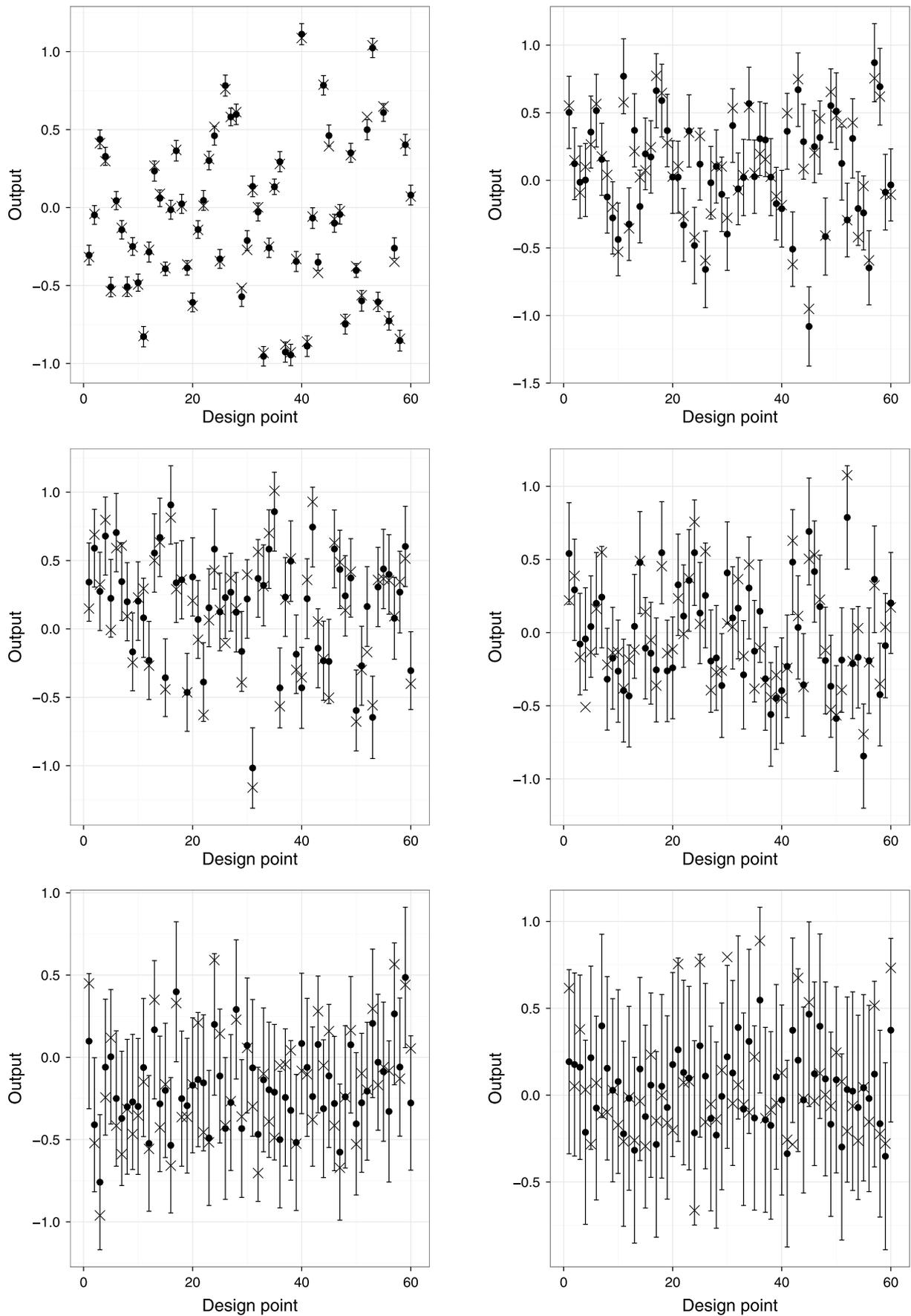
Fig. 5. Validation of the emulator for each principal component in order: 1 (left, top), 2 (right, top), 3 (left, middle), 4 (right, middle), 5 (left, bottom), 6 (right, bottom). Black dot is emulator mean (predicted value) for a given run in the test set, error bars give a 95% probability interval around this mean. Cross shows actual value of principal component for this run.
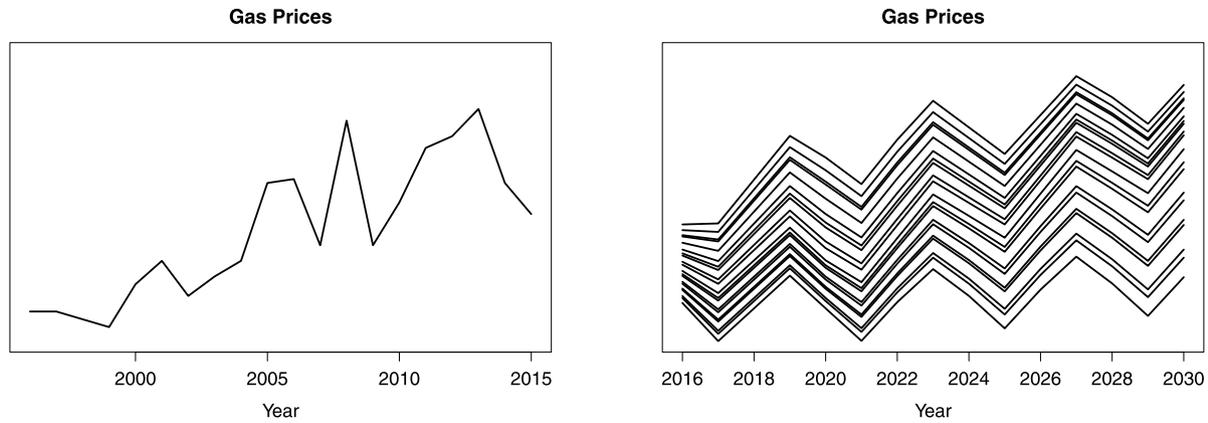
Gas Prices

Gas Prices

Fig. 6. Left hand side—Historic annual gas prices (adjusted for inflation) [39]. Right hand side—volatile gas price inputs tested for design 3c (amplitude 8). Both charts have the same scale.
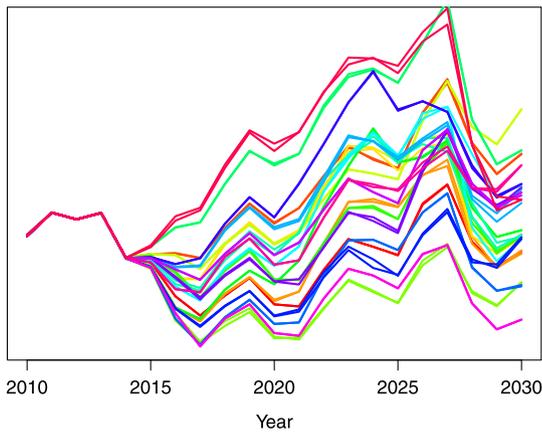

Fig. 7. Wholesale electricity price projections for the model evaluations in designs 3c and 3d, with and without variation in coal price. Each colour is associated with two wholesale price projections, one with variation in coal price and one using the standard central assumption. Lines of the same colour are associated with the same input assumptions (except for coal price). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
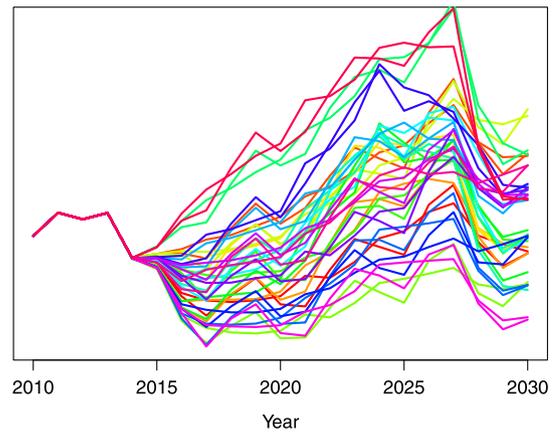
Fig. 8. Wholesale electricity price projections for the model evaluations in designs 3a and 3c, with and without gas price volatility of amplitude 8. Each colour is associated with two wholesale price projections, one with volatility in gas price and one using the standard central assumption. Lines of the same colour are associated with the same input assumptions (except for gas price). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5.1. Modelling structural discrepancy

To incorporate structural discrepancy relating to coal and gas prices in our uncertainty specification, we model our uncertainty in the wholesale price time series $\mathbf{w}(\mathbf{x})$ as

$$\mathbf{w}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T L^T + \boldsymbol{\epsilon}_{PCA} + \boldsymbol{\epsilon}_g + \boldsymbol{\epsilon}_c, \tag{5}$$

where

- $\mathbf{f}(\mathbf{x})$ is a 6-dimensional vector where each dimension $f_i(\mathbf{x})$ is the emulator modelling the ith principal component of the wholesale electricity prices at input $\mathbf{x}$ (given by (1)).
- L is a 6×21 dimensional loadings matrix used to transform the principal components to wholesale prices.
- $\boldsymbol{\epsilon}_{PCA}$ is a 21-dimensional random vector with zero mean and covariance matrix $I\sigma_{PCA}^2$, where I is the identity matrix and σ_{PCA}^2 is a 21-dimensional vector of constant variance terms. $\boldsymbol{\epsilon}_{PCA}$ represents the uncertainty in each year of wholesale prices due to the emulation of 6 principal components rather than the full 21.
- $\boldsymbol{\epsilon}_g$ is a 21-dimensional vector with zero mean and covariance matrix $I\sigma_g^2$, representing the uncertainty due to smoothed gas price inputs.

- $\boldsymbol{\epsilon}_c$ is a 21-dimensional vector with zero mean and covariance matrix $I\sigma_c^2$, representing the uncertainty due to use of a fixed times series of coal price inputs.

In (5) we assume that $\boldsymbol{\epsilon}_{PCA}$, $\boldsymbol{\epsilon}_g$ and $\boldsymbol{\epsilon}_c$ do not depend on the input $\mathbf{x}$. We also assume that $\boldsymbol{\epsilon}_{PCA}$, $\boldsymbol{\epsilon}_g$ and $\boldsymbol{\epsilon}_c$ are independent of one another and of $\mathbf{f}(\mathbf{x})$. The 21 components in each of $\boldsymbol{\epsilon}_{PCA}$, $\boldsymbol{\epsilon}_g$ and $\boldsymbol{\epsilon}_c$ are assumed to be independent of one another. It is likely that some, if not all, of these assumptions do not hold in practice. Informal evidence that errors are correlated year-on-year can be seen in Fig. 7. When the error is positive in one year, it is generally positive in the next. Given the small size of the errors in comparison to the overall variation in wholesale prices due to different input assumptions, this assumption of independence is likely to have a negligible impact on the results, but further investigation is required to confirm this.

Values for σ_c^2 and σ_g^2 were estimated using the sample variances of the differences in each year when adding coal price variation and gas price volatility respectively. To reduce the effect of the small sample size (and to remove the periodic effects seen in the gas price volatility results), a lowess curve was used to smooth the sample variances through time. A low and high value for each of σ_c^2 and σ_g^2 was estimated. The low value of σ_g^2 corresponds to volatility with

amplitude 2 and the high value of σ_g^2 to amplitude of 8. One of the model evaluations was found to have a very large deviation in years 2028, 2029 and 2030 when coal price was varied. The low value for σ_c^2 was estimated removing this model evaluation. The high value included this model evaluation in the variance estimation. This outlying model evaluation provides some evidence that the error terms in (5) may not be independent of $\mathbf{x}$.

The estimate of the variance of σ_{PCA}^2 was set using results in [40]. The variance for each year was set to the total variance not explained by the first six principal components (which was 0.28 or 1.3% of the overall variance) divided by the difference between the number of dimensions (21) and the number of principal components (6), giving a variance in each year of 0.0189. This variance is of the scaled principal components and translates to a standard deviation of approximately 0.15 in 2015 up to 1.33 in 2030 in the wholesale prices.

The uncertainty specification in (5) accounts for uncertainty in wholesale prices at untested $\mathbf{x}$ (through the emulator), uncertainty arising from the use of PCA and structural uncertainty arising from using smoothed gas prices and a fixed series of coal prices. Other sources of structural discrepancy will exist and have not been assessed here, but the methodology described provides a proof of concept for further study of structural discrepancy.

6. Uncertainty and sensitivity analysis

In this section the uncertainty specification in (5) is combined with a probability distribution over the four model inputs to investigate uncertainty in wholesale electricity price projections. The emulator used for this investigation was fitted using all 210 model evaluations, i.e. the 150 in the first design combined with the 60 in the test set.

A truncated multivariate Normal distribution over the four shift parameters was used to represent uncertainty in the input parameters. The truncation was used so that the EU ETS price remained above zero. The mean, variance and correlations of the multivariate Normal distribution (where x_1 = gas price, x_2 = EU ETS price, x_3 = demand and x_4 = offshore wind deployment) were set to:

$$\mathbb{E}[X] = (0, 0, 0, 0)^T, \ \text{Var}[X] = (a^2, a^2, a^2, a^2)^T$$
$$\text{Corr}(x_1, x_3) = b, \ \text{Corr}(x_2, x_3) = c, \ \text{Corr}(x_i, x_j) = 0$$
$$\text{for all other pairs of inputs}, \tag{6}$$

for constants a, b and c. The values of a, b and c were varied to investigate the effect of these parameters on the overall uncertainty. Setting $a = 0.5$ corresponds to the belief that the high and low scenarios approximately form a 95% probability interval, with the mean given by the central projection. Correlations were set based on discussions with the model developers and represent the interaction between different inputs. For example, setting a value close to one for b would imply that a high gas price is associated with high demand. It is possible to combine any joint distribution over the inputs with the emulator to assess the effect of this uncertainty on the wholesale prices.

The R package [41] was used to draw a Monte Carlo sample of size 3000 from the distribution over the inputs given in (6). The expected value of the emulator at each of the 3000 Monte Carlo samples and for each of the 6 principal components was then obtained using (1) and the Bayes linear updating Eq. (3). Similarly, the 3000×3000 covariance matrix for each principal component, giving the covariance of the emulator for all pairs of inputs in the Monte Carlo sample was estimated using (1) and the Bayes linear updating Eq. (4). A multivariate Normal distribution with this mean and covariance matrix was used to approximate the joint distribution of the emulator for each principal component over the Monte Carlo sample. This joint distribution describes the uncertainty arising from use of an emulator rather than the underlying

computer model itself. A single draw from the joint distribution can be thought of as a simulation of the 3000 outputs associated with the 3000 inputs in the Monte Carlo sample. Samples of size 5000 were then drawn from the multivariate Normal distributions for each principal component. The principal components were then transformed back to wholesale electricity price projections. Each of the 5000 samples then consisted of 3000 wholesale electricity price projections, one for each of the Monte Carlo draws over the input distribution given by (6). The 3000 samples over the input distribution incorporate parametric uncertainty, whereas the 5000 samples of the emulator account for the uncertainty in the output of the computer model at the 3000 draws tested. These wholesale electricity price projections could then be used to estimate the mean and variance in each year arising both from parametric uncertainty over distribution (6) and from uncertainty over the distribution of the emulator (1). The uncertainty specification given in (5) was used to add structural discrepancy to these variances.

In the Monte Carlo procedure described above, the uncertainty arising from use of an emulator is explicitly modelled. It is important to model this uncertainty because the actual principal components of the computer model output at each of the Monte Carlo draws are unknown and have been estimated using the emulator. This estimation is necessary because the time taken to run the underlying computer model means that the computer model itself cannot be run at each of the Monte Carlo draws. An emulator makes it possible to quantify the uncertainty that arises because the computer model itself cannot be used. Without quantifying this uncertainty using an emulator, the uncertainty would not be modelled and hence it would be impossible to assess the consequences of this uncertainty on the wholesale prices.

Fig. 9 shows the results of the uncertainty analysis for $a = 0.5$, $b = 0.3$, $c = 0.3$, split by the errors included in the analysis. In each plot, the mean of the projections is shown by a solid black line. This black line represents the mean wholesale price projection integrated over uncertainty in the input parameters, structural discrepancy and uncertainty arising from the limited number of model evaluations (i.e. uncertainty due to use of an emulator rather than the underlying computer model). The uncertainty in this mean, caused by use of an emulator rather than the model itself to do the analysis, is shown as a 95% probability interval with dashed black lines. The 95% probability interval incorporating uncertainty in the input assumptions as well as structural uncertainty due to gas price volatility and coal price variation is shown with solid grey lines. The uncertainties in these approximations of the variance, again caused by use of an emulator, is shown as a 95% probability interval with dashed grey lines. In all plots, the means and variances are calculated individually for each year (i.e. they are pointwise means and variances). There is no uncertainty shown up to and including 2014, as all inputs are known to then. Given the time taken to perform individual runs of the computer model, it would not be possible to produce the plots shown here without using an emulator to approximate the computer model output.

In Fig. 9, the effect of incorporating ϵ_{PCA}, ϵ_g and ϵ_c into the analysis can be seen. In the upper left plot, none of these errors are included in the analysis. As can be seen, the uncertainty bounds (both solid lines and dashed lines) are smaller. In the other plots, the incorporation of ϵ_{PCA} in the analysis increases the uncertainty due to emulation (i.e. the black dashed lines). The size of these uncertainty bounds could be reduced by incorporating more principal components. Including ϵ_g and ϵ_c causes the probability interval given by the solid grey lines to increase. There is little difference between the upper right and lower left plots, suggesting that it is only the high variance assumptions for coal and gas volatility that have an impact on results.

In Fig. 10, the effect that the values of a, b and c have on the wholesale price projections is investigated. Increasing the correlation between gas and demand (i.e. the value of b) slightly increases
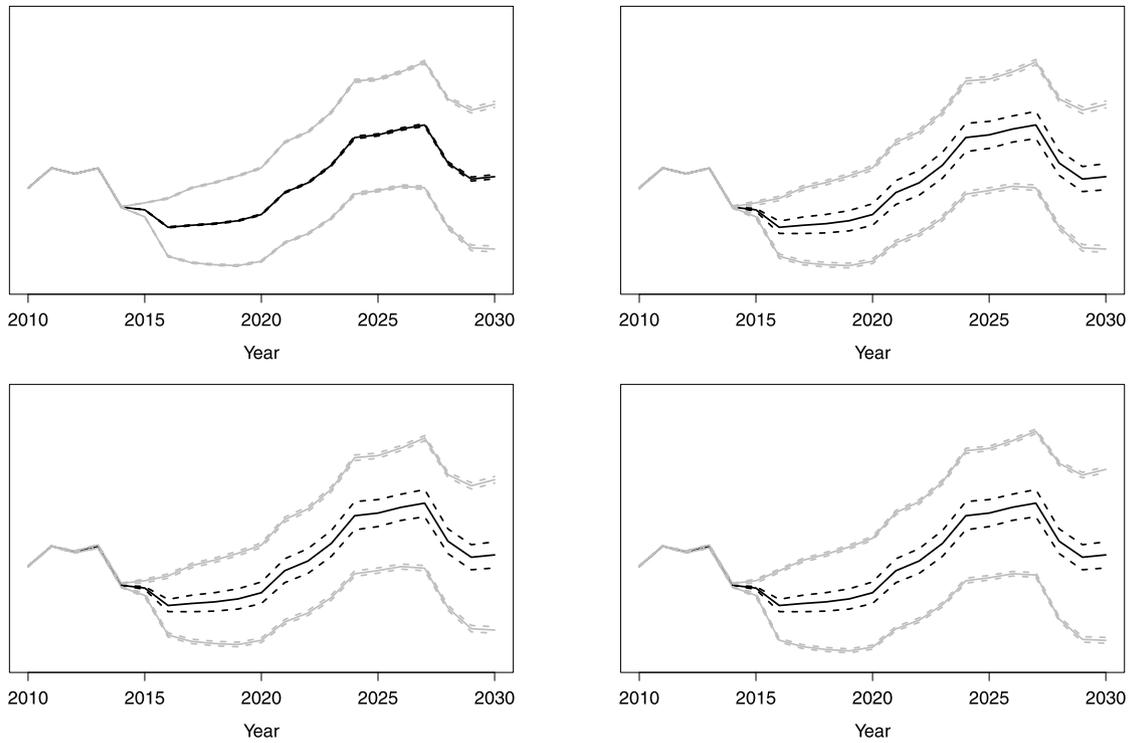
Fig. 9. Estimated mean wholesale electricity price projection (solid black line) with 95% probability interval accounting for emulator uncertainty (dashed black lines), estimated variance due to gas price, demand, EU ETS price, offshore deployment and model discrepancy (solid grey line) with 95% probability interval accounting for emulator uncertainty (dashed grey lines). Model discrepancies considered: none (upper left); PCA approximation (upper right), low coal price variance, low gas price volatility and PCA approximation (lower left), high coal price variance, high gas price volatility and PCA approximation (lower right). The scale of the *y*-axis is the same across all graphs.
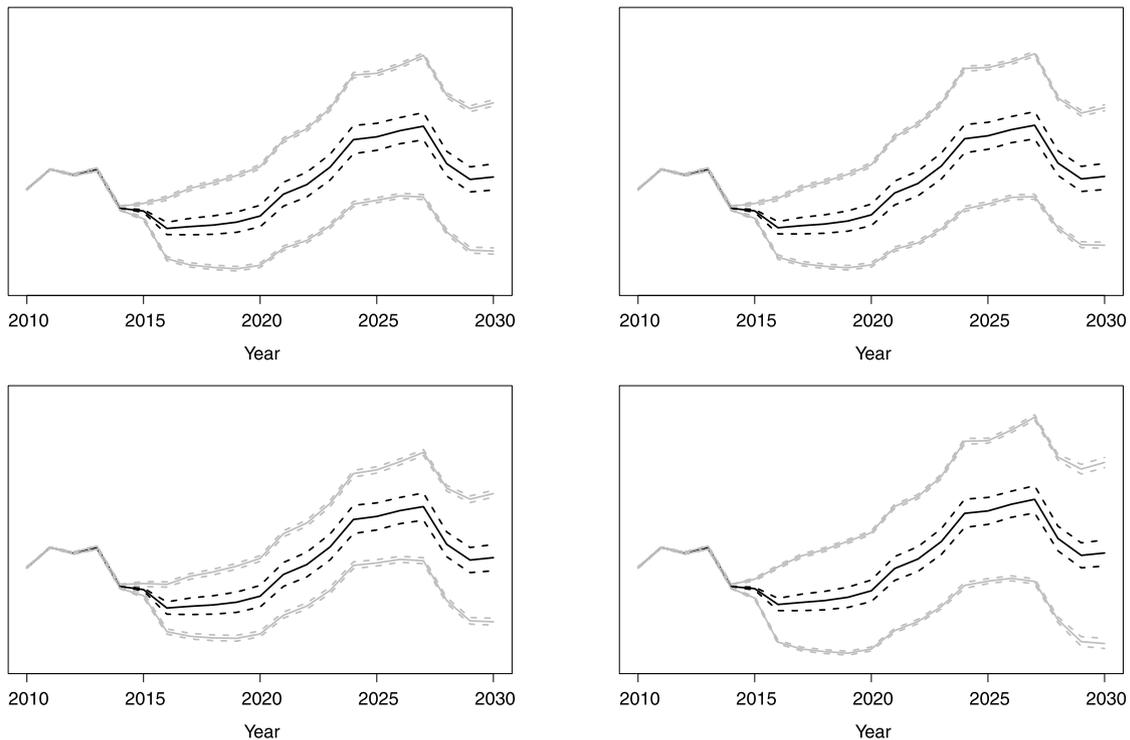

Fig. 10. Estimated mean (solid black line) with 95% probability interval accounting for emulator uncertainty (dashed black lines), estimated variance due to gas price, demand, EU ETS price, offshore deployment and model discrepancy (solid grey line) with 95% probability interval accounting for emulator uncertainty (dashed grey lines). In each plot the parametric uncertainty given in (6) has been set in the following way: $a = 0.5$, $b = c = 0$ (upper left), $a = 0.5$, $b = -0.3$, $c = 0.3$ (upper right), $a = 0.4$, $b = c = 0.3$ (lower left), $a = 0.6$, $b = c = 0.3$ (lower right). All have low coal price variance, low gas price volatility and PCA approximation error included. The *y*-axis scale is the same across all graphs.
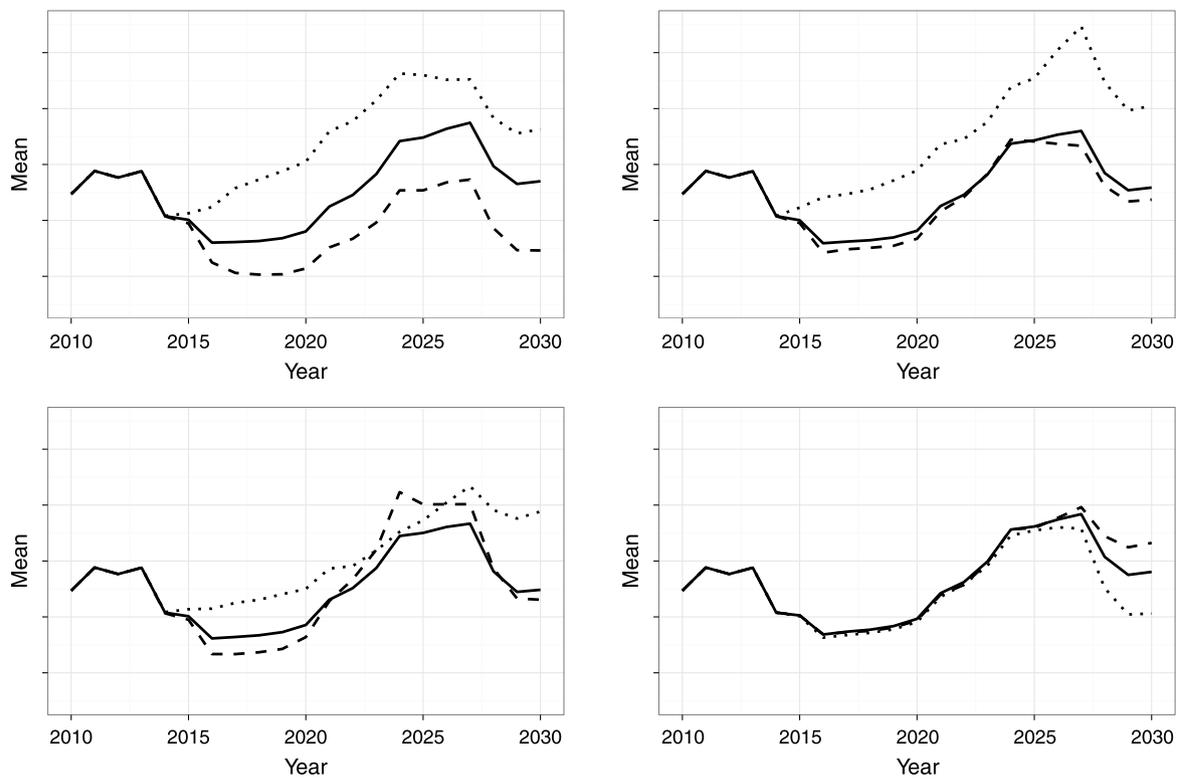
Fig. 11. Estimated mean wholesale electricity prices curve for fixed gas price assumption (top left), EU ETS price (top right), demand (bottom left) and offshore deployment (bottom right). For gas price, demand and offshore deployment the fixed values are −1.2 (dashed), −0.1 (solid) and 1.2 (dotted). For EU ETS price, the fixed values are −1 (dashed), 0 (solid) and 1.2 (dotted). In all cases, the mean is estimated by integrating over uncertainty in all other parameters, where this uncertainty is specified by the conditional distributions of (6). The scale of the y-axis is the same across all graphs.

the overall variation in wholesale prices. A bigger increase is seen when increasing the assumed standard deviation of the model inputs (a). This suggests that careful consideration of the standard deviation of model inputs is needed to ensure an appropriate representation of uncertainty in future wholesale electricity prices.

The effect that each of the four shift parameters has on the model output was investigated by holding each individual parameter fixed at a range of values and integrating over uncertainty in the remaining parameters using the Monte Carlo procedure described above. The uncertainty in the remaining parameters is specified by the conditional distributions of (6), with $a = 0.5$, $b = c = 0.3$. Draws from these conditional distributions were made using results in [42]. The same process can be used to investigate sensitivity to groups of inputs, by holding several inputs fixed at once and integrating over uncertainty in the remaining inputs.

The mean wholesale electricity price curves for a range of fixed values of the four shift parameters are shown in Fig. 11. For gas price, demand and offshore deployment, the plots show the mean wholesale electricity price projection when the shift parameter in question is fixed to −1.2 (dashed line), −0.1 (solid line) and 1.2 (dotted line). For the EU ETS price, the shift parameter is fixed to −1 (dashed line), 0 (solid line) and 1.2 (dotted line). A variety of values were tested and those displayed were selected as being generally representative of the behaviour of the wholesale electricity price time series as each input parameter varies.

Fig. 11 shows that each shift parameter has a different effect on the shape of the modelled mean wholesale prices curve. The biggest effect is seen with the gas price, which seems to shift the wholesale prices curve, keeping the shape roughly consistent for all shift parameters. High EU ETS prices have a similar effect up to around 2025 but after then the wholesale prices continue rising instead of flattening off. Demand and offshore wind seem to have an impact on the extent of the modelled price drop post 2027,

with the amount of offshore wind having little effect before then. Demand also seems to have an effect on whether the price drops post 2015.

Assessing uncertainty in wholesale electricity price projections as demonstrated in this section allows better real-world decisions to be made. Consider for example the decision of whether to invest in a new thermal generating plant. This decision would be made based on projected future cash flows, which would incorporate assumptions about the future wholesale electricity price throughout the lifetime of the plant. Using a scenario-based analysis for decision-making would restrict consideration of the wholesale prices to a small set of future scenarios. The method described here instead estimates a central projection (i.e. the solid black lines in Figs. 9 and 10) which incorporates parametric uncertainty over all possible scenarios, as determined by the probability distribution over the input space. The risk arising from this parametric uncertainty and from structural discrepancy is also assessed. Making a decision based only on the central wholesale price projection would ignore these risks and could result in the building of a plant when there is a high probability that this plant will not be profitable.

7. Conclusion

This paper has presented a study of uncertainty associated with a complex computer model used for the projection of wholesale electricity prices. Three sources of uncertainty were considered. The first was parametric uncertainty, namely uncertainty in four model inputs (gas price, demand, EU ETS price and offshore wind deployment). The second source of uncertainty considered was structural discrepancy, arising because the model used is an approximation of the real-world and so even at some best set of inputs, there will still be some discrepancy between the model and

the real-world. Specifically, structural discrepancy arising from the use of smoothed gas price projections (which in reality are volatile) and from treating coal prices as known was studied. The third source of uncertainty considered was uncertainty as to the value of the model output at untested model inputs. As the model takes one hour to run, assessing parametric uncertainty necessarily requires an approximation of the model output at a large collection of model inputs. A model for the error associated with the approximation is required to fully account for the extra uncertainty this approximation introduces.

An emulator, fitted to 210 model evaluations, was used to model uncertainty in model outputs at inputs that had not been tested. This emulator was shown to have a good fit to the model when validated using a test set. A small number of model evaluations were used to assess the size of the structural discrepancy relating to the use of smoothed gas price projections and from fixing the coal price input assumptions. It was found that introducing volatility into the gas price input assumptions has a larger effect on wholesale electricity price projections than varying the fixed time series of coal price assumptions. The emulator was then combined with a probability distribution over the model inputs and a model for structural discrepancy to investigate the impact that these uncertainties have on wholesale electricity price projections.

The quantification of uncertainty associated with the output of a complex computer model is a critical step that needs to be taken when real-world decisions are being made based on this output. Without having a full picture of the uncertainties associated with a model output, it is not possible to relate this output to the real-world, and hence to make decisions which are relevant for the real-world. This paper has provided a case study for the quantification of uncertainty in projections of wholesale electricity prices, demonstrating steps that can be taken to quantify uncertainties logically and consistently, even when time constraints restrict the number of available model runs.

Acknowledgements

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.segan.2017.11.003.

References

[1] The EERA joint programme in energy systems integration, http://www.eera-set.eu/wp-content/uploads/JP_ESI_Description_of_Work_Dec-7th-2015.pdf .(Accessed 31 July 2017).

[2] L.M. Hall, A.R. Buckley, A review of energy systems models in the UK: Prevalent usage and categorisation, Appl. Energy 169 (2016) 607–628.

[3] Energy and emissions projections 2016, https://www.gov.uk/government/publications/updated-energy-and-emissions-projections-2016. (Accessed 01 August 2017).

[4] Select Committee on Economic Affairs, The price of power: reforming the electricity market, 2017. https://www.publications.parliament.uk/pa/ld201617/ldselect/ldeconaf/113/113.pdf. (Accessed 17 June 2017).

[5] M. Xu, A.L. Wilson, C.J. Dent, Calibration and sensitivity analysis of long-term generation investment models using Bayesian emulation, Sustainable Energy, Grids Netw. 5 (2016) 58–69.

[6] A. Lawson, M. Goldstein, C.J. Dent, Bayesian framework for power network planning under uncertainty, Sustainable Energy, Grids Netw. 7 (2016) 47–57.

[7] M.-C. Hu, B.F. Hobbs, Analysis of multi-pollutant policies for the U.S. power sector under technology and policy uncertainty using MARKAL, Energy 35 (12) (2010) 5430–5442.

[8] D. Eager, B. Hobbs, J. Bialek, Dynamic modeling of thermal generation capacity investment: Application to markets with high wind penetration, IEEE Trans. Power Syst. 27 (4) (2012) 2127–2137.

[9] B.F. Hobbs, M.C. Hu, J.G. Inon, S.E. Stoft, M.P. Bhavaraju, A dynamic analysis of a demand curve-based capacity market proposal: The PJM reliability pricing model, IEEE Trans. Power Syst. 22 (1) (2007) 3–14.

[10] W. McDowall, E. Trutnevyte, J. Tomei, I. Keppo, UKERC Energy Systems Theme: reflecting on scenarios, 2014. http://www.ukerc.ac.uk/publications/ukerc-energy-systems-theme-reflecting-on-scenarios.html. (Accessed 17 June 2017).

[11] W. Usher, N. Strachan, Critical mid-term uncertainties in long-term decarbonisation pathways, Energy Policy 41 (2012) 433–444.

[12] W. Usher, N. Strachan, An expert elicitation of climate, energy and economic uncertainties, Energy Policy 61 (2013) 811–821.

[13] A.H. van der Weijde, B.F. Hobbs, The economics of planning electricity transmission to accommodate renewables: using two-stage optimisation to evaluate flexibility and the cost of disregarding uncertainty, Energy Econ. 34 (6) (2012) 2089–2101.

[14] I. Konstantelos, G. Strbac, Valuation of flexible transmission investment options under uncertainty, IEEE Trans. Power Syst. 30 (2) (2015) 1047–1055.

[15] H. Tekiner, D.W. Coit, F.A. Felder, Multi-period multi-objective electricity generation expansion planning problem with Monte-Carlo simulation, Electr. Power Syst. Res. 80 (12) (2010) 1394–1405.

[16] N.E. Koltsaklis, P. Liu, M.C. Georgiadis, An integrated stochastic multi-regional long-term energy planning model incorporating autonomous power systems and demand response, Energy 82 (2015) 865–888.

[17] J.H. Roh, M. Shahidehpour, L. Wu, Market-Based generation and transmission planning with uncertainties, IEEE Trans. Power Syst. 24 (3) (2009) 1587–1598.

[18] O. Edenhofer, K. Lessmann, C. Kemfert, M. Grubb, J. Köhler, Induced technological change: Exploring its implications for the economics of atmospheric stabilization: Synthesis report from the innovation modeling comparison project, Energy J. 27 (2006) 57–107.

[19] J. Price, I. Keppo, Modelling to generate alternatives: A technique to explore uncertainty in energy-environment-economy models, Appl. Energy 195 (2017) 356–369.

[20] J. Oakley, A. O'Hagan, Bayesian inference for the uncertainty distribution of computer model outputs, Biometrika 89 (4) (2002) 769–784.

[21] J.A. Cumming, M. Goldstein, Bayes linear uncertainty analysis for oil reservoirs based on multiscale computer experiments, in: A. O'Hagan, M. West (Eds.), The Oxford Handbook of Applied Bayesian Analysis, Oxford University Press, 2010, pp. 241–270.

[22] P. Craig, M. Goldstein, A. Seheult, J. Smith, Pressure Matching for Hydrocarbon Reservoirs: A Case Study in the Use of Bayes Linear Strategies for Large Computer Experiments, in: C. Gatsonis, J.S. Hodges, R.E. Kass, R. McCulloch, P. Rossi, N.D. Singpurwalla (Eds.), Case Studies in Bayesian Statistics, Vol. 3, Springer, 1997, pp. 37–93.

[23] M. Kennedy, C. Anderson, S. Conti, A. O'Hagan, Case studies in Gaussian process modelling of computer codes, Reliab. Eng. Syst. Saf. 91 (10–11) (2006) 1301–1309.

[24] I. Vernon, M. Goldstein, R.G. Bower, Galaxy formation: a Bayesian uncertainty analysis, Bayesian Anal. 5 (4) (2010) 619–670.

[25] R.E. Kass, L. Wasserman, The selection of prior distributions by formal rules, J. Amer. Statist. Assoc. 91 (435) (1996) 1343–1370.

[26] M. Goldstein, D. Wooff, Bayes Linear Statistics, Theory and Methods, John Wiley & Sons, 2007.

[27] Ofgem, Wholesale energy markets in 2016, https://www.ofgem.gov.uk/system/files/docs/2016/08/wholesale_energy_markets_in_2016.pdf. (Accessed 17 June 2017).

[28] A. O'Hagan, Bayesian analysis of computer code outputs: A tutorial, Reliab. Eng. Syst. Saf. 91 (10–11) (2006) 1290–1300.

[29] J.A. Cumming, M. Goldstein, Small sample Bayesian designs for complex high-dimensional models based on information gained using fast approximations, Technometrics 51 (4) (2009) 377–388.

[30] Electricity market reform - contracts for difference. https://www.gov.uk/government/collections/electricity-market-reform-contracts-for-difference. (Accessed 13 February 2017).

[31] M.D. Mckay, R.J. Beckman, W.J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics 21 (1) (1979) 239–245.

[32] R. Carnell, lhs: Latin Hypercube Samples, R package version 0.13, 2016. https://CRAN.R-project.org/package=lhs.

[33] I. Jolliffe, Principal Component Analysis, Springer, 2002.

[34] D. Higdon, J. Gattiker, B. Williams, M. Rightley, Computer model calibration using high-dimensional output, J. Amer. Statist. Assoc. 103 (482) (2008) 570–583.

[35] S. Conti, A. O'Hagan, Bayesian emulation of complex multi-output and dynamic computer models, J. Statist. Plann. Inference 140 (3) (2010) 640–651.

[36] P. Craig, M. Goldstein, A. Seheult, J. Smith, Bayes linear strategies for matching hydrocarbon reservoir history, in: J. Bernardo, J. Berger, A. Dawid, A. Smith (Eds.), Bayesian Statistics, Vol. 5, Clarendon Press, Oxford, 1996, pp. 69–95.

[37] D. Eddelbuettel, R. François, Rcpp: Seamless R and C++ integration, J. Statist. Softw. 40 (8) (2011) 1–18.

[38] D. Eddelbuettel, C. Sanderson, RcppArmadillo: Accelerating R with high-performance C++ linear algebra, Comput. Statist. Data Anal. 71 (2014) 1054–1063.

[39] BP UK NBP natural gas prices. http://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy/natural-gas/natural-gas-prices.html. (Accessed 17 November 2016).

[40] M.E. Tipping, C.M. Bishop, Probabilistic principal component analysis, J. R. Stat. Soc. Ser. B Stat. Methodol. 61 (3) (1999) 611–622.

[41] S. Wilhelm, B.G. Manjunath, tmvtnorm: Truncated multivariate normal and student t distribution, R package version 1.4-10, 2015. http://CRAN.R-project.org/package=tmvtnorm.

[42] W.C. Horrace, Some results on the multivariate truncated normal distribution, J. Multivariate Anal. 94 (1) (2005) 209–221.