

EU PVSEC PAPER

# Probabilistic evaluation of solar photovoltaic systems using Bayesian networks: a discounted cash flow assessment

Philip A. Leicester<sup>1\*</sup>, Chris I. Goodier<sup>2</sup> and Paul Rowley<sup>1</sup>

<sup>1</sup> Centre for Renewable Energy Systems Technology (CREST), School of Electronic, Electrical and Systems Engineering, Loughborough University, Loughborough, LE11 3TU, UK

<sup>2</sup> School of Civil and Building Engineering, Loughborough University, Loughborough, LE11 3TU, UK

## ABSTRACT

Solar photovoltaic (PV) technology is now a key contributor worldwide in the transition towards low-carbon electricity systems. To date, PV commonly receives subsidies in order to accelerate adoption rates by increasing investor returns. However, many aleatory and epistemic uncertainties exist with regard to these potential returns. In order to manage these uncertainties, an innovative probabilistic approach using Bayesian networks has been applied to the techno-economic analysis of domestic solar PV. Empirical datasets from over 600 domestic PV systems, together with national domestic electricity usage datasets, have been used to generate and calibrate prior probability distributions for PV yield and domestic electricity consumption, respectively, for typical urban housing stock. Subsequently, conditional dependencies of PV self-consumption with regard to PV generation and household electricity consumption have been simulated via stochastic modelling using high temporal resolution demand and PV generation data. A Bayesian network model is subsequently applied to deliver posterior probability distributions of key parameters as part of a discounted cash flow analysis. The results illustrate the sensitivity of PV investment returns to parameters such as PV self-consumption, PV degradation rates and geographical location and quantify inherent uncertainties when evaluating the impact of sector-specific PV adoption upon economic indicators. The outcomes are discussed in terms of the value and impact of this new Bayesian approach in terms of supporting robust and rigorous policy and investment decision-making, especially in post-subsidy contexts globally. © 2016 The Authors. *Progress in Photovoltaics: Research and Applications* published by John Wiley & Sons Ltd.

## KEYWORDS

return on investment; discounted cash flow analysis; PV yield analysis; Bayesian networks; self-consumption; building stock modelling

### \*Correspondence

Philip A. Leicester, Centre for Renewable Energy Systems Technology (CREST), School of Electronic, Electrical and Systems Engineering, Loughborough University, Loughborough, LE11 3TU, UK.

E-mail: p.a.leicester@lboro.ac.uk

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Received 6 July 2015; Revised 21 December 2015; Accepted 4 January 2016

## 1. INTRODUCTION

Photovoltaic (PV) technologies are seen as a potentially valuable means of contributing significantly to a number of energy policy objectives, including those related to environmental impact, security of supply and energy affordability. As such, PV benefits from a range of financial support mechanisms worldwide. In the UK, for example,

these instruments include the Feed-in Tariff (FIT), Renewables Obligation and most recently for larger systems, Contracts for Difference [1]. This support, together with deep technology cost reductions, has attracted significant investment since 2010, resulting in a very rapid increase in the rate of deployment and installed capacity of solar PV systems such that by 2015, the capacity of grid-connected PV in the UK now exceeds 7 GW [2].

Given the relevance of building mounted or building-integrated PV (BMIPV) to issues such as household fuel affordability and peak load management, the impacts of significantly increased PV penetration merit further study. However, as yet unquantified, uncertainties exist with regard to these impacts and benefits including returns on investment. Such uncertainty represents a risk for decision and policy makers, and investors [3].

In the BMIPV sector, sources of uncertainty with respect to solar PV performance are multi-faceted and include technical factors pertaining to the ongoing and long-term efficacy of the installed system, environmental aspects, such as stochastic variations in expected levels of the incident solar resource or ambient temperature and human factors, such as the quantity and timing of electricity consumption for different households. The interactions of such factors ensure that every deployment context is different and makes predictions of how uncertainties propagate within and between domain boundaries (technical, environmental, financial and societal) a multi-disciplinary problem. Therefore, predicting specific outcomes to a high degree of confidence is extremely difficult with certainty [4].

In such contexts, it is possible to create conceptual models (or abstractions of the real world) in order to explore the causal relationships between parameters in a systemic model. Often, approaches such as causal mapping [5] and 'soft system methodology' [6] are qualitative in nature, and although they provide valuable insights into a multidisciplinary domain (and serve to structure the system in question) [7], they do not provide decision and policy makers with the quantitative analysis capability that is often required. As an alternative, a number of deterministic modelling environments exist, which presume mechanistic relationships between parameters. However, using such techniques means uncertainty can only be explored primarily via sensitivity approaches [8]. Furthermore, the complexity of such socio-technical-economic entities as BMIPV systems, together with their inherently stochastic nature, means that determining a mechanistic relationship (empirical or otherwise) between parameters is often difficult or impractical, particularly at the interface between knowledge domains. Secondly, for a sensitivity analysis, the uncertainty is exogenous to the model as each parameter is varied outside of the model's definition.

A number of multidisciplinary research projects have treated this problem by endogenising the uncertainty into the model itself by introducing variables as probability distributions. Moreover, the mechanistic relationship between parameters can be replaced by probabilistic relationships defined by conditional probabilities. These techniques also offer the benefit of qualitative problem structuring methods by incorporating intuitive visualisation in the form of graphical models [9].

The subject of the research presented here has a number of attributes, which are suited to such a combined quantitative approach using probabilistic parameters, including uncertainties in the parameters that define technology performance and those that define the social, economic and environmental context of BMIPV deployment. In this

multidisciplinary problem context, there are numerous relevant probabilistic parameters and a clear requirement to interrelate these between various domains. As such, probabilistic approaches offer a promising modelling solution.

Furthermore, it is often difficult to quantify uncertainties and thus to propagate uncertainties between different knowledge domains in a meaningful manner; one solution is to develop an integrated probabilistic model that endogenises uncertainties in order to elicit meaningful understanding of key performance indicators (KPIs) for BMIPV.

## 2. BAYESIAN NETWORK MODELS: THEORY AND APPLICATION

Bayesian networks (BNs) are one type of probabilistic model that can be applied to address problems of the type under study in this research [10]. A BN model is constructed using a directed acyclic graph (DAG), which is a collection of nodes used to represent system variables linked by directed arcs, as shown in Figure 1. The arcs denote a causal or observational relationship between the variables; and thus, a BN model delivers a representation of a real-world system comprising both qualitative and quantitative components. These components can be learnt using expert knowledge or derived from data using learning algorithms [11].

The BN is therefore an expedient tool for probabilistic integrated modelling. Firstly, the graphical model provides a user-intuitive representation of the variables and the conditional dependencies between them, in a multi-dimensional parameter space; secondly the chain rule (Eqn (1)) provides a method of creating a joint probability distribution (JPD) for the system by learning the constituent marginal and conditional probability distributions; and thirdly, once the BN is defined, observations made at one or more nodes can be used to update the probability distributions at target nodes of interest using belief propagation algorithms [12]. A disadvantage is that the underlying probability theory requires discrete random variables, although in practice, continuous variables can still be used if they are discretised [10].

A BN can be described as a couple  $(G, P)$ , where  $G = (V, E)$  is a DAG with a set of nodes  $V = \{V_1, V_2, \dots, V_n\}$  representing random variables, and  $E$  is a set of directed edges (or arcs), which represent conditional dependencies between them. These are encapsulated as conditional

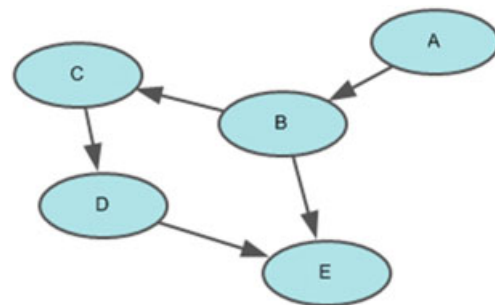


Figure 1. An example of a directed acyclic graph.

probability tables (CPT). It can be shown that the JPD,  $P(V_1, V_2, \dots, V_n)$ , of the BN can be factorised using the chain rule, represented by Eqn (1) [11], where  $\pi_{V_i}$  is the set of parent nodes of node  $V_i$ . Thus, each term in the product is a CPT, or in the case where the set of parent nodes is empty, its marginal distribution.

$$(V_1, V_2, \dots, V_n) = \prod_{i=1}^n P(V_i | \pi_{V_i}) \quad (1)$$

Both exact and approximate algorithms are computationally intensive, an issue that is exacerbated by larger numbers of nodes with larger numbers of discrete states. This can be appreciated when one considers that a BN has  $C_S = \prod_{i=1}^n C_i$  possible system states, where  $C_i$  is the cardinality of node  $V_i$ . Thus, a BN with 10 variables each with 10 discrete states has  $C_S = 10^{10}$  system states. Whilst for a small BN model, an intuitive graphical model communicates the independencies and conditional independencies between variables, for a large model with several dozen or more variables, this oversight is lost. Autonomous BN objects can be connected using common variables to form a larger object-oriented Bayesian network (OOBN) [13]. This allows BNs that model several distinct knowledge domains to be integrated into a single integrated knowledge representation.

An observation applied to one or more nodes is described as evidence, of which there are two key forms. Firstly, an instantiation of a variable  $V$  to state  $v_x$  given evidence  $e$  such that  $P(V=v_x|e)=1$ ; this is referred to as ‘hard evidence’ or a ‘hard finding’. Secondly, the evidence establishes a new probability distribution on the observed variable. This is referred to as uncertain evidence of which there are two types—*likelihood* evidence where the observation is uncertain and *probabilistic* evidence where the evidence specifies the new local distribution [14].

Once evidence has been applied to a variable, probabilistic reasoning algorithms update the probability distributions of all dependent nodes to yield new posterior distributions. This facilitates both prognostic and diagnostic inference, depending on whether evidence is applied to a key input or output parameter, respectively. As such, BNs can be used as a tool for multi-criteria decision support [15] and reasoning under uncertainty. In an OOBN, beliefs are propagated between the constituent BN objects, providing a powerful trans-disciplinary knowledge representation and inference tool [16].

A number of commercial and open-source software applications exist for creating BN models [14]. These provide a graphical user interface for constructing the graphical model, a method to import or learn marginal and conditional probability distributions and features to apply either hard or uncertain evidence to one or more nodes. Belief propagation is automated, and probability distributions are rapidly updated. Uncertainty is by definition endogenised because all variables are input as prior distributions; this renders outputs on nodes of interest with ready-made uncertainty analysis.

In this paper, four BN components of an OOBN are described, which were constructed using the commercial Netica BN software [17]. The design and acquisition of data for these are discussed in the next section.

### 3. OBJECT-ORIENTED BAYESIAN NETWORK MODEL: DESIGN AND CONSTRUCTION

In this analysis, solar PV is considered in the context of domestic buildings and its inhabitants, the attributes of which influence both household electricity consumption (for example, related to the number and the behaviour of the inhabitants) and the PV energy yield (for example, related to the location, geometry and shading characteristics of the building’s roof). These in turn have a direct influence on the amount of self-consumption of PV energy, which is a measure of the quantity of generated electricity consumed within the building. In this context, the concept of four model components is useful, each of which can be represented as an object in an OOBN integrated model (Figure 2). This shows the underlying parameters required as inputs by each BN object, whilst the dependencies between parameters that determine the BN graph structure may be deduced from datasets contained with the literature or further investigated using empirical or numerical modelling methods. The purpose of the OOBN model is to obtain a probabilistic analysis of electricity generation, consumption, self-consumption and export, which produce inputs for a probabilistic analysis of techno-economic indicators such as net present value (NPV), internal rate of return or payback. Figure 2 illustrates the interfaces between the four model components (or objects), via which probabilistic data are propagated between knowledge domains. Note that although we are concerned here with a discounted cash flow analysis, the OOBN objects contain specific parameters (such as gas demand and household income), which facilitate the analysis of additional KPIs such as incidence of fuel poverty [4].

#### 3.1. Building stock sub-model

Specific parameters in the building stock model are known to influence domestic energy consumption and are utilised in a number of building energy consumption predictive or compliance modelling environments such as the UK Standard Assessment Procedure model [18]. For example, the building’s floor area is assumed to be a relatively strong predictor of energy consumption together with the building’s age and built form [19]. Also, the rooftop incident irradiation (and therefore PV system yield) is determined by the geometric relationship between the roof (as defined by its slope and orientation) and the path of the sun integrated over each day of the year [20]. The assemblage of these parameters in an OOBN model is shown in Figure 3. For the OOBN sub-model, spatially disaggregated building stock data were obtained for four small census areas

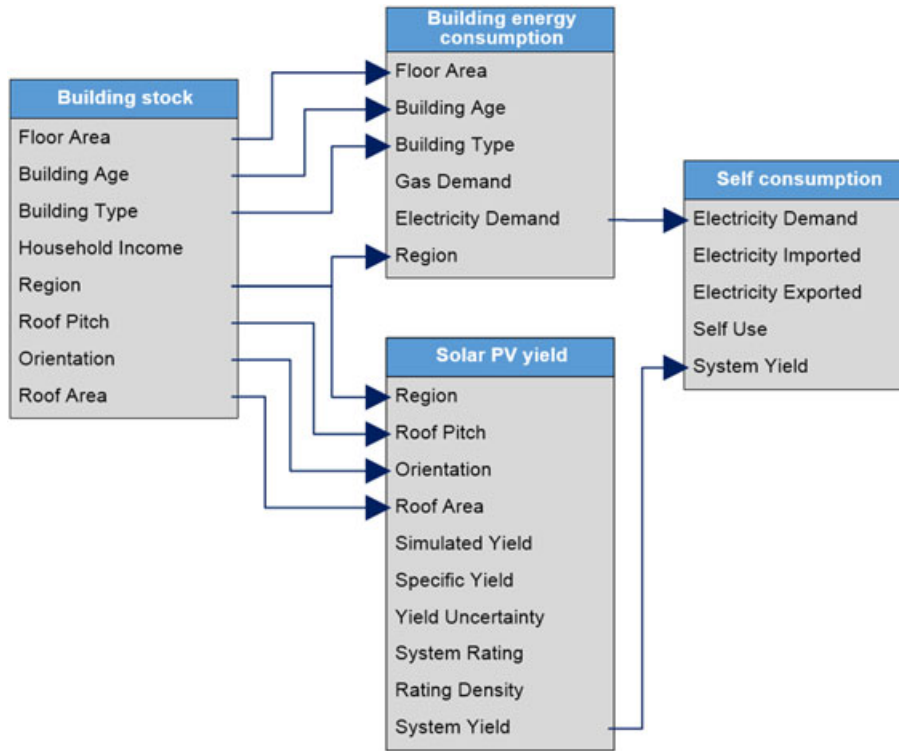


Figure 2. Entity relationship diagram representation showing the object-oriented Bayesian network sub-models and the interfaces between them. PV, photovoltaic.

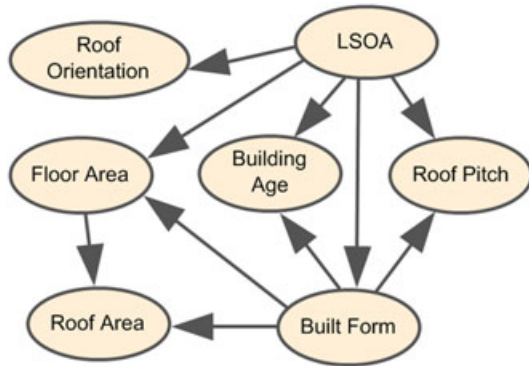


Figure 3. Bayesian network for the building stock sub-model. LSOA, lower super output areas.

derived by the UK office for national statistics known as lower super output areas (LSOAs) [21]. These comprise on average 600 dwellings and 1000 inhabitants [22]. This dataset provides discrete variables for building age and built form (detached, semi-detached, row-houses, flats etc.). Building footprint areas,  $F$ , were obtained from a graphical information system (GIS) dataset of building line elements [23]. These were polygonised and their areas measured using GIS spatial analysis algorithms [24]. The number of unique households within a building polygon,  $N_h$ , was determined by spatially cross referencing with a national address database [25]. The number of storeys,  $S$ ,

for each dwelling was determined by a ‘virtual street walk’ using Google Earth street view [26]. The floor area per domestic household ( $A$ ) is estimated using:

$$A = \frac{F \cdot S}{N_h} \tag{2}$$

Lidar data [27] was used to ascertain roof slope, orientation and area. Subsequently, roof elements were matched with corresponding building data using spatial queries in GIS software to associate roof and building attributes into a unified stock dataset for subsequent application in the OOBN. Not all dwellings possess a suitable roof for PV siting because of such factors as near-object shading, or because the household was part of an apartment block or similar building. A further check using Google Earth aerial photography was used to determine the eligibility of a roof for solar PV; complex hip or intersecting roofs and roofs with dormer windows, skylights or other obstructions were rejected, and the dwelling assigned a zero area roof. Typically, two-thirds of households in each LSOA were found to have a viable roof of an azimuth between 90°(E) and 270°(W) [4].

### 3.2. Solar photovoltaic energy yield sub-model

Using the orientation, slope and geographic co-ordinates of all dwellings in each LSOA, an estimate of annual specific

yield was made using the PVGIS simulation tool using the satellite application facility on climate monitoring (CMSAF) irradiance model [28]. In order to estimate the uncertainty inherent in the PVGIS simulation, calibration was carried out by comparing simulated specific yields for 600 real-world systems for which empirical specific yields were recorded in the Sheffield Microgeneration Database [29]. PVGIS overestimated the empirical yields with a mean bias error of 36 kWh/year/kWp, with a root-mean-square error of 101 kWh/year/kWp. Thus, the estimated yields for the housing stock were modified in the model to reflect the bias and uncertainty using an uncertainty node in the node. This method of accounting for measurement uncertainty is known as the ‘measurement idiom’ [30]. PV system ratings were calculated from roof area data obtained as an output from the building stock sub-model, whilst the rating density function is derived from the aforementioned empirical dataset. The marginal distribution for the latter was determined from metadata for the modules in the Sheffield Microgeneration Database (Figure 4). The roof area and rating density distributions are used to predict the system rating distribution assuming that the whole area of each suitable roof pitch is utilised. Finally, the annual system yield is calculated from the system rating and the annual specific yield. The final BN object is shown in Figure 5.

**3.3. Building electricity consumption sub-model**

Household electricity consumption is predicted by using building attributes from the building stock sub-model. The probabilistic modelling of domestic electricity consumption was conducted in parallel with gas consumption, although only the former analysis is presented here. Although the model includes building attributes to which electricity consumption is only slightly sensitive such as building age and region [19], these have been retained in the simplified electricity consumption BN sub-model shown in Figure 6.

The data to furnish this sub-model with the CPT needed to predict electricity consumption were extracted from the UK’s National Energy Efficiency Data framework anonymised dataset [31]. This consists of records of annual

electricity readings from 2006 to 2012 for over 4m dwellings along with the associated relevant building attributes. These data can be used directly to learn the CPTs using a ‘counting method’ approach [32]. Typically, each node should have no more than three parents because this leads to very large CPTs and requires a large amount of learning data. Fortunately, the large sample size of the National Energy Efficiency Data set facilitated good quality electricity

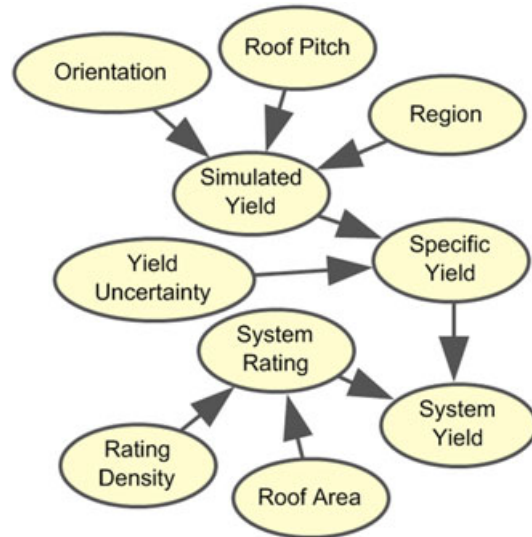


Figure 5. Bayesian network for the solar photovoltaic yield model.

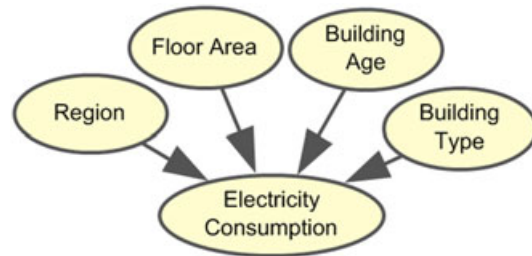


Figure 6. Bayesian network sub-model for electricity consumption.

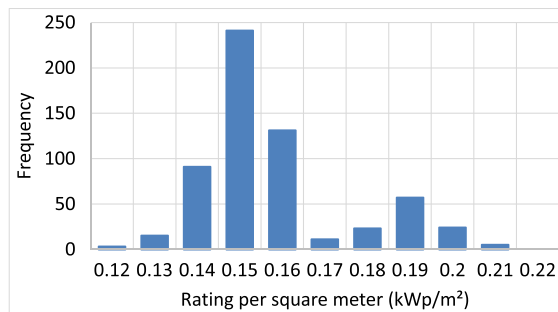


Figure 4. Rating density function of solar photovoltaic modules deployed in the Sheffield Microgeneration Database sample.

consumption distributions in both the resulting prior and posterior distributions following the selection of evidence on building attribute nodes.

### 3.4. Photovoltaic energy self-consumption sub-model

Self-consumption of PV-derived energy arises from the temporal overlap of the instantaneous electricity demand and the generation by the PV system [33]. Both profiles are highly stochastic with large fluctuations occurring on sub-minute timescales [34]. The coincidence of demand and generation at these timescales has to be measured over a significant duration (at least annually) in order to obtain data upon which techno-economic inferences can be based. Furthermore, there is a need to account for the wide variety of electricity demand profiles and installed PV system characteristics in order to be able to construct a valid CPT, which can subsequently be used to predict PV self-consumption from annual consumption and generation data. A number of empirical studies exist (e.g. [35]), but none that provides sufficient data to construct such a probabilistic relationship. To address this, simulated 1-min time-resolution demand and generation data were used to aggregate self-consumption for over 20 000 dwellings for a whole year. These simulations were generated using the a stochastic demand and PV generation model developed by Richardson *et al.* [34], which was modified to reflect a realistic broad variation in electricity demand from 2000 to 10 000 kWh/year whilst varying the PV system rating from 0.5 to 5.0 kWp [36]. The resultant simulated dataset was used to generate the required CPT in the BN sub-model. Nodes for the annual electricity import and export were subsequently added to this sub-model (Figure 7).

Finally, the OOBN was constructed using these four sub-models. Each interconnecting interface between sub-models (Figure 2) was configured such that the probability distribution of the input node  $P(X_i)$  was always equal to that of the output node  $P(X_o)$ , such that  $P(X_i|X_o) = P(X_o)$ . The last component defines the three required target nodes (namely, electricity generation, export and self-consumption), which are used as inputs for a probabilistic discounted cash flow analysis discussed in the next section.

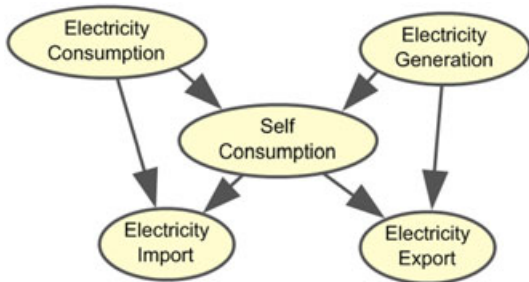


Figure 7. Self-consumption Bayesian network sub-model.

## 4. DISCOUNTED CASH FLOW ANALYSIS

In a discounted cash flow analysis, all future cash flows are discounted to deliver their sum (net value) in the present day known as the NPV (Eqn (3)).

$$NPV = \sum_{n=0}^{\lambda} \frac{V_n}{(1+i)^n}, \quad (3)$$

where  $V_n$  is the net cash flow at time interval  $n$  and  $i$  is the discount rate, which includes the interest rate at which an alternative method of investing the initial sum  $V_0$  could accrue value, and  $\lambda$  is the lifetime of the investment project.

For a domestic PV investment, negative cash flows include the large initial ( $n=0$ ) capital expenditure,  $C_0$ , and subsequent expenditure, during interval  $n$ , for maintenance  $M_n$  and repair  $R_n$  over its estimated lifetime. Positive cash flows arise from the monetisation of generated energy  $E_n$  during each interval. Thus, the net cash flow in interval  $n$  is given by Eqn (4).

$$V_n = E_n - M_n - R_n \quad (4)$$

Under the UK's FIT subsidy scheme,  $E_n$  is given by the sum of the value of generated electricity  $G_n$ , exported electricity  $X_n$  and avoided imported electricity  $A_n$  (Eqn (5)).

$$E_n = G_n + X_n + A_n \quad (5)$$

These values, in year  $n$ , are given in Eqns (6), (7) and (8), where  $Y_n$  is the annual yield,  $F_{G_n}$  is the generation tariff,  $F_{X_n}$  is the export tariff and  $T_{E_n}$  is the electricity tariff during interval  $n$ .  $S$  is the self-consumption fraction.

$$G_n = Y_n \cdot F_{G_n} \quad (6)$$

$$X_n = \frac{Y_n}{2} F_{X_n} \quad (7)$$

$$A_n = Y_n \cdot S \cdot T_{E_n} \quad (8)$$

To account for inflation, FIT rates in the UK are incremented commensurate with the annual Retail Price Index (RPI) [37]. The cost of domestic energy often varies at a different rate than other goods, so a distinct energy inflation rate (EIR) is used [38]. Equations 9 to 11 describe the tariffs in year  $n$ , relative to year 0, assuming average inflation rates.

$$F_{G_n} = F_{G_0} (1 + RPI)^n \quad (9)$$

$$F_{X_n} = F_{X_0} (1 + RPI)^n \quad (10)$$

$$T_{E_n} = T_{E_0} (1 + EIR)^n \quad (11)$$

A further important parameter is the degradation in performance of a PV system,  $d$ , over its lifetime [39]. An unresolved question with regard to the degradation rate is whether it is linear (the same incremental loss in yield each year) or geometric (the same fractional loss each year) [39]. In this analysis, a geometric rate is assumed; thus, if  $Y_0$  is the first year yield, then the yield in year  $n$  is given by Eqn (12).

$$Y_n = Y_0(1 - d)^n \quad (12)$$

Substituting Eqns 6 to 12 in to Eqn (5) yields the income from monetised electricity generation in year  $n$  (Eqn (13)).

$$E_n = Y_0(1 - d)^n \left( F_{G_0} + \frac{F_{X_0}}{2} \right) (1 + RPI)^n \quad (13)$$

$$+ Y_0(1 - d)^n T_{E_0} S (1 + EIR)^n$$

Assuming no additional non-warranted expenditures for repair and maintenance, and neglecting decommissioning costs, the total NPV is given by Eqn (14).

$$NPV = \sum_{n=0}^{\lambda} \left( Y_0 \left( F_{G_0} + \frac{F_{X_0}}{2} \right) \alpha^n + Y_0 T_{E_0} S \beta^n \right) - C_0, \quad (14)$$

where:

$$\alpha = (1 - d)(1 + RPI)(1 + i) \quad (15)$$

$$\beta = (1 - d)(1 + EIR)(1 + i) \quad (16)$$

If  $\alpha$  and  $\beta$  are assumed to be constant over the lifetime of the technology, then the sums of the terms in  $\alpha^n$  and  $\beta^n$  can be represented by sums of geometric progressions [40]. The initial capital expenditure  $C_0$  is replaced by the system rating  $R$  multiplied by the installation cost per unit rating  $C_u$ . Substituting into Eqn (14) yields Eqn (17).

$$NPV = Y_0 \left( F_{G_0} + \frac{F_{X_0}}{2} \right) \left( \frac{1 - \alpha^\lambda}{1 - \alpha} \right) \quad (17)$$

$$+ Y_0 T_{E_0} S \left( \frac{1 - \beta^\lambda}{1 - \beta} \right) - R \cdot C_u$$

The parameters used in Eqn (17) are summarised in Table I. Of these, the self-consumption fraction  $S$ , the initial system yield  $Y_0$  and the system rating  $R$  are represented as probabilistic nodes in the OOBN. Because the dependencies between these parameters are encapsulated by the JPD of the OOBN (shown in Eqn (1)), they can each be taken as independent inputs for a BN sub-model whose JPD is calculated using Eqn (17). Further uncertainties can be endogenised in the model by considering the marginal distributions of remaining parameters, or by treating them as constants, which can be varied for specific modelling scenarios. In

**Table I.** Parameters for net present value calculation in Eqn (17).

Variable	Name	Unit	Type
$S$	Self-consumption fraction	Fraction (per year)	Probabilistic input
$Y_0$	Initial system yield	kWh/year	
$R$	System rating	kWp	
$d$	Annual degradation rate	Fraction (per year)	Probabilistic marginal
$C_u$	System cost	£/kWp	Constant
RPI	Retail Price Index	Fraction (per year)	
EIR	Energy inflation rate	Fraction (per year)	
$i$	Discount rate	Fraction (per year)	
$F_{G_0}$	Initial FIT generation tariff	£	
$F_{X_0}$	Initial FIT export tariff	£	
$T_{E_0}$	Initial electricity tariff	£	
$\lambda - 1$	Life time of technology	Years	

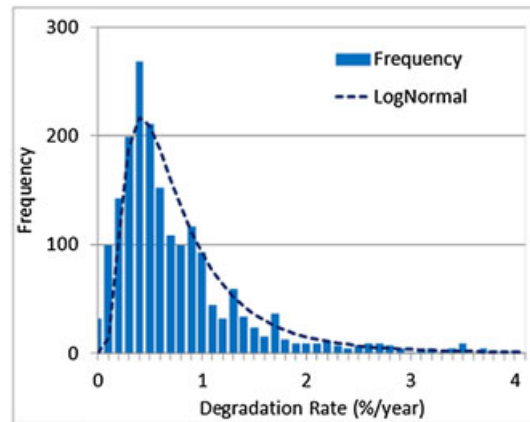
RPI, Retail Price Index; EIR, energy inflation rate; FIT, Feed-in Tariff.

the following section, the uncertainty of the annual degradation rate is described, and values for the remaining parameters (which are treated as constants) are presented.

## 5. RESULTS AND DISCUSSION

### 5.1. Degradation rate analysis

The appropriate inclusion of the PV module degradation rate in the discounted cash flow analysis is of prime importance, given its impact upon PV system lifetime energy yield. Jordan and Kurtz [39] have conducted a review of published literature in which almost 2000 long-term degradation rates for modules or entire systems were assessed to produce the frequency distribution shown in Figure 8. The average and median values for this analysis were 0.8 and



**Figure 8.** Distribution of long-term degradation rates fitted to a lognormal distribution (after Jordan and Kurtz [39]).

0.5%/year, respectively. A qualitatively similar distribution has been used previously to carry out levelized cost of electricity calculations for solar PV using a Monte Carlo approach [41].

Whilst an initial rapid light-induced deterioration of yield over the first few days of exposure is documented in the literature [42,43], the assumption that the annual degradation rate is a gradual process is supported by observations of year-on-year degradation as opposed to catastrophic failures. This long-term gradual decrease in efficiency occurs because of a number of degradation processes caused by thermal and mechanical shocks, and physico-chemical changes, which result in physical damage to module components, and corrosion following humidity ingress [44].

Equation (17) assumes a geometric degradation rate, whereas cited rates are usually calculated as the degradation rate divided by the number of years over which it is measured because insufficient data points are generally available to infer otherwise. This allows a simpler formula to be used (based upon the sum of a geometric progression) rather than that based upon a more complex arithmetico-geometric series. However, the discrepancy between a geometric and an arithmetic (linear) degradation is only 2.2% after 20 years at an annual degradation rate of 1%. The majority of reported degradation rates are less than this, clustered around a value of 0.5%, at which this discrepancy falls to only 0.5%. Thus, given the intrinsic uncertainty in degradation rates, a geometric degradation rate was assumed, and this has been incorporated into the composite discount factors as represented by Eqns (15) and (16).

### 5.2. Photovoltaic cost data analysis

The cost of domestic solar PV reduced rapidly from 2010 (when the UK FIT scheme commenced) from typically £5000 to less than £2000 per kWp in 2015. Mean prices between 2014 and 2015 have fallen from £2229 to £1971 per kWp for systems in the <4-kWp band [45]. The inherent variability of prices by supplier/installers results in a natural distribution of installed cost, but another factor is the way prices are determined; this is based on the notion that some costs are independent of rating (for example, scaffolding costs), thus leading to fixed and marginal costs [45]. The effect of this is to render the price per kWp for smaller systems higher than that for larger systems. With the observed distribution of system ratings skewed towards the larger systems in the band, this results in a positively skewed distribution for cost per kWp (Figure 9). Thus, it is necessary to include both fixed and marginal costs for domestic PV systems in order to account for the higher fixed costs per kWp for smaller systems.

### 5.3. Definition of inflation and discount rates

Both the RPI and the EIR are integral components of the discount factors  $\alpha$  and  $\beta$  used in Eqn (17). In common with parallel national FIT schemes, the former is the UK government's preferred method of incrementing the FIT each

year to account for inflation; and thus, maintain the value of the incentive over its 20-year duration [37]. The RPI consists of a composite index, which measures price variations for a wide range of consumer items. Since 2010, the RPI applied to decrement both the generation and export tariff has ranged from 4.8% to 1.6%, whilst over the same period, the electricity component of the RPI, which measures the percentage change in the price of domestic electricity, has fluctuated from 10.6% to 4.7% [46]. The latter, which is indicative of the EIR, has been extremely volatile over the past 25 years subject to rapid reduction and negative values corresponding to price reductions during the 1990s [47], and a rapid increase between 2003 and 2008 (Figure 10). Equation (18) utilises an average figure for the period of interest, an assumption also made by other researchers [38]. The average RPI between 2014 and 1988 is 3.5% with a standard deviation of 2.0%, whilst the average EIR is 4.3% with a standard deviation of 6.1%.

The NPV calculation should ideally utilise the actual year-on-year inflationary value (RPI or EIR) in each year, rather than an average value over the lifetime of the calculation. Given the high variability, particularly of the EIR, the error that this introduces into Eqn (17) can be deduced.

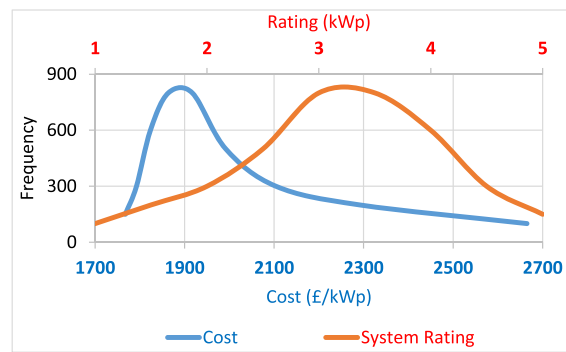


Figure 9. The distribution of cost per kWp for an empirical distribution of UK photovoltaic ratings based on a fixed cost of £1122 and a marginal cost of £1543 for 2014 to 2015 (after Parsons Brinckerhoff, [45]).

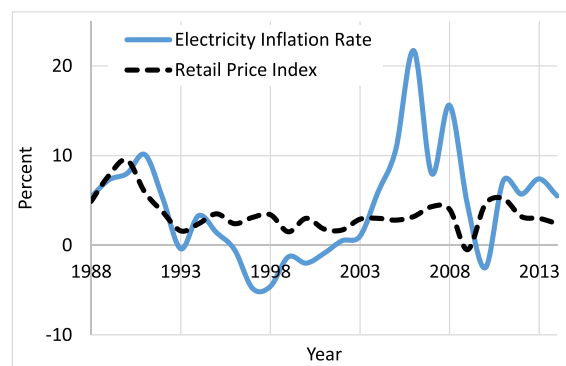


Figure 10. UK Retail Price Index and electricity inflation rate 1988 to 2014.



This is given by the difference between using year-on-year values and calculating the sum of the geometric progression using Eqn (18), or by using an average value as a geometric factor and the formula for the sum of a geometric progression. This is shown in Eqn (19), where  $I_n$  is the RPI or EIR in year  $n$  and  $\bar{I}$  is the average of the index over the period of  $\lambda$  years. A Monte Carlo simulation using a randomised sequence of values for the period 1988 to 2014 gives a mean standard error of 30% for the EIR (standard deviation 15%), the positive value signifying that the exact method using Eqn (18) is higher. For the RPI, the standard error was only 4% (standard deviation 4%).

$$\sum_{n=0}^{\lambda} (A(1 - I_n)^n) \tag{18}$$

$$\frac{1 - (1 + \bar{I})^{\lambda}}{\bar{I}} \tag{19}$$

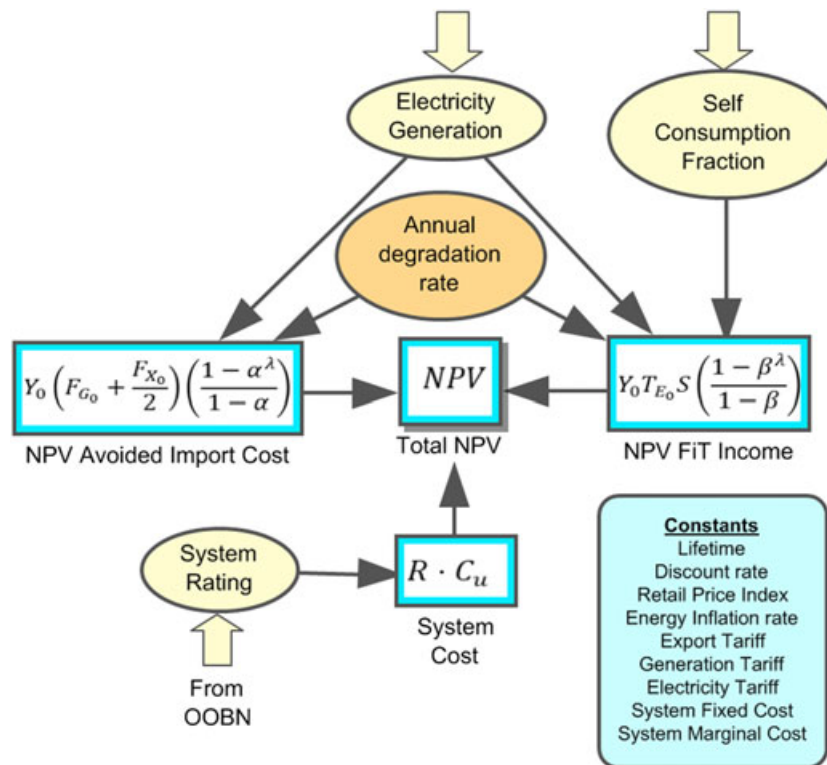
The choice of discount rate is generally contentious because, in private sector investment, it reflects the hurdle rate of return at which a project is deemed viable for investment. To incentivise the adoption of PV using the FIT, the UK's department of energy and climate change (DECC) have set the tariffs to deliver a hurdle rate of 5–8%. In a recent UK Government review, it has been proposed to adjust tariffs to deliver a 4% internal rate of return [48]. The model is demonstrated here

using the UK Treasury's preferred social discount rate of 3.5% [49]. In practice, the constant could be set to any desirable hurdle rate dependent on who is making the investment.

### 5.4. Tariff considerations

There are three tariffs used in Eqn (17), namely, the initial generation, the export and the retail electricity tariff, respectively. The FIT tariffs have undergone significant reductions commensurate with the ~68% PV system cost reductions that have occurred since 2010 [50]. For systems up to 4 kWp, the generation tariff reduced from 43.3 p/kWh in 2010 to 21 p/kWh in March 2012, with another reduction to 16.0 p/kWh only 5 months later. At this juncture, a more responsive approach to degression—a systematised quarterly reduction in tariffs—was introduced, which allows accurate prediction of tariff reductions as long as deployment targets have been met. If deployment levels are low, then the tariff reductions may be skipped for up to two quarters. The current generation and export tariffs are 13.4 p/kWh and 4.85 p/kWh, respectively, for systems installed on or after 1 April 2015.

Average retail electricity prices have increased from 12.6 p/kWh in 2010 to 17.5 p in 2014 [51]. As described in the previous section, fuel prices are volatile and therefore likely to be highly uncertain going forward.



**Figure 11.** Bayesian network sub-model to calculate net present value (NPV) showing the deterministic nodes with their defined equations and the interface nodes, which connect to the rest of the model. See Section 5.1 for description of the equations and their parameters. OOBN, object-oriented Bayesian network; FIT, Feed-in Tariff.

Furthermore, unlike the FIT rates, electricity tariffs are subject to significant market uncertainties. For example, in terms of available tariffs, these might be lower-cost long-term contracts or high-cost card meter consumers [51]. Thus, the self-consumption contribution to the financial impact is subject to further uncertainty; and therefore, the OOBN model assumes a constant value whilst allowing this to be varied to explore a range of electricity cost scenarios.

### 5.5. Bayesian network and discounted cash flow analysis

A DAG represented by Figure 11 was constructed consistent with Eqn (17). Three interface nodes, marked with the broad input arrows, receive probabilistic inputs from the OOBN discussed in Section 4. Four deterministic nodes have been created, the equations for which are shown in the node frame in Figure 11. These show the

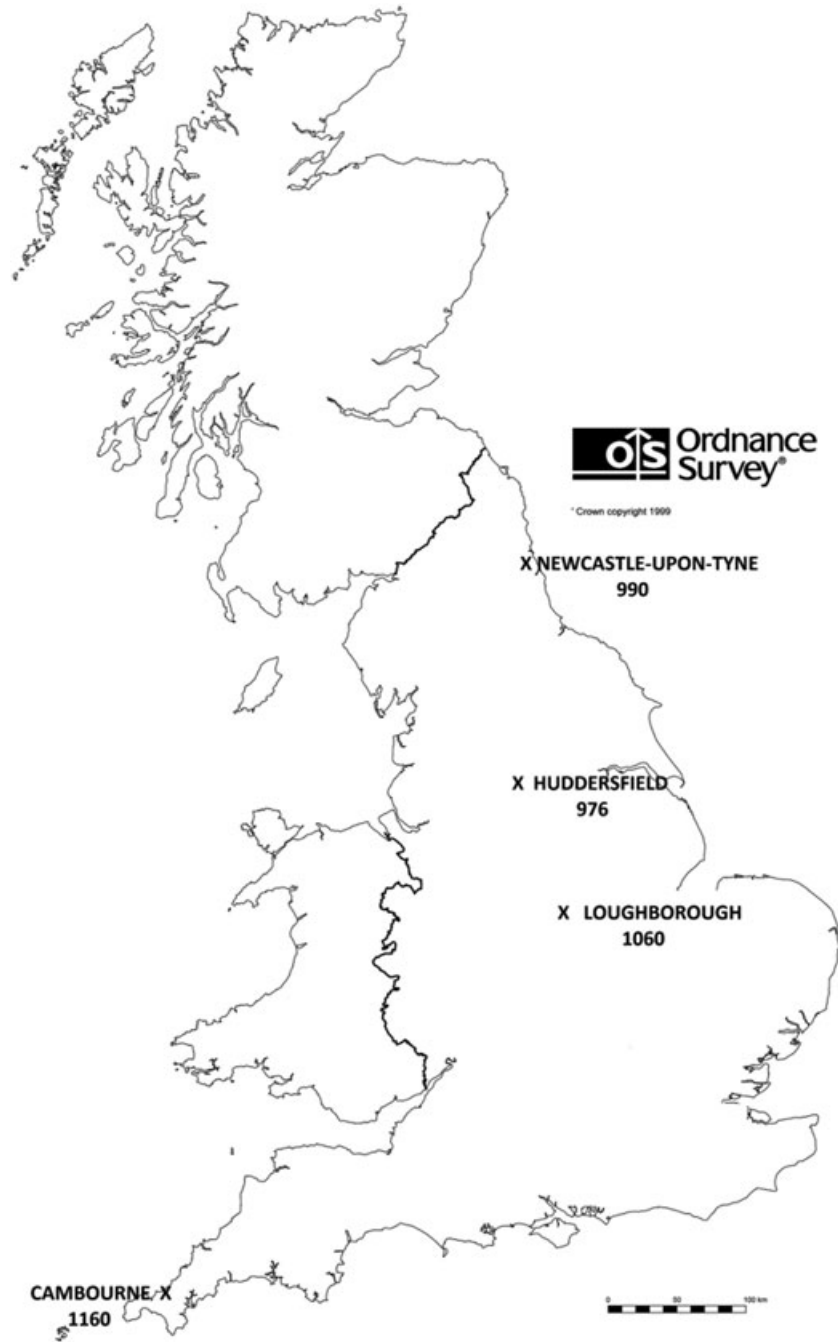
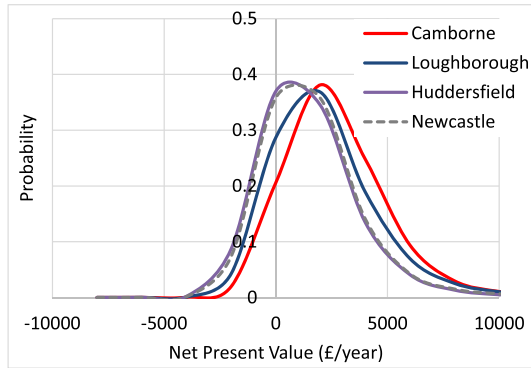


Figure 12. Location of the four geographic areas and horizontal irradiation used in this study.



**Figure 13.** Net present value distributions for each LSOA with no other evidence applied to any nodes and constants as in Table II.

**Table II.** Values of constants for Eqn (17) used to give the net present value distributions as shown in Figure 13.

Name	Unit
System cost	£1900/kWp
Retail Price Index	3%
Energy inflation rate	4%
Discount rate	4.5%
Initial FIT generation tariff	£0.125/kWh
Initial FIT export tariff	£0.05/kWh
Initial electricity tariff	£0.18/kWh
Life time of technology	20 years

FIT, Feed-in Tariff.

separate contributions of avoided import costs, FIT income and system costs to the overall NPV value. The input variables for these are received as inputs from other nodes, or are given as a constant. The CPTs for the deterministic nodes are calculated using 1000 simulations per discrete state. Once calculated, the model responds to evidence entered in any node elsewhere in the OOBN to recalculate a new posterior for the NPV components and the total. Furthermore, any of the parameters may be easily adjusted to evaluate various scenarios.

The building stock model of the OOBN allows the selection of a specific local geographic area for which the building stock data have been encoded. The four LSOAs selected for this case study are shown in Figure 12, together with global horizontal irradiation values in kWh/m<sup>2</sup>/year obtained from the PV GIS climate SAF model [50].

Figure 13 shows NPV distributions calculated by the model for each of the four LSOA census areas used in this study. For this scenario, the constants used in Eqn (17), documented in Table I, are fixed to the values given in Table II. Here, a generation tariff close the current generation and export tariffs of £0.0125 and £0.05, respectively, were chosen, and a discount rate of 4.5% selected as a representative hurdle rate. A moderately high electricity inflation rate was chosen (4%) and an RPI of 3%.

The variability of NPV reflects the uncertainty inherent in the specific input parameters, including system rating and yield, orientation and PV energy self-consumption. These in turn are influenced by their predictor variables

**Table III.** Average, standard deviation, coefficient of variation (CV) and percentile points for the resultant net present value (£ over 20 years) for prior distributions for each geographic area and posterior distributions following the application of evidence to input variables.

	Average (£)	Standard deviation (£)	CV %	Percentile (£)				
				10	25	50	75	90
All	2948	2509	85	-1753	-800	688	2157	3906
<b>Area</b>								
Camborne	3770	2652	70	-1226	124	1436	3148	4941
Loughborough	3222	2564	80	-1606	-560	922	2546	4354
Huddersfield	2482	2318	93	-1942	-1131	238	1691	3436
Newcastle	2621	2373	91	-1863	-1025	372	1773	3553
<b>Orientation</b>								
East	1879	2186	116	-2709	-1548	-395	1138	2580
South	3205	2534	79	-1585	-524	929	2487	4114
West	1824	2159	118	-2757	-1578	-436	1080	2439
<b>Rating (kWp)</b>								
2-2.25	2232	1953	87	-1890	-1162	57	1419	2785
3-3.25	3119	2387	77	-1622	-534	910	2395	3933
4-4.25	3883	2775	71	-1357	85	1566	3400	5364
<b>Degradation (%)</b>								
0-0.1	3483	2673	77	-1517	-305	1162	2912	4823
0.4-0.5	3149	2564	81	-1666	-625	874	2470	4186
0.9-1	2750	2442	89	-1827	-949	496	1907	3719
<b>Self-consumption (kWh/year)</b>								
0-200	-252	1571	-624	-3775	-3287	-2475	-1343	-394
600-800	2394	1490	62	-1558	-744	438	1410	1992
1400-1600	6146	1744	28	2171	2910	4135	5294	5989

as encapsulated in the various BN sub-models described in Section 4 and shown in Figure 2.

A key utility of a BN is the facility to enter hard or soft evidence (observations) for one or more parameters in order to constrain the model, and to support detailed analysis, such as that for a more localised assessment. The uncertainty of the remaining parameters is propagated to, and reflected in, the target variables. Thus, a variety of scenarios may be rapidly evaluated in terms of impact upon the discounted cash flow output. Table III shows variations in statistical indicators following the application of hard evidence on the orientation, system rating, degradation rate and self-consumption parameters, respectively.

Net present values are seen to be significantly sensitive to all these parameters whilst maintaining a marked variability across each location. Thus, the influence of orientation shows a difference in NPV of £1353 between systems facing East or West compared with the optimal azimuth (due South). Notably, the system degradation rate has a significant influence, varying from £3483 with a relatively conservative 0.1% annual degradation rate to £2750 for a degradation rate of 1% annually. Self-consumption is very influential on the NPV. A very low self-consumption of less than 200 kWh/year yields an average value of NPV of minus £252, whereas a high value of 1400 to 1600 kWh/year yields plus £6146.

Under many conditions, many systems do not yield a positive NPV at the chosen 4.5% discount rate as shown by the percentile points; often, less than 50% of systems do not make a positive return.

As well as the absolute value of NPV on the application of hard evidence, it is also of interest to observe the variability due to specific BN parameters. To this end, a sensitivity analysis was carried out, which measures the reduction in variance on a target node following application of hard evidence to an input node. The results are shown in Table IV, using NPV as the BN target node and identifies parameters, which result in a significant variance reduction. The greatest sensitivity (and cause of variability in the NPV output) is due to self-consumption,

which causes a 65% variance reduction in the NPV node on the application of hard evidence.

In summary, the results of this work illustrate the power of an object-oriented BN approach in co-evaluating the impact of a number of interdisciplinary parameters on PV technology performance indicators. In this case, a candidate techno-economic indicator (NPV) provides a valuable illustration of how multi-criteria parameters can be integrated for decision support and how the approach can account for diverse stakeholder perspectives, including those of system designers, investors and policy makers alike. Moreover, a probabilistic assessment of parameters of interest can transparently define risks pertaining to the attainment of specific KPIs in a wide number of simulated scenarios using this BN approach. The prospects for, and identification and robustness of, appropriate PV adoption pathways can thus be enhanced by deliberative policy and decision-making using empirical evidence where uncertainties have been previously obscured.

## 6. CONCLUSIONS

A BN model has been applied to a discounted cash flow analysis for domestic PV in order to manage inherent multi-domain uncertainties and deliver a rigorous and robust financial evaluation in the form of a probabilistic analysis of investment returns. The results indicate the sensitivity of returns to a number of key parameters, including PV degradation rate and the level of PV self-consumption together with PV yield, which is in turn related to geographical location and building geometry, as well as PV system technical factors.

A significant contribution to the research field is the integration of highly variable self-consumption, predicted by typical system yields and domestic electricity consumption [36], in order to quantify the variability of NPV. This probabilistic approach enables an enhanced understanding of risk for investors and policy makers. Thus, model further permits the exploration of FIT tariff scenarios to test the impact of subsidy changes on investment returns from a risk perspective; hereto policy makers have relied on deterministic models, which fail to account for the requisite variety inherent within the input parameters. This is demonstrated by the many scenarios in Table III where there is a high probability of not attaining the set discount rate. In this way, it is, for example, possible to model the required system prices under a variety of tariff and self-consumption scenarios to test the risk of attaining desired hurdle rates.

The approach enhances the analysis of complex stochastic parameters such as long term tariff fluctuations and human factors, such as electricity consumption patterns in order to yield detailed insights of relevance to a wide range of private and public sector stakeholders. The implications for policy makers grappling with multiple uncertainties whilst planning transitions to post-subsidy market and legislative frameworks are significant. The work also contributes significantly to the methodological

**Table IV.** A variance reduction sensitivity analysis for the net present value node.

Node	Variance reduction ( $\times 10^6$ )	Percent
Net present value	20.7	100
Self-consumption	13.5	65
System yield	9.09	43.8
Rating	5.92	28.5
System cost	5.73	27.6
Electricity consumption	5.58	26.9
Roof area	4.91	23.7
System cost per kWp	3.03	14.6
Specific yield	2.58	12.4
Yield uncertainty	1.32	6.36
Annual degradation rate	0.197	0.95

and intellectual development of the understanding and measurement of the techno-economic impact of PV deployment. The model can be readily adapted to test the impact on the techno-economics of PV under alternative post-subsidy financial models, such as full net metering or changes to direct self-consumption by virtue of demand-shifting or the integration of energy storage.

## ACKNOWLEDGEMENTS

The authors wish to acknowledge the financial support of the Engineering and Physical Sciences Research Council (EPSRC) of the UK through EPSRC grants EP/K022229/1 (WISE PV—Whole System Impacts and Socio-economics of wide scale PV integration) and EP/K02227X/1 (PV2025—Potential Costs and Benefits of Photovoltaics for UK-Infrastructure and Society).

## REFERENCES

1. DECC. UK Solar PV Strategy Part 1: Roadmap to a Brighter Future, 2013.
2. DECC. 'Solar Photovoltaics deployment': Feed-in Tariff Statistics, June 2015. Available from: <https://www.gov.uk/government/statistics/solar-photovoltaics-deployment> [accessed on 3 July 2015].
3. Leicester PA, Goodier C, Rowley P. Using a Bayesian network to evaluate the social, economic and environmental impacts of community deployed renewable energy. In *Proceedings of CISBAT, Clean Technology for Smart Cities and Buildings*, Lausanne, Scartezzini JL (ed.). 4–6 September 2013; 10 pp. Available from: <https://dspace.lboro.ac.uk/2134/14472> [accessed on 3 July 2015].
4. Leicester P. The Development of Object Oriented Bayesian Networks to Evaluate of the Social, Economic and Environmental Impacts of Solar PV. PhD Thesis, Loughborough University, 2015.
5. Goodier CI, Austin SA, Soetanto R, Dainty A. Causal mapping and scenario building with multiple organisations. *Futures* 2010; **42**(3): 219–229.
6. Checkland P. Soft systems methodology: a thirty year retrospective. *Systems Research and Behavioural Science* 2000; **17**: S11–S58.
7. Mingers J, Rosenhead J. *Rational Analysis for a Problematic World Revisited*. Wiley: Chichester, 2001.
8. Saltelli A. *Global Sensitivity Analysis*, 1st edn. John Wiley: Chichester, 2008.
9. Koller D, Friedman N. *Probabilistic Graphical Models*. MIT Press: Cambridge, MA, 2009.
10. Jensen F, Nielsen T. *Bayesian Networks and Decision Graphs*. Springer: New York, 2007.
11. Neapolitan R. *Learning Bayesian Networks*. Pearson Prentice Hall: Upper Saddle River, NJ, 2004.
12. Pearl J. Reasoning under uncertainty. *Annual Review of Computer Science* 1990; **4**(1): 37–72.
13. Koller D, Pfeffer A. Object-oriented Bayesian networks. In *Thirteenth Annual Conference on Uncertainty in Artificial Intelligence*, Providence, Rhode Island, 1997; 302–313.
14. Mrad A, Delcroix V, Piechowiak S, Leicester P, Abid M. An explication of uncertain evidence in Bayesian networks: likelihood evidence and probabilistic evidence. *Applied Intelligence* 2015; **43**(4): 802–824.
15. Delcroix V, Sedki K, Lepoutre F. A Bayesian network for recurrent multi-criteria and multi-attribute decision problems: choosing a manual wheelchair. *Expert Systems with Applications* 2013; **40**(7): 2541–2551.
16. Weidl G, Madsen A, Israelson S. Applications of object-oriented Bayesian networks for condition monitoring, root cause analysis and decision support on operation of complex continuous processes. *Computers & Chemical Engineering* 2005; **29**(9): 1996–2009.
17. Norsys. *Netica Application, (Version 5.12) Computer Software*. Norsys Software Corp: Vancouver, 2014. Available from: <http://www.norsys.com> [accessed on 3 July 2015].
18. Kelly S, Crawford-Brown D, Pollitt M. Building performance evaluation and certification in the UK: is SAP fit for purpose? *Renewable and Sustainable Energy Reviews* 2012; **16**(9): 6861–6878.
19. Kelly S. Do homes that are more energy efficient consume less energy?: A structural equation model of the English residential sector. *Energy* 2011; **36**(9): 5610–5620.
20. Wieder S. *An Introduction to Solar Energy for Scientists and Engineers*. Wiley: New York, 1982.
21. The Geoinformation Group. UK Buildings. Available from: [www.geoinformationgroup.co.uk/](http://www.geoinformationgroup.co.uk/) [accessed on 3 July 2015].
22. ONS. Super Output Area (SOA). Office for National Statistics, 2015. Available from: <http://ons.gov.uk/> [accessed on 3 July 2015].
23. OS MasterMap Topography Layer [GML geospatial data], Coverage: Various, Updated, Ordnance Survey, GB. Using: EDINA Digimap Ordnance Survey Service. Available from: <http://edina.ac.uk/digimap> [Downloaded: June 2014].
24. Quantum GIS. 2007. Available from: <http://qgis.org/> [accessed on 3 July 2015].
25. AddressBase. Ordnance Survey. Available from: <https://www.ordnancesurvey.co.uk> [accessed on 3 July 2015].
26. Google Earth. <https://earth.google.com> [accessed on 3 July 2015].

27. Nguyen H, Pearce J, Harrap R, Barber G. The application of LiDAR to assessment of rooftop solar photovoltaic deployment potential in a municipal district unit. *Sensors* 2012; **12**(12): 4534–4558.
28. European Commission Joint Research Centre. Photovoltaic Geographical Information System (PVGIS) Version 4, [online computer program], 2013. Available from: <http://re.jrc.ec.europa.eu/pvgis/> [accessed on 3 July 2015].
29. Colantuono G, Everard A, Hall L, Buckley A. Monitoring nationwide ensembles of PV generators: limitations and uncertainties. The case of the UK. *Solar Energy* 2014; **108**: 252–263.
30. Fenton N, Neil M. *Risk Assessment and Decision Analysis With Bayesian Networks*. Taylor & Francis: Boca Raton, 2012.
31. Department of Energy and Climate Change. National Energy Efficiency Data-Framework, [computer file]. UK Data Archive [distributor]: Colchester, Essex, July 2014. SN: 7518: 10.5255/UKDA-SN-7518-1.
32. Spiegelhalter D, Dawid A, Lauritzen S, Cowell R. [Bayesian analysis in expert systems]: rejoinder. *Statistical Science* 1993; **8**(3): 277–283.
33. Cao S, Sirén K. Impact of simulation time-resolution on the matching of PV production and household electric demand. *Applied Energy* 2014; **128**: 192–208.
34. Richardson I, Thomson M, Infield D, Clifford C. Domestic electricity use: a high-resolution energy demand model. *Energy and Buildings* 2010; **42**(10): 1878–1887.
35. DTI. Domestic Photovoltaic Field Trials—Final Technical Report. Department of trade and Industry, London, 2006.
36. Leicester P, Rowley P, Goodier I. Probabilistic Analysis of Solar PV Self-consumption Using Bayesian Network Models. IET Renewable Power Generation, Accepted for publication 30 November 2015. Available from: <https://dSPACE.lboro.ac.uk/dSPACE-jspui/handle/2134/19923> [accessed on 3 July 2015].
37. Cherrington R, Goodship V, Longfield A, Kirwan K. The feed-in tariff in the UK: a case study focus on domestic photovoltaic systems. *Renewable Energy* 2013; **50**: 421–426.
38. Cucchiella F, D’Adamo I, Gastaldi M, Koh S. Renewable energy options for buildings: performance evaluations of integrated photovoltaic systems. *Energy and Buildings* 2012; **55**: 208–217.
39. Jordan D, Kurtz S. Photovoltaic degradation rates—an analytical review. *Progress in Photovoltaics: Research and Applications* 2013; **21**(1): 12–29.
40. Riley K, Hobson M, Bence S. *Mathematical Methods for Physics and Engineering*. Cambridge University Press: Cambridge, 2006.
41. Darling S, You F, Veselka T, Velosa A. Assumptions and the levelized cost of energy for photovoltaics. *Energy and Environmental Science* 2011; **4**(9): 3133.
42. Dunlop ED. Lifetime Performance of Crystalline Silicon PV Modules. In *Proceedings of the 3rd World Conference on Photovoltaic Energy Conversion*, Osaka, Japan, 2003; 2927–2930.
43. Kroposki B, Hansen R. Technical Evaluation of Four Amorphous Silicon Systems at NREL. In *Proceedings of the 26th PV Specialists Conference*, Anaheim, CA, 1997; 1357–1360. DOI: 10.1109/PVSC.1997.654342.
44. Kaplanis S, Kaplani E. Energy performance and degradation over 20 years performance of BP c-Si PV modules. *Simulation Modelling Practice and Theory* 2011; **19**(4): 1201–1211.
45. Solar PV Cost Update. Prepared by Parsons Brinkerhoff for the Department of Energy and Climate Change. Parsons Brinkerhoff: London, May 2012.
46. ONS. Retail Price Index data, 1986 to 2014, Office for National Statistics, 2015. Available from: <http://ons.gov.uk/> [accessed on 3 July 2015].
47. Helm D. Energy policy: security of supply, sustainability and competition. *Energy Policy* 2002; **30**(3): 173–184.
48. Department of Energy and Climate Change. Impact Assessment: Periodic Review of FITs 2015. UK Government, DECC, London, 2015.
49. HM Treasury. The Green Book: Appraisal and Evaluation in Central Government. UK Government, HM Treasury, London, 2010.
50. OFGEM Generation & Export Payment Rate Table for Photovoltaic Installations—FIT Year 6 (2015/16)
51. DECC. Domestic Energy Price Statistics, March, 2015. Available from: <https://www.gov.uk/government/collections/domestic-energy-prices> [accessed on 3 July 2015].