Comparing district heating options under uncertainty using stochastic ordering☆

Victoria Volodina [a],[*], Edward Wheatcroft [b], Henry Wynn [b]

[a] *Clinical Operational Research Unit, University College of London, 4 Taviton Street, London, WC1H 0BT, UK*
[b] *London School of Economics, Houghton Street, London, WC2A 2AZ, UK*

ARTICLE INFO

ABSTRACT

District heating is expected to play an important role in the decarbonisation of the energy sector in the coming years since low carbon sources such as waste heat and biomass are increasingly being used to generate heat. The design of district heating often has competing objectives: the need for inexpensive energy and meeting low carbon targets. In addition, the planning of district heating schemes is subject to multiple sources of uncertainty, such as variability in heat demand and energy prices. This paper proposes a decision support tool to analyse and compare system designs for district heating under uncertainty using stochastic ordering (dominance) so that decision-makers can make robust decisions. The uncertainty in input parameters of the energy system model together with general scenarios are introduced to generate distributions of net present costs and emissions for each design. To perform inference about the induced distributions of outputs, we apply the orderings in the mean and dispersion. The proposed approach is demonstrated in an application to the waste heat recovery problem in district heating in Brunswick, Germany. The results obtained show that heat pump, a low carbon design option, is more robust in comparison to combined heat and power (CHP) and a mix of CHP and heat pump under all three scenarios, highlighting that robustness is an attractive feature of low-temperature waste heat recovery.

1. Introduction

The transition to a net-zero economy is an urgent challenge, and many countries have agreed to put in place national action plans to become carbon-neutral by 2050 or sooner [1,2]. A significant aspect of the response is an increase in funding for heat decarbonisation and energy efficiency projects. In the context of tight government budgets and the uncertainty associated with the implementation of green technologies, a comprehensive handling of risk is crucial in assessing the viability of these projects. The aim of this paper is to demonstrate a new approach to dealing with uncertainty for the planning of energy infrastructure projects in the form of stochastic ordering (dominance).

The context of this paper is waste heat recovery, in which heat from industrial and urban sources is used as a zero-carbon alternative to fossil fuels in district heating networks. In particular,

we consider a district heating project based on a real system in Brunswick (Braunschweig), Germany, in which heat from a data centre will be used as an input to a district heating network for a newly constructed residential and commercial area in the city. The system is one of four demonstrators on the ReUseHeat project [3]. The main objective of ReUseHeat is to aid *replication* of projects and provide advice to other investors. Here, "based on" refers to the matching of variables, that is matching the model inputs and outputs to the metrics observed in the real-world. We note that the modelling results were not used in the design of the actual system. However, the model conclusions derived can provide assistance for the design of similar projects.

A wide range of methods have been adopted to support decisions in the energy systems domain under uncertainty. For example, [4] pointed out the importance of uncertainty analysis in energy systems. Different approaches include the use of membership functions (fuzzy methods) and probability density functions (stochastic methods) for describing the uncertainty of input parameters. The effect of input parameters on model outputs can be quantified by simple statistical measures such as the standard deviation, mean squared error, confidence intervals and their multivariate and Bayesian counterparts. In addition, there is a recent interest in the use of scenarios to capture wider uncertainty issues [5].

* Corresponding author.
 E-mail addresses: v.volodina@ucl.ac.uk (V. Volodina),
E.D.Wheatcroft@lse.ac.uk (E. Wheatcroft), H.Wynn@lse.ac.uk (H. Wynn).

Nomenclature

Abbreviations

CHP	Combined heat and power
COP	Coefficient of performance
FSD	First-order stochastic dominance
KS	Kolmogorov–Smirnov
MV	Mean-Variance
NPC	Net Present Cost
OSeMOSYS	Open Source Energy Modelling System
SD	Stochastic dominance

Mathematical variables and notations

$\mathrm{vol}(\cdot)$	volume
α	significance level
$\boldsymbol{X}, \boldsymbol{Y}$	vectors of random variables
$\triangle(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1})$	the k-dimensional simplex whose vertices are $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1}$
$\forall$	for all
$\mathbb{E}[\cdot]$	expectation
$\mathbb{R}^d$	a coordinate space of dimension d over the real numbers.
$\preceq_\triangle$	dispersion ordering
$\preceq_{st}$	first-order stochastic ordering
sup	supremum, the least upper bound of a set
$F_X(z), F_Y(z)$	cumulative density functions (cdfs) for X and Y
$F_{\boldsymbol{X}}(\boldsymbol{z}), F_{\boldsymbol{Y}}(\boldsymbol{z})$	cumulative density functions (cdfs) for $\boldsymbol{X}$ and $\boldsymbol{Y}$
$g(\cdot)$	increasing (non-decreasing) function
X, Y	univariate random variables

Another general approach for decision support in energy systems is optimisation frameworks (stochastic optimisation), which operate by controlling the output variability of a component or system, while keeping the output minimum (maximum) on target [6]. In particular, in mean–variance (MV) and capital asset pricing models, widely used in finance and economics to assess different investment options, the aim is to achieve maximum yield and minimum volatility. Mean–variance portfolio theory has been employed for the modelling of uncertainties and risk in energy system planning [7,8]. These decision approaches rely on the MV rule, which states that, when two alternative prospects are considered, a rational decision-maker will choose an option with higher expected return and lower variance. However, the MV rule leads to a number of paradoxes such as the fact that it is sometimes unable to rank the two choices when a clear preference between them exists [9].

We note that the approaches described above operate with statistical summaries, which can fail to provide specific information about the distribution of outputs. We are keen to preserve the notion of robustness (uncertainty) by considering full probability distributions using stochastic ordering (dominance). Stochastic ordering is a special methodological framework, which underlines decision making under uncertainty and has been exploited in a number of fields, including robust design [10], portfolio theory [11] and signal processing [12]. In particular, we note that stochastic dominance has been used previously to compare

the performance of algorithms in power and energy system optimisation [13]. An extensive review of stochastic ordering is given in [14].

A stochastic ordering is a partial ordering on distributions. In this paper we consider first-order stochastic dominance (FSD) [9]. Orderings provide a useful framework for comparing distributions in the following sense: suppose we are aiming to maximise the return on our investment and have two choices to make regarding the investment portfolio (Portfolio X and Portfolio Y). We assume that there is uncertainty associated with returns on investment portfolios and we can construct probability distributions of the return in each case. If the distribution of choice X first-order stochastically dominates that of choice Y, for any given level of return, the probability that the profit exceeds that particular value is higher for choice X than it is for choice Y. In other words, if we are only interested in maximising the probability of the return exceeding some chosen value, we should always choose Portfolio X. This seems like a sensible and intuitive way to look at decision making under uncertainty and this is our approach here.

Stochastic orderings roughly divide into two classes: orderings that denote shift (in mean) and those that denote variability (dispersion). We note that the ordering does not prescribe uncertainty measures, but we claim that it provides a platform for uncertainty analysis in the following sense. Given an ordering, a suitable metric is one which is order-preserving with respect to that ordering. This means that it is a function whose expectation is ordered in the same direction as the stochastic underlying ordering. Thus both the mean and the median themselves are order-preserving with respect to first-order stochastic dominance and both the standard deviation and the Gini coefficients are ordered with respect to some well-known dispersion orderings. Following this condition, stochastic ordering allows us to extend the range of metrics for risk and uncertainty beyond the standard measures.

This paper is structured as follows. Section 2 covers background regarding district heating. Section 3 provides an introduction to stochastic orderings. Section 4 describes a series of simple computer experiments on a specially selected set of system designs, choices of input variations and broader scenario-based alternatives. Section 5 draws conclusions based on a selection of stochastic orderings. Section 6 contains the concluding remarks.

2. Problem specification

2.1. District heating

District heating is a system in which heat is produced by some centralised source and distributed via a network of insulated pipes. It is particularly well-developed in northern Europe and Scandinavia [15]. Historically, district heating has been powered by the burning of fossil fuels such as coal and gas. However, more recently, it has been seen as an opportunity to decarbonise the heating sector via the use of waste heat from industry and other sources, with an increasing focus on the opportunities of recovering waste heat from low temperature, urban sources such as, for example, metro stations [3]. The prevalence of low temperature sources and their location near areas of heat demand provides an opportunity in the wider agenda of carbon reduction [16].

A major difference between high and low temperature heat recovery is that, in the latter case, the heat typically needs to be upgraded before it is suitable for use in the network. This requires the use of a heat pump to increase the temperature to the required level. The installation of heat pumps poses additional technical challenges due to a lack of maturity in the technology and a lack of experience in installation and maintenance. In

Table 1

Description of design options in the study.

Design type	Description
Design Option 1	Combined Heat and Power (CHP) is employed to meet both the baseload and seasonal heat demand.
Design Option 2	A heat pump is employed to meet baseload heat demand and CHP is used to meet seasonal heat demand.
Design Option 3	A heat pump is employed with a small amount of storage to meet both the baseload and seasonal heat demand.

addition, heat pumps run on electricity and this creates additional operational costs and a vulnerability to increases in the price of electricity. This is considered to be a major risk in low temperature heat recovery [3].

At present, low temperature heat recovery is not widespread. There are a number of reasons for this, including a lack of political and commercial awareness, a lack of interest from heat 'owners', the immaturity of the technology and a lack of a legal and regulatory framework [3]. However, one of the biggest barriers, in our opinion, to the widespread rollout of low temperature waste heat recovery is a gap between the risk assessment required by financial institutions and that which is typically provided by project developers. With the methodology presented in this paper, we aim to contribute towards closing this gap.

2.2. Heat recovery in Brunswick

We construct a simple model based on the Brunswick (Braunschweig) in Germany within the ReUseHeat project [3]. The aim of the demonstrator is to supply heat to the newly developed residential area with 400 housing units, using waste heat from a nearby data centre. The diagram in Fig. 1 shows that excess heat from a data centre needs to be upgraded to a suitable temperature with a heat pump. The new housing units will also be connected to the existing city-wide network, which, at present, supplies 45 percent of residents in the city using a gas-powered combined heat and power (CHP) unit. The intention is then that heat from the data centre will cover the baseline demand (demand that is present throughout the year such as for hot water) and heat from the CHP will cover seasonal demand.

For the study described in this paper, we plan to assess three different design options to meet the demand for heat, which are summarised in Table 1. Each design option is evaluated in terms of its Net Present Costs (mln €) and CO_2-equivalent emissions (in metric tonnes).

To construct a model and run simulations, we employ an open source optimisation model for energy planning called Open Source Energy Modelling System (OSeMOSYS) [17]. OSeMOSYS is a deterministic, linear optimisation model that obtains the energy supply mix that minimises the Net Present Cost (NPC), subject to a number of chosen constraints. We have selected OSeMOSYS for two main reasons. Firstly, it provides us with the flexibility to operate on a local (city) level and to define our own scenarios. Secondly, OSeMOSYS is an open source model and therefore freely available for comparative project modelling.

3. Uncertainty and stochastic orderings

We consider the use of stochastic ordering for describing and comparing uncertainty. The rationale is that stochastic orderings are weaker than a limited list of specific metrics. In particular, stochastic ordering is a partial ordering, whereas the commonly used metrics of uncertainty, such as standard deviation, impose total ordering.

Suppose X and Y are two real-valued random variables representing the values of NPC (or emission levels) associated with two energy system designs, which we refer to as design options 1 and 2. We construct a probability distribution for each output under each design option $F_X(z)$ and $F_Y(z)$ respectively.

We can say that design option 1 has first-order stochastic dominance (SD) over design option 2, written $Y \preceq_{st} X$, if:

$$F_X(z) \leq F_Y(z) \quad \forall z \in \mathbb{R}, \tag{1}$$

$$F_X(z) < F_Y(z) \quad \text{for at least one value of } z \in \mathbb{R}.$$

Marshall and Arnold [19] presented the following conditions equivalent to $X \preceq_{st} Y$:

(i) $\mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)]$ for all increasing (non-decreasing) functions $g(\cdot)$;

(ii) $g(X) \preceq_{st} g(Y)$ for all increasing (non-decreasing) functions $g(\cdot)$;

The class of functions $g(\cdot)$ that satisfies (i.) and (ii.) is said to be the set of order-preserving functions with respect to a stochastic ordering. From Eq. (1), we conclude that, to look for first-order SD, we can simply plot the empirical cdfs. If one always lies below the other, we can claim that it dominates with respect to any increasing function.

3.1. Dispersion ordering

We consider dispersion orderings (sometimes known as dispersive orderings), which are a particular case of stochastic orderings. Dispersion measures the extent to which a distribution is stretched or squeezed, and therefore dispersion orderings compare the spread of the probability distributions [20,21]. In practice, decision-makers are interested to consider multiple metrics at the same time. Dispersion orderings can be extended for the multivariate variable case based on the idea of independent copies.

Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two d-dimensional random vectors, i.e. $\boldsymbol{X} = (X_1, \ldots, X_d)$ and $\boldsymbol{Y} = (Y_1, \ldots, Y_d)$. For instance, X_1 and Y_1 are two random variables representing the values of NPC for design option 1 and design option 2 respectively, whereas X_2 and Y_2 are random values that correspond to emission levels for these design options. Let us construct multivariate cdfs $F_{\boldsymbol{X}}(\boldsymbol{z})$ and $F_{\boldsymbol{Y}}(\boldsymbol{z})$ and define a function, $\phi(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_{k+1})$, that describes the separation between them. In [22], it is shown that, if we take

$$\phi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1}) = (\text{vol}\{\triangle(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1})\})^2,$$

where $\triangle(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1})$ is the k-dimensional simplex (in d dimensions) whose vertices are $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1} \in \mathbb{R}^d$, then $\mathbb{E}[\phi(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{k+1})]$ is a function of the variance–covariance matrix Σ of the underlying distribution. This prompts a dispersion ordering $\boldsymbol{X} \preceq_\triangle \boldsymbol{Y}$ defined by

$$\text{vol}\{\triangle(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{k+1})\} \preceq_{st} \text{vol}\{\triangle(\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{k+1})\}, \tag{2}$$

where $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{k+1}$ and $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{k+1}$ are independent draws from the $F_{\boldsymbol{X}}(\boldsymbol{z})$ and $F_{\boldsymbol{Y}}(\boldsymbol{z})$, respectively. If Eq. (2) holds, we can say that $\boldsymbol{X}$ is less dispersive than $\boldsymbol{Y}$. In other words, less uncertainty is associated with the distribution of the random vector $\boldsymbol{X}$ than the random vector $\boldsymbol{Y}$.

The case in which $k = d$ was introduced by Oja [23] and discussed in [24] and is referred to as the Simplex ordering. Here, we remove the requirement that $k = d$ and refer to the approach as the Generalised Simplex ordering. Note, again, that whenever we see the ordering $\preceq_{st}$, we can write down the class of order-preserving functions, $g(\cdot)$, in the present case
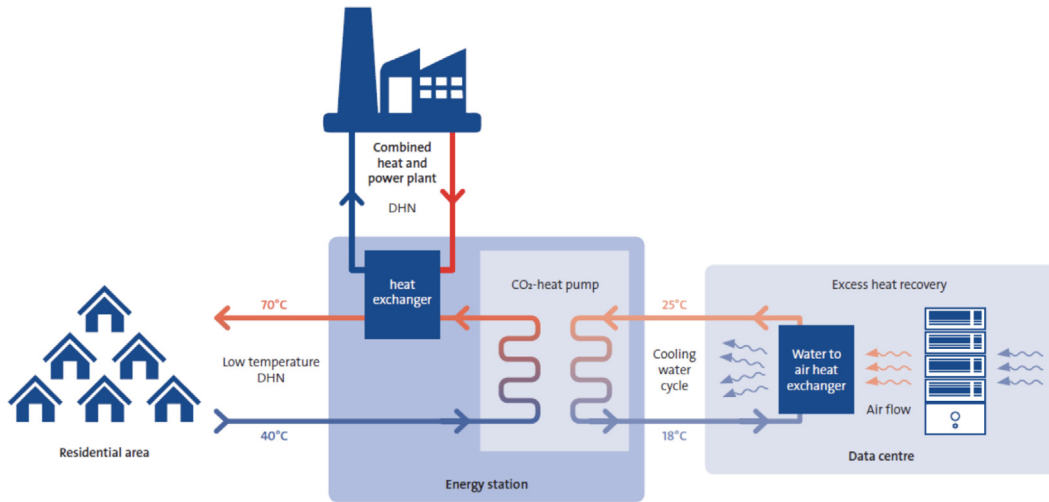
Fig. 1. An outline of the Brunswick demonstrator.
Source: BS|ENERGY. Adopted from [18].

of $\text{vol}\{\triangle(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{k+1})\}$. Dispersion orderings based on Hausdorff distance [25] and Mahalanobis distance [26] are possible alternatives to the Simplex ordering.

We would also like to add small remarks about the computation of simplex volume. Firstly, the simplex volume is a scale-free, homogeneous measure and therefore no pre-processing is required for the Generalised Simplex ordering. We also note that computing the simplex volume can become computationally expensive, in particular when we are operating with large data sets. To deal with this, [21] proposed the use of bootstrap resamples from the data set to obtain the distances. Despite this, we believe the properties of the Generalised Simplex ordering make it an attractive choice and we provide a demonstration of its use in Section 5,

3.2. Quantifying the difference between cdfs

It is useful to quantify the extent of the difference between two cdfs when one dominates another. Here, we use a Kolmogorov–Smirnov (KS) distance measure. For two random variables X and Y with cdfs $F_X(z)$ and $F_Y(z)$ respectively, the KS distance between X and Y is given by

$$D = \sup_z |F_X(z) - F_Y(z)|, \quad z \in \mathbb{R}.$$

Noting that the value of D lies in the range [0, 1], the statistic tells us the maximum difference between the two cdfs. If D is small, the extent to which one dominates the other is small and vice versa.

3.3. Hypothesis testing for stochastic orderings

In the studies presented in Section 5, the distributions of the outputs are produced using an energy systems simulation (OSeMOSYS [17]) and therefore we can only approximate the underlying cdf of the output variables that would be produced with an infinite sample size. Thus we make use of the Kolmogorov–Smirnov (KS) test proposed by [27] which defines the null hypothesis to be the case that Y first-order stochastically dominates X and the alternative hypothesis that the null hypothesis is false. If the null hypothesis can be rejected, we do not have enough evidence that one dominates the other.

Formally, let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be samples from F_X and F_Y, respectively. We proceed to define the hypotheses for testing first-order stochastic dominance as

$$H_0 : X \preceq_{st} Y \quad \text{against} \quad H_1 : X \npreceq_{st} Y, \tag{3}$$

where $\npreceq_{st}$ stands for the negation of $\preceq_{st}$. Following [27,28], the hypothesis testing in Eq. (3) is equivalent to

$$H_0 : F_Y(z) \leq F_X(z) \text{ for all } z \in \mathbb{R},$$
$$H_1 : F_X(z) < F_Y(z) \text{ for some } z \in \mathbb{R}.$$

Let us consider the two d-dimensional random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$, we propose the following test for dispersion ordering

$$H_0 : \boldsymbol{X} \preceq_{\triangle} \boldsymbol{Y} \quad \text{against} \quad H_1 : \boldsymbol{X} \npreceq_{\triangle} \boldsymbol{Y},$$

where $\npreceq_{\triangle}$ stands for the negation of $\preceq_{\triangle}$. This hypothesis testing is equivalent to

$$H_0 : \text{vol}\{\triangle(\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{d+1})\} \preceq_{st} \text{vol}\{\triangle(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{d+1})\},$$
$$H_1 : \text{vol}\{\triangle(\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{d+1})\} \npreceq_{st} \text{vol}\{\triangle(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{d+1})\}.$$

To test stochastic ordering, we proceed to adopt a one-tailed Kolmogorov–Smirnov (KS) test.

This approach is similar to that used by [21] who tested the Hausdorff and Simplex dispersion orderings.

We note that an assumption of the KS test is that the samples should be independent of each other and, although, in each sample, the simulations are performed independently, the input parameters induced by the scenarios are common to both and therefore the assumption fails. Similar technical difficulties arise with the bootstrap methodology. Nonetheless, we note that the effect of this failed assumption is small and diminishes with sample size (which is generally large). We therefore believe that these results are informative for our analysis and provide a useful diagnostic technique in general.

Another issue requiring us to be cautious is the use of multiple testing. In our experiments, we consider three options ($N = 3$) and, as a result, we perform three pairwise post-hoc tests to assess the differences between them. To account for this multiple testing, we make use of the Bonferroni correction. Therefore, if we specify a significance level of $\alpha = 0.05$, the corresponding significance level for each individual hypothesis is adjusted to $\alpha/N = 0.05/3 = 0.0167$.

4. Experimental design

We demonstrate the use of stochastic ordering to assess the aforementioned waste heat recovery project in Brunswick (see Section 2.2 for more details). Three different design options are considered, each of which differ according to the technology
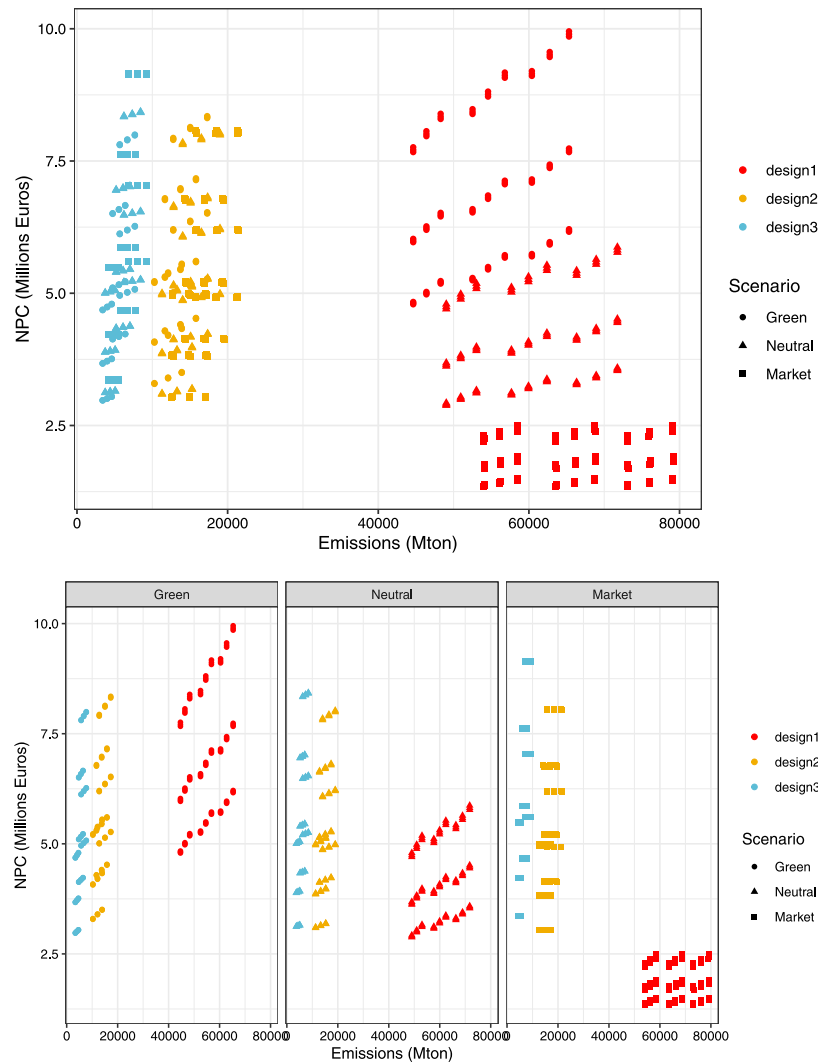
Fig. 2. *First row*: NPC (mln €) against CO_2-equivalent emissions (Mton) for all three design options under the three scenarios. *Second row*: NPC against emissions for all three design options, plotted separately for each scenario.

mix employed for supplying domestic heat. These are outlined in Table 1. Each design option is evaluated in terms of its Net Present Cost (NPC) (mln €) and CO_2-equivalent emissions (in metric tonnes). Here, we analyse the uncertainty in the outputs of the model induced by variations in four different inputs:

1. Operational costs.
2. Discount rate.
3. Coefficient of Performance (COP) for the heat pump (units of heat delivered per unit of electricity).
4. Emission Activity Factor (the emissions produced (in metric tonnes) from operating a particular technology in the energy system).

Variations in the four input variables are expected to impact both the NPC and emissions levels. For each input variable, we specify three levels: low, medium and high. We then perform simulations with a full factorial design (often known as a fully crossed design) so that all possible combinations across the model inputs are considered [29]. This gives a total of 81 simulation runs.

We consider each of the design options under three different scenarios. Scenario Analysis is a widely used uncertainty tool in energy systems studies [5]. Examples of scenarios related to energy include the Future Energy Scenarios (FES) published by the UK National Grid [30] and World Energy Scenarios [31], both of which consider different pathways to decarbonisation.

We define three scenarios that differ in terms of selected elements of government climate policy and consumer engagement with green technology. A description of the three scenarios is given in Table 2. In the Green scenario, a combination of high consumer engagement and government incentive schemes are used to rapidly reach net zero. In the Market scenario, there is a reliance on market forces and only limited government intervention in the energy system planning. Finally, the Neutral scenario captures the middle ground between the two. Since the operational lifetime of a typical heat pump is around twenty years, this time horizon is used for the study.

Under each scenario, we produce model simulations that, for a given set of model inputs, generate the volume of emissions and NPC over a 20 year period as model outputs. In Section 5.1, we consider the empirical cdfs of the model outputs under different scenarios and design options. Then, in Section 5.2, we consider the Generalised Simplex dispersion ordering to compare the effects of different scenarios and design options on uncertainty.

5. Results

Fig. 2 shows scatter plots of the two model outputs obtained for each of the three design options and under each scenario.
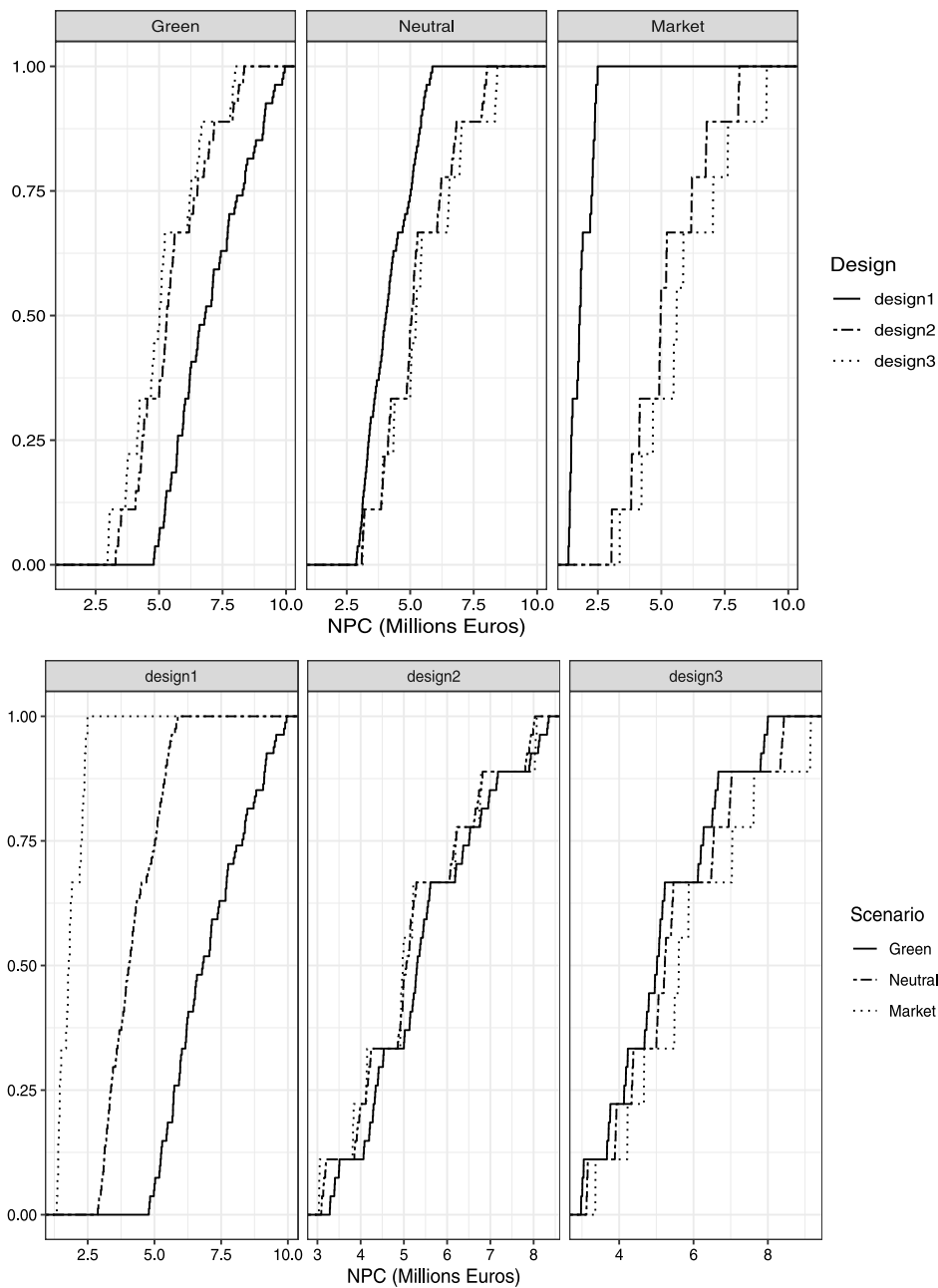
Fig. 3. Empirical cdfs for NPC for (i) all three design options plotted together for each individual scenario (*first row*), (ii) all three scenarios plotted together for each individual design option (*second row*).

Focusing first on emissions, since the emissions factor for natural gas is higher than for electricity, the highest level of emissions is produced by design option 1, followed by design option 2 and then design option 3. From the second row in Fig. 2, we observe that the different scenarios have only a small impact on the levels of emissions under each of the design options, since government intervention strategies considered in this paper mainly aim at changing behaviour through cost.

Differences in the three scenarios have a much larger impact on NPC. Under the Market and Neutral scenarios, design option 1 corresponds to the lowest NPC values, which can be explained by the high investment cost of the heat pump. However, under the Green scenario, the opposite is true due to a high carbon penalty. Therefore, under the Green scenario, design options 2

and 3 become more attractive alternatives, thus demonstrating the value of considering different scenarios when informing decision-makers.

5.1. Orderings in the mean

Fig. 3 shows the empirical cdfs of the NPC of each design option under each of the three scenarios (first row) and of each scenario under the three design options (second row). In Table 3 values of the Kolmogorov–Smirnov (KS) distance measure between the cdfs of different design options under each scenario are shown. The numbers in the table represent the distance between the cdfs when the design option denoted in the column dominates that denoted in the row and indicate the extent to which one design option dominates the other. In each cell, there are three numbers that correspond to the Market, Neutral and Green
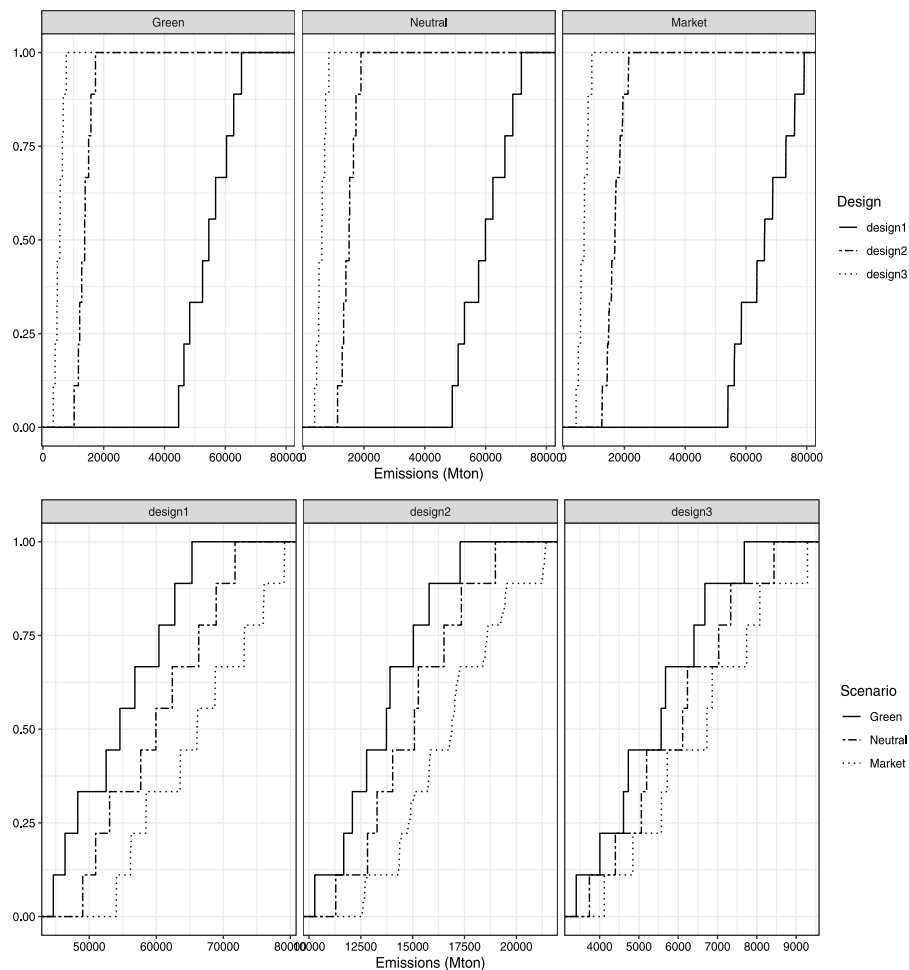
Fig. 4. Empirical cdfs for CO_2-equivalent emissions for (i) all three design options plotted together for each individual scenario (*first row*), (ii) all three scenarios plotted together for each individual design option (*second row*).

Table 2
Description of the selected scenarios in the study.

Scenario	Description
Green scenario	Penalty per Mton of emissions: 100 € per Mton. Annual change in baseload and seasonal heat demand: −1%. Increasing gas prices and decreasing electricity prices.
Neutral scenario	Penalty per Mton of emissions: 40 € per Mton. Baseload and seasonal heat demand fluctuate around the central projections. Gas and electricity prices stay within the central projected values.
Market scenario	No penalty per Mton of emissions. Annual change in baseload and seasonal heat demand: 1%. Decreasing gas prices and increasing electricity prices.

Table 3
Kolmogorov–Smirnov distances between empirical cdfs of NPC and CO_2-equivalent emissions for pairs of design options. The KS values correspond to the Market, Neutral and Green scenarios and are only shown when the cdf for the design option in the row is dominated by that in the column. NA values indicate that the cdfs cross, and first-order SD does not exist in either direction.

		Design 1	Design 2	Design 3
NPC	Design 1		1/0.37/–	1/0.41/–
	Design 2	–/–/0.48		0.33/NA/–
	Design 3	–/–/0.57	–/NA/0.23	
Emissions	Design 1		–/–/–	–/–/–
	Design 2	1/1/1		–/–/–
	Design 3	1/1/1	1/1/1	

scenarios, respectively. Cases in which there is no dominance in this direction are denoted with a dash. In addition, we test the first-order stochastic ordering and provide the corresponding *p*-values in Table 4. Cases where the cdfs cross and first-order SD clearly does not exist are denoted by an 'na'.

From the first row of Fig. 3, we observe that, under the Market scenario, design option 3 has first-order SD over design option 2 which has first-order SD over design option 1. Under the Green

scenario, the ordering is reversed, mainly due to the increased financial support from policy makers for renewable energy sources. Under the Neutral scenario, the cdfs for design option 2 and design option 3 cross, so first-order SD cannot discriminate between these two options. However, we observe that design option 1 is dominated by the other two options. We therefore conclude that, if the sole aim is to minimise NPC, under the Green scenario, design option 3 would be the preferred option whilst design option 1 would be the preferred option under the Neutral and Market scenarios. The reason for this difference is that, under the Green scenario, the government introduces high penalties for generation of emissions, which makes design option 1 an unattractive option for providing heat.
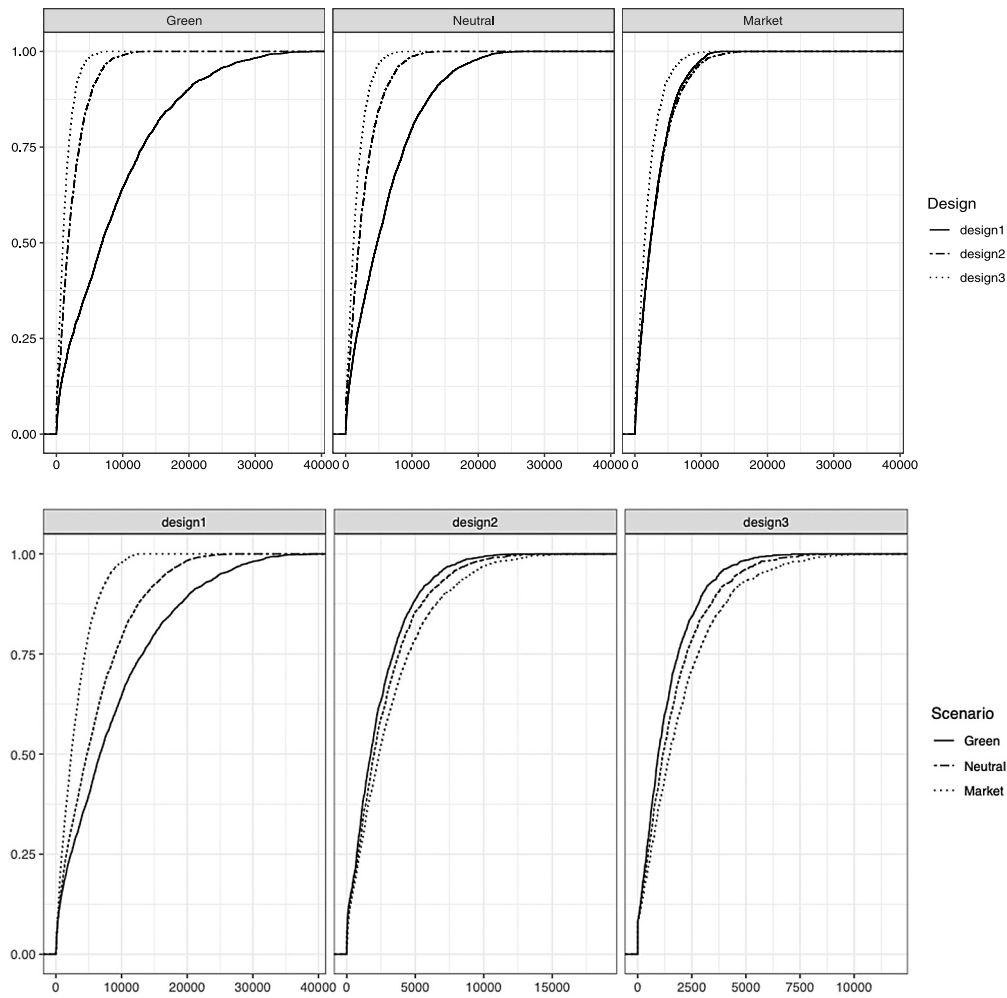
Fig. 5. *First row*: empirical cdfs of the Generalised Simplex metric for all three design options plotted together for each individual scenario. *Second row*: empirical cdfs of the Generalised Simplex metric for all three scenarios plotted together for each individual design option.

Table 4
Observed *p*-values for NPC and CO_2-equivalent emissions using first-order stochastic ordering. The three *p*-values in each table entry correspond to the Market, Neutral and Green scenarios and an asterisk denotes that stochastic dominance is rejected at the chosen significance level. NA values indicate that the cdfs cross, and therefore no hypothesis test is performed.

		Design 1	Design 2	Design 3
NPC	Design 1		1/1/0*	1/1/0*
	Design 2	0*/< 0.001*/1		1/NA/0.012*
	Design 3	0*/< 0.001*/< 0.001*	< 0.001*/NA/1	
Emissions	Design 1		0*/0*/0*	0*/0*/0*
	Design 2	1/1/1		0*/0*/0*
	Design 3	1/1/1	1/1/1	

Table 5
Kolmogorov–Smirnov distances between empirical cdfs of NPC and CO_2-equivalent emissions for pairs of scenarios. The KS values correspond to design option 1, design option 2 and design option 3 and are only shown when the cdf for scenario in the column lies below that in the row. NA values indicate that cdfs cross, and first-order SD does not exist.

		Market	Neutral	Green
NPC	Market		1/NA/–	1/NA/–
	Neutral	–/NA/0.33		0.78/0.20/–
	Green	–/NA/0.33	–/–/0.18	
Emissions	Market		–/–/–	–/–/–
	Neutral	0.33/0.33/0.22		–/–/–
	Green	0.56/0.56/0.44	0.33/0.33/0.22	

Table 3 shows that in the Market and Neutral scenarios design option 2 dominates design option 1 with a KS distance of 1 and 0.37, respectively. We conclude that there is greater difference between design option 2 and design option 1 under the Market scenario than under the Neutral scenario. In the Green scenario, the dominance is reversed and design option 1 dominates design option 2 with a KS distance of 0.48. Focusing on Table 4, for NPC, the *p*-values related to the tests with the null hypothesis that design option 1 first-degree stochastically dominates design option 2 under Market, Neutral and Green scenarios are 0/ < 0.001/1. We therefore reject the null hypothesis for the Market and Neutral scenarios whilst there is not enough evidence to reject the null hypothesis for the Green scenario. These results are in line with the findings observed in Fig. 3.

Table 6

Observed p-values for NPC and CO_2-equivalent emissions using first-order stochastic ordering. The p-values correspond to design option 1, design option 2 and design option 3, where p-values with asterisk are less than or equal to significance level. NA values indicate that cdfs cross, and first-order SD does not exist and we cannot perform hypothesis tests.

		Market	Neutral	Green
NPC	Market		1/NA/< 0.001*	1/NA/< 0.001*
	Neutral	0*/NA/1		1/1/0.06
	Green	0*/NA/1	0*/0.042/1	
Emissions	Market		< 0.001*/< 0.001*/0.018	0*/0*/0*
	Neutral	1/1/1		< 0.001*/< 0.001*/0.018
	Green	1/1/1	1/1/1	

Focusing now on the effects of each scenario on the three design options, under design option 1 the Green scenario dominates the Neutral scenario which dominates the Market scenario. In this case, the ordering indicates that the highest NPC values are generated under the Green scenario. For design option 3, on the other hand, the ordering between scenarios is reversed, and we conclude that the highest NPC values correspond to the Market scenario. Under design option 2, the cdf for the Neutral scenario dominates that of the Green scenario whilst the cdf for the Market scenario crosses the other two cdfs and thus there is no ordering. Overall, the difference in orderings may lead a planner to think carefully about their choice of design option under different possible futures. If, for example, they consider policy featured in the Green scenario to be likely, they may choose design option 2 or 3 to mitigate that risk.

In Table 5, we present KS distances for the NPC and emissions associated with each scenario under the selected design options to quantify how different these empirical cdfs are. Notably, for NPC, under design option 1, we observe high values of the KS distance between pairs of scenarios. This demonstrates a substantial scenario effect which is much larger than for the other two design options. Table 6 provides p-values for the one-sided KS test. The high p-values indicate that there is not enough evidence to reject the stochastic ordering in design option 1 and design option 3 observed in Fig. 3.

Fig. 4 shows the empirical cdfs corresponding to emission levels. The empirical cdfs of emissions for each design option under each of the three scenarios (top row) demonstrate that, under all three scenarios, design option 1 dominates design option 2, which dominates design option 3. In Table 3, we observe KS values equal to 1, which indicates that empirical cdfs of emissions under all three scenarios significantly differ from each other. These findings are supported by p-values provided in Table 4. Notably, the specified scenarios differ in terms of the extent to which policy is driven by 'green' considerations and confirms our expectation that waste heat recovery is a carbon reducing technology under most reasonable policy decisions.

5.2. Generalised Simplex dispersion ordering

In Section 5.1, we demonstrated how to assess the impacts of different design options by considering orderings in the mean. We now focus on dispersion, and note that a dispersive probability distribution indicates a high level of uncertainty (risk), which is assumed to be a negative attribute. Here, we focus primarily on the effects of scenarios on each of the design options.

Fig. 5 displays empirical cdfs for the Generalised Simplex ordering. Those of the three design options under each of the three different scenarios are shown in the top row, whilst those of the three scenarios under each of the three design options are shown in the second row. We observe that under the Green and Neutral scenarios there is a clear ordering: design option 3 is lesser than design option 2 which is lesser than design option 1 in the Generalised Simplex dispersion order. Under the Market

Table 7

Kolmogorov–Smirnov distances between Generalised Simplex empirical cdfs for pairs of design options. The KS values correspond to the Market, Neutral and Green scenarios and are only shown when the cdf for design option in the column lies below that in the row. NA values indicate that cdfs cross, and first-order SD does not exist.

	Design 1	Design 2	Design 3
Design 1		NA/–/–	–/–/–
Design 2	NA/0.33/0.49		–/–/–
Design 3	0.19/0.49/0.63	0.19/0.21/0.24	

Table 8

Observed p-values for different design options using the Generalised Simplex dispersion ordering. The p-values correspond to the Market, Neutral and Green scenarios, where those with an asterisk are less than or equal to the chosen significance level. NA values indicate that cdfs cross, and first-order SD does not exist and we cannot perform hypothesis tests.

	Design 1	Design 2	Design 3
Design 1		NA/0*/0*	0*/0*/0*
Design 2	NA/0.98/1		0*/0*/0*
Design 3	1/1/1	1/1/1	

scenario, design option 3 is dominated by the other two design options, though the cdfs of design options 1 and 2 cross. We can conclude that under the Generalised Simplex ordering, design option 3 is the most robust option under all three scenarios.

Table 7 shows the KS distances between the empirical cdfs associated with the pairs of design options under the Market, Neutral and Green scenarios, respectively, and the p-values in Table 8 correspond to the KS tests of the dispersion ordering. The KS distances between the design options in Table 7 are on average smaller under the Market Scenario than under the Green Scenario, suggesting that the difference in robustness is profound for the case in which government prioritise green issues. In all cases in which the dispersion ordering is demonstrated, the difference according to the KS test is significant.

For a planner, the above results are informative in that they demonstrate the relative robustness of different design options under each of the three scenarios, which is an important criterion for decision-makers.

6. Concluding remarks

The importance of informing decision-makers and stakeholders about the uncertainty associated with design choices is increasingly recognised in energy systems [4]. The present paper demonstrates the use of stochastic ordering (dominance) in the context of local energy planning under uncertainty. In particular, we have introduced variability at two levels: local variability in inputs and more general scenarios to account for changes in consumer behaviour and the political environment. Based on a generated data set, we have considered stochastic orderings both in shift and dispersion.

In our analysis of the waste heat recovery system in Brunswick, we have demonstrated that, under a Green scenario,

which assumes active government policy to meet the 2050 net-zero carbon target, lower CO_2-equivalent emissions can be produced at a lower cost by employing a heat pump with heat from a data centre. In addition, we have found using the Generalised Simplex ordering that this design choice is more robust (less volatile) compared to the other options under all three scenarios. We argue that, if it can be shown to be true in general, this robustness is an attractive feature of low-temperature waste heat recovery. We note that, in a number of cases, the cdfs cross and therefore first-order SD failed to discriminate between these respective choices. One possible next step is to consider higher-orders of stochastic dominance, which can be performed as an extension to the presented analysis.

We consider the proposed approach for analysing and comparing different options for providing residential heat in district heating as a competitive alternative to traditional uncertainty metrics used in energy systems. Our non-parametric method offers orderings in uncertainty based on distribution functions, which are easy to visualise and communicate to decision-makers and other stakeholders. In addition, very limited information is required from decision-makers and experts to construct SD, contrary to the traditional mean–variance analysis where the specification of a utility function is necessary [32,33]. An extension to the proposed approach is to introduce new uncertainty metrics by specifying order-preserving functions of the presented orderings.

CRediT authorship contribution statement

Victoria Volodina: Formal analysis, Software, Visualisation, Writing – original draft, Writing – review & editing. **Edward Wheatcroft:** Formal analysis, Writing – original draft, Writing – review & editing, Resources. **Henry Wynn:** Conceptualisation, Writing – original draft, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Dataset and code related to this article can be found at https://github.com/vicvolodina93/OrderHeating/tree/master, hosted at Github.

References

[1] European Commission, The European Green Deal, Eur. Comm. (ISSN: 1098-6596) 53 (9) (2019) 24, http://dx.doi.org/10.1017/CBO9781107415324.004, URL https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52019DC0640&from=EN.

[2] Jos Delbeke, Artur Runge-Metzger, Yvon Slingenberg, Jake Werksman, The Paris agreement, in: Towards a Climate-Neutral Europe: Curbing the Trend, 2019, pp. 24–45, http://dx.doi.org/10.4324/9789276082569-2.

[3] Kristina Lygnerud, Edward Wheatcroft, Henry Wynn, Contracts, business models and barriers to investing in low temperature district heating projects, Appl. Sci. (Switzerland) (ISSN: 20763417) 9 (15) (2019) http://dx.doi.org/10.3390/app9153142.

[4] Alireza Soroudi, Turaj Amraee, Decision making under uncertainty in energy systems: State of the art, Renew. Sustain. Energy Rev. (ISSN: 13640321) 28 (2013) 376–384, http://dx.doi.org/10.1016/j.rser.2013.08.039.

[5] Edward Wheatcroft, Henry Wynn, Chris J. Dent, Jim Q. Smith, Claire L. Copeland, Daniel Ralph, Stan Zachary, The scenario culture, 2019, pp. 1–25, URL http://arxiv.org/abs/1911.13170.

[6] Stefano Moret, Frédéric Babonneau, Michel Bierlaire, François Maréchal, Decision support for strategic energy planning: A robust optimization framework, Eur. J. Oper. Res. (ISSN: 03772217) 280 (2) (2020) 539–554, http://dx.doi.org/10.1016/j.ejor.2019.06.015.

[7] Anastasia Ioannou, Andrew Angus, Feargal Brennan, Risk-based methods for sustainable energy system planning: A review, Renew. Sustain. Energy Rev. (ISSN: 18790690) 74 (July 2016) (2017) 602–615, http://dx.doi.org/10.1016/j.rser.2017.02.082, URL http://dx.doi.org/10.1016/j.rser.2017.02.082.

[8] Anindya Bhattacharya, Satoshi Kojima, Power sector investment risk and renewable energy: A Japanese case study using portfolio risk optimization method, Energy Policy (ISSN: 03014215) 40 (1) (2012) 69–80, http://dx.doi.org/10.1016/j.enpol.2010.09.031.

[9] Haim Levy, Myles Robinson, Stochastic Dominance: Investment Decision Making Under Uncertainty, Vol. 34, Springer, New York, ISBN: 9783319217079, 2006, http://dx.doi.org/10.2307/3009513.

[10] Laurence W. Cook, Jerome P. Jarrett, Optimization using multiple dominance criteria for aerospace design under uncertainty, AIAA J. 56 (12) (2018) 4965–4976.

[11] Jan Annaert, Sofieke Van Osselaer, Bert Verstraete, Performance evaluation of portfolio insurance strategies using stochastic dominance criteria, J. Bank. Finance (ISSN: 03784266) 33 (2) (2009) 272–280, http://dx.doi.org/10.1016/j.jbankfin.2008.08.002, URL http://dx.doi.org/10.1016/j.jbankfin.2008.08.002.

[12] Cihan Tepedelenlioglu, Adithya Rajan, Yuan Zhang, Applications of stochastic ordering to wireless communications, IEEE Trans. Wirel. Commun. 10 (12) (2011) 4249–4257.

[13] Gianfranco Chicco, Andrea Mazza, Assessment of optimal distribution network reconfiguration results using stochastic dominance concepts, Sustain. Energy Grids Netw. (ISSN: 23524677) 9 (2017) 75–79, http://dx.doi.org/10.1016/j.segan.2016.12.005.

[14] K. Mosler, Marco Scarsini, Stochastic Orders and Applications: A Classified Bibliography, Springer-Verlag, 1993.

[15] Danish Energy Agency, Regulation and planning of district heating in Denmark, 2016, p. 27, URL http://www.ens.dk/sites/ens.dk/files/climate-co2/Global-Cooperation/Publications/Publications/regulation_and_planning_of_district_heating_in_denmark.pdf.

[16] Urban Persson, Helge Averfalk, Accessible urban waste heat: Deliverable 1.4, 2018.

[17] Mark Howells, Holger Rogner, Neil Strachan, Charles Heaps, Hillard Huntington, Socrates Kypreos, Alison Hughes, Semida Silveira, Joe DeCarolis, Morgan Bazillian, Alexander Roehrl, OSeMOSYS: The open source energy modeling system. An introduction to its ethos, structure and development, Energy Policy (ISSN: 03014215) 39 (10) (2011) 5850–5870, http://dx.doi.org/10.1016/j.enpol.2011.06.033, URL http://dx.doi.org/10.1016/j.enpol.2011.06.033.

[18] Brunswick, https://www.reuseheat.eu/brunswick/. (Accessed 11 January 2022).

[19] Albert W. Marshall, Barry C. Arnold, Springer Series in Statistics Inequalities : Theory of Majorization, Media, ISBN: 9780387400877, 1979.

[20] Moshe Shaked, Dispersive ordering of distributions, J. Appl. Probab. 19 (2) (1982) 310–320.

[21] Guillermo Ayala, María Concepción López-Díaz, Miguel López-Díaz, Lucía Martínez-Costa, Methods and Algorithms to Test the Hausdorff and Simplex Dispersion Orders with an R package, Methodol. Comput. Appl. Probab. (ISSN: 15737713) 17 (3) (2015) 661–675, http://dx.doi.org/10.1007/s11009-013-9386-z.

[22] Luc Pronzato, Henry P. Wynn, Anatoly A. Zhigljavsky, Extended generalised variances, with applications, Bernoulli (ISSN: 13507265) 23 (4A) (2017) 2617–2642, http://dx.doi.org/10.3150/16-BEJ821.

[23] H. Oja, Descriptive statistics for multivariate distributions, Statist. Probab. Lett. 1 (October) (1983) 327–332.

[24] Alessandra Giovagnoli, Henry P. Wynn, Multivariate dispersion orderings, in: Statistics & Probability Letters, Technical Report, 22, 1995, pp. 325–332.

[25] Miguel López-Díaz, An indexed multivariate dispersion ordering based on the Hausdorff distance, J. Multivariate Anal. (ISSN: 0047259X) 97 (7) (2006) 1623–1637, http://dx.doi.org/10.1016/j.jmva.2006.02.004.

[26] Luc Pronzato, Henry P. Wynn, Anatoly A. Zhigljavsky, Simplicial variances, potentials and Mahalanobis distances, J. Multivariate Anal. (ISSN: 10957243) 168 (2018) 276–289, http://dx.doi.org/10.1016/j.jmva.2018.08.002, URL DOI: 10.1016/j.jmva.2018.08.002.

[27] Daniel McFadden, Testing for stochastic dominance, in: Studies in the Economics of Uncertainty, Springer, 1989, pp. 113–134.

[28] Garry F. Barrett, Stephen G. Donald, Consistent tests for stochastic dominance, Econometrica 71 (1) (2003) 71–104.

[29] E.P. George, William Gordon Hunter, J. Stuart Hunter, Statistics for Experimenters: Design, Innovation, and Discovery, Wiley, 2005.

[30] National Grid, Future energy scenarios, Energy (July) (2014) 220, URL https://www.nationalgrid.com/NR/rdonlyres/C7B6B544-3E76-4773-AE79-9124DDBE5CBB/56766/UKFutureEnergyScenarios2012.pdf.

[31] World Energy Council (WEC), World Energy Scenarios 2019 : Exploring Innovation Pathways to 2040, WEC Publications, 2019, pp. 1–152, URL https://www.worldenergy.org/assets/downloads/2019_Scenarios_Full_Report.pdf.

[32] Stefano Canessa, John G. Ewen, Matt West, Michael A. McCarthy, Terry V. Walshe, Stochastic dominance to account for uncertainty and risk in conservation decisions, Conserv. Lett. 9 (4) (2016) 260–266.

[33] Harry M. Markowitz, G. Peter Todd, Mean-Variance Analysis in Portfolio Choice and Capital Markets, Vol. 66, John Wiley & Sons, 2000.