Li, Pei Hao; Zamanipour, Behzad; Keppo, Ilkka

Revealing technological entanglements in uncertain decarbonisation pathways using bayesian networks

Published in:
Energy Policy

DOI:
10.1016/j.enpol.2024.114273

Published: 01/10/2024

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Li, P. H., Zamanipour, B., & Keppo, I. (2024). Revealing technological entanglements in uncertain decarbonisation pathways using bayesian networks. *Energy Policy*, *193*, Article 114273.
https://doi.org/10.1016/j.enpol.2024.114273

Energy Policy 193 (2024) 114273

Contents lists available at ScienceDirect

Energy Policy

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

Revealing technological entanglements in uncertain decarbonisation pathways using bayesian networks

Pei-Hao Li [a], Behzad Zamanipour [b,*], Ilkka Keppo [a,b]

[a] UCL Energy Institute, University College London, Central House, 14 Upper Woburn Place, London, WC1H 0NN, UK
[b] Department of Mechanical Engineering, Aalto University, Otakaari 1 B, 02150, Espoo, Finland

ARTICLE INFO

Keywords:
Uncertainty
Decarbonisation pathways
Bayesian networks
Energy system model

ABSTRACT

To effectively meet the ambitious objectives set by the Paris Agreement, gaining a deeper understanding of the relationships between the key technologies involved in mitigation activities is pivotal. This research uses Bayesian Network (BN) methodology on a large ensemble of energy system model runs, aiming to shed light on the complex interdependencies, and related uncertainties, among the various technologies within the pathways. We specifically focus on tracking the evolution and interconnectedness of technology portfolios over time, enabling dynamic assessments of the impacts linked to specific deployment strategies. The results suggest that prioritizing early-stage transitions within the building sector is imperative and the consistent deployment of district heating emerges as a pivotal element in the long-term plans for decarbonisation. In the power sector, the rising trends in electrification and the substantial growth in low-carbon power plants and wind energy deployment, underscore the urgency for adaptable strategies within the power sector. Notably, the integration of bioenergy with carbon capture and storage (BECCS) also emerges as a crucial technology, offering a means to counterbalance emissions from carbon-intensive industries. The BN-based approach provides decision makers a powerful tool for comprehensive, informed, and systematic planning as they navigate towards a carbon-neutral future, but it is also crucial to acknowledge the reliance of our analysis on assumptions inherent in energy system models. Studies using different assumptions and model structures are needed to confirm the generalizability of our findings.

1. Introduction

To achieve the goals of the Paris Agreement, global energy systems should be deeply decarbonised in the coming decades to reach net-zero CO2 emissions by 2050 (Masson-Delmotte, 2018). Global energy systems thus need to be dramatically restructured with low-carbon technologies to reduce low-carbon emissions from all energy supply and use sectors. The development pathways for energy system development, however, have always been subject to various types of economic, technological, political, behavioural, and other uncertainties (Hughes et al., 2013), complicating the planning of actions of the different stakeholders from citizens to investors and policymakers.

The perhaps most common way of considering that uncertainty in long-term energy system analysis is through scenarios (Guivarch et al., 2022), often quantified using energy system optimization models, such as TIMES (Loulou et al., 2005) and OSeMOSYS (Howells et al., 2011), that suggest, under the specified assumptions, the most cost-effective pathways to transform the existing energy system and reach the set long-term decarbonisation targets. Using a set of scenarios with varying assumptions (Riahi et al., 2017) illustrates a range of different energy system outcomes and configurations, ideally demonstrating the strategies that are robust across the different scenarios. Such scenarios alone, however, tell us little about the specific relationships between the technologies that succeed – or fail – in the scenarios. For example, do certain technologies require other technologies to succeed and thus co-evolve with them? Do others compete and therefore crowd each other out? In other words, what kind of lock-in, or lock-out, relationships might exist in the scenarios (Unruh, 2000). And how do these relationships function across time, i.e. do we need to deploy certain technologies first, to deploy -or lock out - others later? Our study aims to address this gap and use Bayesian Networks to investigate how technologies and technology portfolios evolve and link in time.

* Corresponding author.
E-mail address: behzad.zamanipour@aalto.fi (B. Zamanipour).

https://doi.org/10.1016/j.enpol.2024.114273
Received 29 September 2023; Received in revised form 8 July 2024; Accepted 23 July 2024
Available online 29 July 2024

1.1. Literature review

Various families of methodologies have been proposed to evaluate the uncertainties in the energy system models. Sensitivity analysis was used by Moksnes et al. (2019) to capture the influences of uncertain developments of electricity infrastructure for South America in 324 scenarios. This approach was also adopted by Fais et al. (2016) to create 32 scenarios to explore the technology uncertainty of the decarbonisation of the UK energy system.

Narrative-informed approach is another method to analyse various uncertainties. For instance, Moallemi et al. (2017) developed scenario categories, considering the influences of government or market forces and societal needs. For each category, 2500 transition pathways were created with the Exploratory Modelling Workbench (Kwakkel, 2017) for India's electricity sector.

Pye et al. (2015), in turn, used Monte Carlo sampling to evaluate the impacts of uncertainties on the decarbonisation pathways for 80% reduction in GHG emissions by 2050. In total, 500 pathways were generated to better represent the statistical characteristics of uncertain inputs. The same approaches have also been applied to generate 1800 uncertain pathways for three major scenarios, considering the impacts of climate ambition and the availability of CCS (Pye et al., 2019).

Moreover, Li and Trutnevyte (2017) linked a whole energy system model (UK TIMES model) to a power sector pathway generator model (D-EXPANSE) to generate 800 pathways for the UK power sector using a combined approach of Monte Carlo sampling and Modelling-to-Generate-Alternatives. The combined approach can reflect both the parametric uncertainties (e.g. technology performance and costs) and structural uncertainties in energy system modelling. For a more comprehensive review of uncertainty assessment with energy system models, please refer to Yue et al. (2018).

A wide range of statistical approaches have been applied in previous studies to reveal consistent patterns across uncertain pathways (i.e. scenario discovery) for robust decision-making. In the past, descriptive statistics and visualisation have been widely applied to gain insights from a large ensemble of uncertain decarbonisation pathways. For instance, Li and Trutnevyte (2017) drew out the installed capacity of individual power technologies over time across 800 pathways to show the impacts of uncertainties on the required capacity of power plants for achieving low-carbon targets. Likewise, Price and Keppo (2017) generated a series of plots, such as line charts, bar charts, and error-bar charts, to visualise the variations (e.g. energy production and consumption) across 16 near cost optimal scenarios for robust policy-making to decarbonise the global energy system. Pizarro-Alonso et al. (2019) used cobweb plots to depict the impacts of the six most influential parameters on the electricity price and installed wind capacity for the Danish electricity system in 2050 across 100 simulations, considering uncertainties of 22 input parameters. Multivariate linear regression models have also been applied by Pye et al. (2015) to determine the impacts of uncertain inputs on two key output metrics, total system costs and GHG emissions, based on standardised regression coefficients.

The most well-known scenario discovery approaches are the Patient Rule Induction Method (PRIM) (Friedman and Fisher, 1999) algorithm and the Classification and Regression Tree (CART) (Gordon et al., 1984). These can reveal drivers for a few target metrics, such as GHG emissions and total energy system costs. PRIM algorithm searches for combinations of input variables that can best explain the group within which data points have similar output characteristics. The meaningful input variables are chosen through statistical data-mining searches, considering a trade-off between interpretability (i.e. density) and coverage of different combinations of determinants. Moksnes et al. (2019) applied the PRIM algorithm to identify the most influential input assumptions to the key outcomes, such as total investment costs and GHG emissions, for possible future developments of electricity infrastructure in South America, considering the uncertainties of electricity demand, fossil fuel price, learning curve, discount rate, CO2-emission cap, and hydropower. PRIM was also utilized by Li et al. (2023) to determine which important indicators and their combinations have the most impact on achieving climate objectives of below 2 °C or below 1.5 °C. CART, on the other hand, can be used to discover combinations of key drivers leading to specific scenario outcomes (Guivarch et al., 2016). For instance, Herran et al. (2019) used CART to identify the three most influential policy measures to distinguish the performance of decarbonisation policies in terms of carbon emissions and energy consumption per household. Nevertheless, PRIM and CART are only able to reveal the linear relationships between input variables and output variables of interest. Non-linear interactions between model parameters thus failed to be captured by these methods (Quinn et al., 2017).

Clustering algorithms have been frequently adopted to group pathways based on a limited number of key output metrics prior to scenario discovery analysis, such as PRIM. For instance, Moksnes et al. (2019) divided 324 scenarios into three or four clusters based on certain key characteristics of interest, such as total system costs or total GHG emissions, using the Gaussian Mixture Model (GMM).

Csereklyei et al. (2017) grouped energy mixes in 28 EU member states in the past 40 years into 7 types using a model-based clustering algorithm (i.e. GMM) to investigate the drivers for the transition from the dominant carbon-intensive fuels to low-carbon fuels.

Pye et al. (2019) applied a hierarchical clustering algorithm to identify interdependency among key technological metrics in 2050 based on the correlations between metrics across uncertain pathways. Li et al. (2020) further verified that the k-means algorithm is the most robust approach to characterising a large ensemble of uncertain decarbonisation pathways (600 in total), and applied the algorithm to identify five distinctive pathways with various decarbonisation strategies, such as focusing on low-carbon power, CCS, or green hydrogen production.

The development of technologies can, however, also be co-evolutionary or substitutional over time. The deployment of technology at the earlier stage might influence the flexibility or suitability of deployment of co-evolutionary or substitutional technologies. For example, it has been shown that the evolution of heating infrastructures can be seen as a series of path-dependent processes with rising returns to adoption as fuel sources, infrastructures, and end-use technologies coevolve to improve system performance (Gross and Hanna, 2019). It is thus essential to understand the temporal evolution of technology portfolios so that policymakers can know how to deploy and substitute technologies over time to achieve a specific kind of low-carbon future. In the previous studies, the temporal evolution of technologies was mostly based on a few scenarios and has not been thoroughly investigated using systematic approaches to date.

Bayesian networks (BNs) are regarded as a valuable tool to investigate uncertain complex systems due to the following four advantages: (1) probabilistic relationship between explanatory variables can be explicitly revealed; (2) prior expert knowledge can be easily incorporated into the model; (3) overfitting of data can be avoided; (4) capability of handling incomplete input datasets as dependencies between variables are embedded in the model (Bassamzadeh and Ghanem, 2017). Therefore, in recent years, BN models have been applied widely to energy-related issues, such as prediction of electricity demand with various temporal resolutions (e.g. 15 min and hourly intervals) (Bassamzadeh and Ghanem, 2017), selection of sustainable charging sites for electric vehicles (Hosseini and Sarder, 2019), estimation of long-term wind power output (Carta et al., 2011), and detection of faults of air conditioning systems (Hu et al., 2018).

Moreover, BNs have been successfully applied to understand consumers' energy-related behaviours, such as user behaviour of cooking-oven usage to predict hourly energy consumption in a household (Basu et al., 2013; Hawarah et al., 2010), potential reduction of electricity consumption in peak load periods (Morris et al., 2015), window opening/closing behaviour of occupants in residential buildings (Barthelmes et al., 2017). BNs have also been used to explore consumers' travel choices. For example, Xie and Waller (2010) applied BNs to investigate the deterministic variables for travel mode choice of work trips, including drive alone, shared ride, public transit, and bicycle and walk, in San Francisco, US. Similarly, Ma et al. (2017) investigated commuting mode choices, such as car and public transport, for the cross-border workers of Luxembourg using BNs, considering socio-demographic, spatial, and commuting trip variables.

Furthermore, BNs have been applied to investigate energy-related risk assessment applications. Cui et al. (2024) analysed the failure probability of hydrogen pipelines to reduce the risk of damage. Chang et al. (2019) conducted a risk analysis using dynamic BN to evaluate the potential uncertainty and risk associated with the hydrogen production unit leakage. In another recent study, a dynamic object-oriented BN was adopted to analyse the safety of hydrogen pipelines (Dao et al., 2024). Additionally, BN has been used for risk analysis for energy projects, such as oil refineries, nuclear power plants, and biodiesel plants (Machado, de Oliveira Ribeiro, and do Nascimento, 2023).

Lately, BNs have been applied to generate scenarios for energy and climate assessments. For example, Small et al. (2019) used expert judgments and surveys to construct BNs to generate scenarios for the adoption of CCS, considering the influences of technological and socioeconomic factors. Düspohl and Döll (2016) used two participatory modelling approaches to build BNs to evaluate the strategies, such as loans and feed-in-tariffs, for promoting renewable electricity generation (e.g. PV) in Germany.

However, to date, BNs have not been applied to explore technological relationships in uncertain decarbonisation pathways for robust decision-making. This study is hence the first using BNs to identify key technologies over strategies for long-term decarbonisation, based on model-generated uncertain pathways.

This study intends to use BNs to reveal the temporal evolution of technologies to further inform policymakers of how to transform the UK's energy systems over time to achieve different low-carbon futures with various technological landscapes. Hierarchical clustering algorithm is also used to compare its results with the ones obtained from the BN-based method. In addition, k-means clustering algorithm is applied to determine representative scenarios in terms of technology portfolios in 2050 and gain insights into the influential technologies in the energy transition.

The rest of the paper is organized as follows: Section 2 presents the methodology of the study and the analysed ensemble of scenarios. Section 3 provides the results and discusses various aspects of the temporal evolution of technologies. Finally, section 4 provides insights into the energy transition pathways for policymakers and suggests directions for future research works.

2. Methodology

An ensemble of decarbonisation pathways for the UK to reach 80% reduction in GHG emissions by 2050 was adopted in this study. Only technological metrics in 2030, 2040, and 2050 were taken into account for simplification. BNs were then built to reveal the probabilistic relationships between technologies in two periods, between 2030 and 2040 and 2040 and 2050, based on the dataset. A weighted importance analysis procedure identified the influential technologies in different periods based on the built BNs. The BNs were further applied to investigate technological transition over two periods for five representative low-carbon scenarios with various technological combinations in 2050, identified by clustering analysis. The methodologies applied are explained in the following sections.

2.1. Bayesian networks

A BN is a directed acyclic graphical (DAG) (i.e. without directed cycles) model, representing the joint probability distribution of a set of random variables by making conditional independent (CI) assumptions (Bassamzadeh and Ghanem, 2017; Murphy, 2012). In the DAG, nodes represent random variables and directed links denote conditional probability distributions between random variables (nodes), linking parent nodes (on which the distribution is conditioned) to child nodes. The links can be interpreted as probabilistic relationships. For instance, in a DAG representing $p(c|a,b)$, there will be two links from nodes a and b to c (Murphy, 2012).

In the most general term, a joint probability distribution over n variables can be presented using the product rule of probability as follows.

$$p(x_1, \ldots, x_n) = p(x_1)p(x_2|x_1)p(x_3|x_1, x_2)\ldots p(x_n|x_1, \ldots, x_{n-1}) \tag{1}$$

This implies that each random variable has relationship with all the other variables. For a large set of variables, the joint probability becomes too complicated to be evaluated efficiently. Therefore, assumptions about conditional independence between variables can be made to simplify the formulation. Topological ordering can then be constructed from any DAG, ordering parent nodes to come before child nodes. Ordered Markov Property further assumes that a node only depends on its immediate parents, not on all predecessors in the ordering. A DAG can then be presented in a factorised expression.

$$p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p(x_i|pa(x_i)) \tag{2}$$

where $p(x_i|pa(x_i))$ is a conditional probability distribution (CPD); $pa(x_i)$ is the collection of parent nodes to node x_i. Equation (2) only holds if the conditional independence (CI) assumptions encoded in the associated DAG are correct. With the CI assumptions, a DAG model can have much fewer parameters than those required by a model without CI assumptions (Bishop and Nasrabadi, 2006; Murphy, 2012).

2.2. Structure learning

Potentially, there is unmanageable number of possible structures of BNs. It is an almost impossible task to find the global optimal structure of BNs to represent the joint probability distribution over a large set of random variables. Learning BN structure is thus regarded as NP-hard (i.e. extremely difficult to solve) (Chickering et al., 2004). Heuristic algorithms are frequently applied to search for the best factorisation of the joint distribution based on a goodness-of-fit score, such as Bayesian Information Criterion (BIC) (Murphy, 2012). BIC, a popular log-likelihood function that can be used for model selection, is adopted to evaluate how the considered structures fit to the decarbonisation pathway dataset. BIC can be expressed as follows (Barthelmes et al., 2017).

$$BIC = \sum_{i=1}^{n} \log p\left(x_i \middle| \prod x_i\right) - \frac{d}{2}\log n \tag{3}$$

where d is the number of variables included in the model; n is the sample size. The first term is the log-likelihood the fitness to the target joint probability distribution based on the samples, given a network structure under consideration. The latter term penalises overfitting to avoid including too many irrelevant variables just to have a better fit to the samples.

Hill-climbing algorithm, a heuristic searching algorithm, is then applied to greedily search for a better network structure to factorise the joint probability distribution by altering the structures in each iteration. The algorithm starts with an empty or randomly generated network. A directed link is then added, deleted, or reversed as long as the action would not lead to a cyclic network and the BIC score of the new network increases. The restructuring of the network continues until the algorithm reaches a network with the maximum BIC score and cannot further improve the score. The final network is regarded as an ideal choice of network structure for the presentation of the joint probability distribution of the samples. It is worth noting that the global optimal structure is not guaranteed to be found with the greedy algorithm as the search procedure is likely to stop at a local optimum network structure (Murphy, 2012; Scutari, 2009).

To avoid the searching of the BN structure being trapped in a poor local optimum, the search procedure was restarted 400 times with random initial structures, on each of which 5 attempts were made to randomly insert/remove/reverse a link to learn BN structures in this study (Scutari, 2009).

Moreover, only links from technologies in the previous year to those in the next year are considered in candidate structures to represent the influences of deployment of technologies in the earlier period on that in the later period.

2.3. Parameter learning

With a known structure of BN, the parameters of the BN can be learned from samples using Bayesian posterior estimation (Murphy, 2012). In this study, all technological metrics are discretized and transformed into discrete variables with five levels, representing the levels of deployments, for simplification. Therefore, the parameters of nodes are conditional probability tables (CPTs) which tabulate $p(x_i|pa(x_i))$. Each row in the table corresponds to a configuration of the parent variables, which has a specific probability distribution. Each entry in the table can be denoted by $\theta_{ijk} = p(X_i = k|pa(X_i) = j)$, representing the probability that node i is in state k, given that its parent nodes are in state j. To simplify the parameter estimation, parameters associated with each node in the BN can be computed independently (i.e. global independence). Moreover, each row of the CPT table for each node can be determined independent of other rows (i.e. local independence). The probability distribution of each row can be represented by Dirichlet prior. Then we can compute the posterior distribution, which is also a Dirichlet distribution, by adding the pseudo counts to the empirical counts to estimate the parameters given the observed samples. Eventually, the mean of the distribution can be estimated as follows.

$$\overline{\theta_{ijk}} = \frac{N_{ijk} + \alpha_{ijk}}{\sum_k (N_{ijk} + \alpha_{ijk})} \tag{4}$$

where α_{ijk} is the prior Dirichlet hyperparameter, known as the pseudo counts. The pseudo count represents an analyst's assumed number of occurrences within an imaginary sample, depicting the case where node i, set at state k, and its parent nodes at state j exist within the prior distribution. With the pseudo counts, the zero-count problem caused by the maximum likelihood estimation can be avoided (Bassamzadeh and Ghanem, 2017; Murphy, 2012).

2.4. Weighted importance analysis

Mutual information between technology metrics was then evaluated to determine the importance of individual technology metrics in terms of the influences on the deployment of other technology metrics in the same year or future years. The influences can be either positive or negative. In information theory, mutual information is a measure that shows the reduction in uncertainty of a variable Q after observing the other variable F (Murphy, 2012). Unlike correlation coefficients, mutual information can reflect both linear and non-linear relationships between variables. It is thus a more general measure than correlation coefficients to represent relationships between technology metrics. The uncertainty of a variable Q can be represented by entropy, defined as follows.

$$H(Q) = -\sum_q P(q)\log_2 P(q) \tag{5}$$

where $P(q)$ is the probability of variable Q in the state of q. The mutual information (i.e. entropy reduction, MI) is calculated as the expected reduction in mutual information of Q, a target variable, from a finding for variable F, calculated as follows (Marcot, 2012).

$$MI = H(Q) - H(Q|F) = \sum_q \sum_f \frac{P(q,f)\log_2[P(q,f)]}{P(q)P(f)} \tag{6}$$

where $H(Q)$ is the entropy of Q before any new findings, $H(Q|F)$ is the entropy of Q after new findings from variable F, and Q is measured in information bits; q is the state of variable Q; f is the state of variable F (Marcot, 2012; Marcot et al., 2006).

The weighted importance of a technology metric i, in turn, is defined as the average of weighted mutual information between a technology metric and other technology metrics in year t.

$$I_{i,t} = \frac{\sum_j Var_j \times MI_{i,j}}{N} \tag{7}$$

where $I_{i,t}$ is the importance of a technology metric i in year t, which can be the same year as the year the metric is in or a future year; j is a technology metric in year t; N is the total number of technology metrics in year t; $MI_{i,j}$ is the reduction in uncertainty of technology metric j after technology metric i is observed; Var_j is the variance of technology metric j across pathways. As some technology metrics vary more widely across pathways than the other technology metrics, policymakers should pay more attention to those more uncertain metrics. Metrics' variances are thus incorporated in the determination of metric importance.

In this study, we used a R package bnlearn (Scutari, 2009) to construct BNs to reflect probabilistic relationships between technologies from decarbonisation pathway ensembles. The same package was also applied to determine parameters of the Bayesian networks. Another R package gRain (Højsgaard, 2012) was then adopted to evaluate the importance of individual technologies in a single year and across years. These two packages have been widely adopted in the past studies using Bayesian networks (Barthelmes et al., 2017; Bassamzadeh and Ghanem, 2017; Song, Semakula, and Fullana-i-Palmer, 2018).

2.5. Ensemble of uncertain decarbonisation pathways

An ensemble of 600 decarbonisation pathways, achieving 80% reduction in GHG emissions relative to 1990 levels by 2050, were taken from a previous study (Pye et al., 2019) for demonstrating our approach. The Energy System Modelling Environment model (ESME) (Heaton, 2014), a technology-rich whole energy systems model for the UK, was used to generate these pathways using Monte Carlo sampling of uncertain techno-economic parameters (Pye et al., 2015). ESME determines the optimal portfolios of technologies in all sectors of the UK energy system to provide sufficient services to meet the demands, with minimum cumulative discounted total energy system costs until 2050, using linear programming. All pathways satisfy the targets of 53% reduction in GHG emissions by 2030 and 80% by 2050. In the Monte Carlo sampling, not only the uncertainties of technology costs and resource costs were taken into account, but also the uncertainties related to technology build rates and resource availability, such as the maximum potential of biomass produced sustainably in the UK or via import from the international market. Low variances of parameters were set for mature technologies, while high variances of parameters were applied to emerging technologies, as uncertainties for those technologies are greater. Probability functions were constructed to represent the uncertainty of the considered parameters using triangular distributions (Pye et al., 2019). For each simulation, values were sampled for 2050 and determined for the intermediate years (prior to 2050) based on interpolation back to the base year (2010) value, following a linear trajectory between 2010 and 2050. Distributions were mostly assumed to be independent, but correlations were considered for some that can clearly be assumed to move together (e.g. a light-duty electric vehicle and an electric car). For more details regarding pathway generation under techno-economic uncertainties, please refer to Pye et al. (2019), including its appendix A1 for the parametrization.

Consequently, some of the relationships seen in the results are likely to be emphasised by the correlations that have been assumed. This is also the case for the deployment of a given technology across time, as our sampling approach means that there is an inbuilt technology specific path-dependency. As the parameter values for 2030 and 2040 depend directly on the values of 2050, results for a given technology across time are more likely to correlate than they would without this assumption. This sampling assumption has less impact on cross-technology comparison connections, however. Nonetheless, in reality, technology costs are likely to be more interlinked than what our correlations capture due to shared components (e.g. (Moglianesi et al., 2023)), similar engineering tasks, and general technological developments.

3. Results and discussion

3.1. Bayesian networks

The constructed BNs are shown in Appendix C, fi gures C.1 and C.2. As the networks are complex and include much probabilistic information, we have extracted data from the BNs to highlight the strong relationships between the metrics and construct a topology for the two networks. These "Technology transition networks" for the two periods (i.e. 2030-2040 and 2040–2050) are illustrated in Figs. 1 and 2, respectively, showing also the variance of, and correlations between, the metrics. The definition of technological metrics can be found in Appendix A. The variance and average values of technology metrics in 2030 and 2050 and the distribution of metric levels for each technology in the two BNs can be found in Appendix B and Appendix C respectively. In both figures, a small group of transport technologies is separated from the rest of the network as no strong correlations between two technology groups can be found without decreasing the likelihood of the network to the pathway ensemble. A Technology Transition Network considering all technology metrics in all three years, 2030, 2040, and 2050, is not presented here since only a very limited number of probabilistic relationships exist in the BN, as the relationships wane dramatically across years.

As shown in Figs. 1 and 2, technology metrics in 2050 seem to have stronger connections with those in 2040 than the connections between metrics in 2040 and 2030, as there are more thick links between 2040 and 2050 than between 2030 and 2040. In other words, the deployment

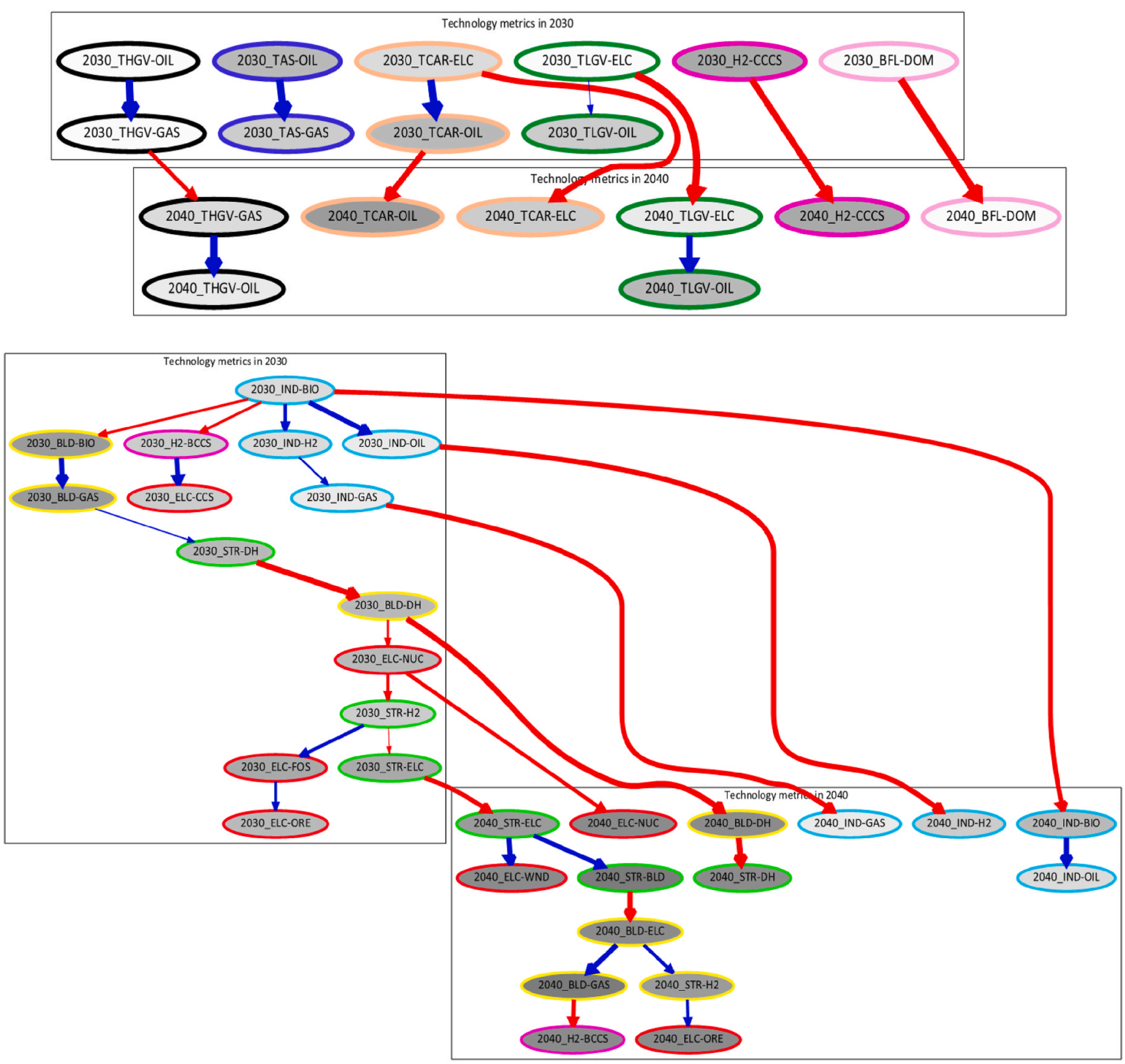


Fig. 1. Technological transition network for the transition between 2030 and 2040. Red arrows indicate a positive correlation between technologies; blue arrows indicate a negative correlation between technologies. Thicker links mean the correlations are stronger; the greyscale of nodes represents the scale of variances of technology metrics, with darker colour indicating a higher variance. The node text colours reflect different sectors. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

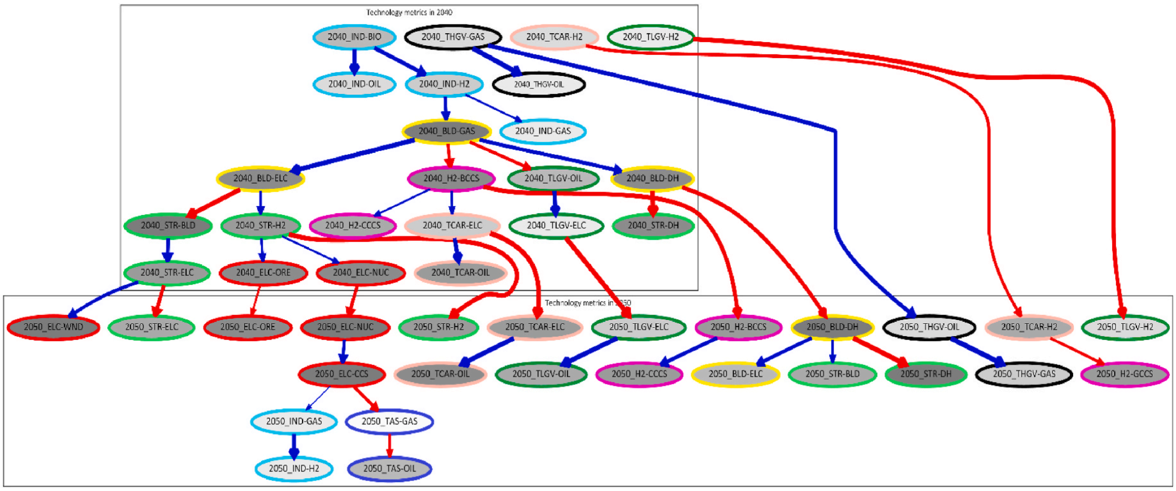


Fig. 2. Technological transition network for the transition between 2040 and 2050. Red arrows indicate a positive correlation between technologies; blue arrows indicate a negative correlation between technologies. Thicker links mean the correlations are stronger. The greyscale of nodes represents the scale of variances of technology metrics, with darker colour indicating a higher variance. The node text colours reflect different sectors. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

of technologies in 2050 is more likely to be influenced by technology combinations in 2040. However, technology combinations in 2030 might have lower influences on the deployment of technologies in 2040. This might imply that technological transition could be more uncertain and inconsistent across pathways in the period between 2030 and 2040, when there are still many overall strategies available for the mitigation, before 2050 decarbonisation targets and the impact of the common starting point for the optimization also features in the results. During the period between 2040 and 2050 results reflect more strongly the optimal trajectory with the given parameter set, with the flexibility for different, feasible technology portfolios further restricted by the stringency of the decarbonisation targets. This leads to the technology interdependency and substitution becoming more evident, and the correlations between technologies, in turn, are amplified.

3.1.1. *Associations with correlations*

The constructed BNs majorly represent strong positive and negative correlations between technology metrics that can maintain a high likelihood to the original pathway ensemble based on BIC. As shown in Figure D. 1 in Appendix D, strong positive correlations can usually be found between the same metrics in nearby two years. For instance, nuclear powers (ELC-NUC), district heating (BLD-DH), and oil vehicles (TCAR-OIL) fall in this category (see F igure D.1). Different metrics in the same year might still have strong correlations due to technology interdependency or substitution. For example, in 2030, hydrogen storage (STR-H2) increases as nuclear power (ELC-NUC) increases. The same positive correlation is also found between electricity consumption in the building sector (BLD-ELC) and energy storage in the same sector (STR-BLD) in 2040. In contrast, electricity consumption (BLD-ELC) and gas consumption (BLD-GAS) in the building sector have a strong negative correlation in 2030. The strong positive correlation between energy storage for district heating in two years has also been reflected with a route from 2030_STR-DH to 2040_STR-DH through 2030_BLD-DH and 2040_BLD-DH. These correlations are also represented as directed links in the Technology Transition Network for metrics in 2030 and 2040 in Fig. 1.

3.1.2. *Comparisons with a clustering analysis-based approach*

As BNs can represent both positive and negative correlations between metrics, the BN-based approach is hence superior to a clustering algorithm-based approach in terms of revealing the hidden technological correlations in uncertain pathways. According to a previous study (Pye et al., 2019), the hierarchical clustering algorithm can group positively correlated technology metrics together based on correlation coefficients between metrics. The same approach was applied to group technology metrics for two periods (between 2030 and 2040 and between 2040 and 2050) for comparisons. Unlike the previous study, technology metrics in multiple years were considered together for groupings. The results are illustrated in Figure E. 1 and Figure E. 2 in Appendix E. The dissimilarity indicates the correlation between metrics and is defined as one minus correlation coefficient between metrics. Lower dissimilarity means there is a stronger positive correlation between metrics. Similar to the built BNs, those with strong positive correlations, such as electric vehicles (TCAR-ELC) in 2040 and 2050 and oil vehicles (TCAR-OIL) in 2040 and 2050, as shown in Figure E. 2. However, the negative correlations between TCAR-ELC and TCAR-OIL are explicitly represented only in the BN, as shown in Fig. 2, but not in the dendrogram created by the clustering algorithm, as illustrated in Figure E. 2. Thus, the results suggest that, unlike clustering, BNs can effectively represent both positive and negative correlations between metrics in a visual way that shows the dynamics of technology transition across years, hidden in the extremely complicated pathway dataset.

3.1.3. *Transport sector*

Technology metrics in the transport sectors are more likely to be independent of the rest of the systems as those metrics are separated from the rest of the networks. This might imply that the correlations between the transport metrics and those in the other sectors might not be strong enough or are not consistent across interconnected transport metrics. As a result, those correlations are discarded from the BNs.

In the first period (2030–2040), the deployment of electric vehicles, including passenger (TCAR-ELC) and light good vehicles (TLGV-ELC), can be considered independently as the electricity consumption might remain limited, compared to electricity consumption in the other sectors. Moreover, GHG budgets are still substantial enough to allow the introduction of new oil-powered vehicles. The substitutional relationship between electric and oil vehicles, however, is evident. The higher market share of electric vehicles (2030_TCAR-ELC and 2030_TLGV-ELC) in 2030 will lead to a higher uptake of electric vehicles (2040_TCAR-ELC and 2040_TLGV-ELC) and a lower share of oil vehicles (2040_TCAR-OIL and 2040_TLGV-OIL) in 2040, as shown in Fig. 1.

As for the second period (2040–2050), with stringent GHG budgets, the increase in new oil vehicles (e.g. TCAR-OIL) can only be possible if bioenergy with CCS (BECCS) metrics (H2-BCCS) ramp up, as shown in

Fig. 2. The correlations between transport metrics and the other metrics become stronger, which means that the technology portfolios in the transport sector should be considered along with technologies in the other sectors. The technological transition dynamic of these metrics over time is basically the same as that in the previous period. On the other hand, the development of hydrogen vehicles (e.g. TCAR-H2) has less correlation with the other sectors and is majorly determined by the adoption of hydrogen production by gas (SMR) with CCS (H2-GCCS).

3.1.4. Building sector

Building metrics (e.g. BLD-ELC and BLD-GAS) show obvious substitutional relationships between one another in the same years, but only the technological transition of district heating (BLD-DH) is clear over time in both periods.

For instance, as shown in Fig. 1, bioenergy consumption (2030_BLD-BIO) has a strong negative correlation with gas consumption (2030_BLD-GAS) in the building sector in 2030. However, there is no strong direct link between 2030_BLD-GAS and 2040_BLD-GAS, which implies that the technological transition of these metrics over time is not obvious. The development of these technologies should be based on the technology portfolios in the other sectors in the same year. Similar dynamics for the building metrics have also been found in the second period, as shown in Fig. 2. District heating (BLD-DH) and electrification (BLD-ELC) can be applied to replace gas consumption (BLD-GAS) in the building sector in 2040, while, in 2050, the substitutional relationship can be found between BLD-ELC and BLD-DH. District heating shows a clear technological transition over time in the second period again, as 2040_BLD-DH and 2050_BLD-DH are directly connected.

The strong technological transition of district heating over time in both periods means that the development of district heating should be persistent to scale up once it is adopted to decarbonise the building sector. Otherwise, the investments in the infrastructure could dramatically increase the total energy system costs as the economy of scale might not be achieved for the heating sector and the district heating system might not be able to be operated cost-effectively.

3.1.5. Power sector

The strong technological transition over two periods can also be observed in the power sector. A high deployment of nuclear power (ELC-NUC) in 2030 likely leads to high shares of nuclear power in both 2040 and 2050, as suggested by the red links between the same technologies in two years. The positive correlation is stronger in the later period (2040–2050) than in the earlier period (2030–2040). This might imply that nuclear power should keep ramping up to decarbonise the power sector as nuclear power is relatively cost-effective in the model. The constant deployment of nuclear power can even further bring down the capital costs of nuclear power due to the accumulation of experienced manpower and the establishment of supply chains.

However, the substitutional relationship is not evident in the power sector. The only exception is the negative correlation between nuclear power (ELC-NUC) and electricity generation with CCS (ELC-CCS) in 2050. As both technologies are capital-intensive, this might suggest that these two low-carbon power technologies compete for limited resources to scale up.

Finally, the power sector also shows strong interactions with other sectors. For instance, in 2040, as the electrification level in the building sector (BLD-ELC) increases, renewable energy metrics (i.e. ELC-WND and ELC-ORD) are also more likely to increase to provide additional electricity to meet the increased demand, as shown in Fig. 1. There are two negative links connecting 2040_BLD-ELC and 2040_ELC-ORD, meaning the two metrics have a positive correlation. Similar relationships can also be found in the later period, as shown in Fig. 2.

Furthermore, the demand profile in the building sector (BLD-ELC) might be similar to the supply curve of wind power (ELC-WND). The need for storage technologies (STR-ELC), in turn, decreases as BLD-ELC increases. Moreover, the BNs also suggest that energy storage in the building sector (STR-BLD) could be a more cost-effective measure to balance electricity supply and demand than STR-ELC does, as indicated by the blue link between STR-BLD and STR-ELC in both figures.

3.2. Representative scenarios in 2050

Representative scenarios in terms of technology portfolios in 2050 are identified with the k-means clustering algorithm to demonstrate how the built BNs can facilitate policymakers to determine transition strategies to implement ambitious technology mixes which will be discussed in Section 3.4.

The k-means clustering algorithm has been found to perform well in determining representative decarbonisation scenarios from a pathway ensemble (Li et al., 2020). This is because long-term decarbonisation pathways are likely to be closely distributed around certain areas in a high-dimensional feature space. The k-means algorithm is thus a suitable procedure to group similar pathways into clusters based on dissimilarity between pathways. The average pathways of the identified clusters can then be regarded as representative ones. More details of the approach can be found in the previous study (Li et al., 2020). The same approach was adopted to discover five representative scenarios in 2050 from the pathway ensemble. These scenarios are shown in Table 1. The plus sign indicates a higher level of the corresponding metric than the average deployment level. The minus sign represents the opposite. A higher number of signs means a stronger deviation of the metric from the mean level. The choice of a proper number of clusters and the deviations of metrics for five identified scenarios can be found in Appendix F. The characteristics of the five scenarios are briefly introduced as follows.

The first scenario (high BECCS) produces more hydrogen by BECCS technologies which create net negative emissions that allow the deployment of more oil vehicles by 2050. In this scenario, the building sector is decarbonised with more district heating. As CCS technologies might be cost-effective, more electricity is generated by power plants with CCS to replace some wind power generation.

The second scenario (high electrification) deeply electrifies the energy system to reduce total GHG emissions, especially in the building and transport sectors. More electric appliances are used to replace fossil-fuel powered appliances to decarbonise the building sector. Furthermore, abundant oil vehicles are replaced with electric vehicles to avoid GHG emissions from oil vehicles. The increased electricity consumption is then met with more low-carbon electricity generated by wind power and power plants with CCS.

As for the third scenario (CCS-central), CCS-fitted technologies are adopted more widely across sectors probably due to the low costs of CCS in this scenario. A high share of electricity is supplied by CCS-fitted power plants, which replace other kinds of plants, such as nuclear and wind power. Furthermore, more hydrogen is produced by BECCS as it could be cheaper in this case. In turn, more oil vehicles remain in the market by 2050.

In the fourth scenario (high hydrogen and electrification), abundant hydrogen is produced with a wide range of production technologies. The higher production of hydrogen enables hydrogen-fuelled technologies in different sectors. More hydrogen is not only used to power more hydrogen vehicles in the transport sector, but also is used in the widely adopted district heating. More low-carbon electricity is generated to meet the increased demand from a higher share of electric vehicles. Specifically, more nuclear power is adopted while less CCS-fitted power plants are in operation.

Finally, the fifth scenario (low CCS) might assume higher costs of CCS so that lower CCS-fitted technologies are adopted. In the power sector, much less electricity is generated by CCS-fitted power plants; other low-carbon power plants are thus adopted to produce more electricity. More light good vehicles can then be powered by the surplus low-carbon electricity in this scenario.

Table 1
Deviation level of metrics in 2050 for five representative scenarios from the corresponding mean metrics in 2050 of all scenarios. DH: district heating; ELC: electricity; GAS: gas; NUC: nuclear; ORE: other renewable energy; WND: wind; H2: hydrogen; BCCS: bioenergy + CCS; CCCS: coal + CCS; GCCS: gas + CCS; AV: aviation; CAR: passenger car; HGV: heavy-good vehicle; LGV: light-good vehicle; +/−: positive/negative variation less than 5 TWh; ++/−: positive/negative variation about 10 TWh; +++/−: positive/negative variation about or more than 15 TWh.

Sector			Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Building		DH	++		–	+++	–
		ELC		+	+	–	
Electricity		CCS	++	++	+++	–	–
		NUC		–	–	+++	++
		ORE			–	+	+
		WND	–	+	–	++	++
Hydrogen		BCCS	+	–	+	–	
		CCCS	–	+	–	++	+
		GCCS				++	
Transport	AV	GAS					
		OIL				–	
	CAR	ELC	–	++	–	+	
		H2			–	++	+
		OIL	++	–	++	–	
	HGV	GAS					
		OIL					
	LGV	ELC				+	+
		H2				+	
		OIL			+	–	–

3.3. Uncertain technological transitions

The BNs then estimated the probability distribution of metrics in the prior years (2030 and 2040) based on the probabilistic relationships between metrics established in the networks, given the identified technological portfolios in the five scenarios. The uncertain metrics in 2040 were firstly evaluated with the BN for the later period (2040–2050), as shown in Fig. 2 and Figure C2. The BN for the first period (2030–2040) (Fig. 1 and Figure C1), in turn, determined the probability distribution of metric levels in 2030 using those estimated uncertain metrics in 2040. A BN library, pyAgrum (AGrUM/pyAgrum Team, 2023), was adopted to conduct the inference for 2030 metrics based on the estimated probabilistic evidence of 2040 metrics. For the sake of visualisation, only the most probable metric levels in previous years (2030 and 2040) for five scenarios are presented in Figs. 3–7. The number of stacked circles shows the level of a metric. As all metrics are transformed into discrete variables with five levels based on the distribution of original metric values (see Section 2.3), metric levels are not comparable across metrics. Empty cells in the figures are for those metrics with no clear probabilistic relationships with other metrics in the BNs. The colour scale of circles represents the certainty of the presented metric levels. A darker colour suggests the corresponding metric level has a higher certainty, with a narrower probability distribution of metric levels, i.e. the variance of the determined probabilistic distribution, given known technology choices, is used to represent the uncertainty of the adoption of a specific technology in individual pathway prototypes. Metrics in 2050 are given and hence marked in the darkest colour.

Clearly, metric levels in 2050 are influenced by the status of other metrics in the same year, as suggested by the probabilistic relationships between metrics in Fig. 2. For instance, nuclear power (ELC-NUC) is much lower as CCS-fitted power plants (ELC-CCS) generate more electricity, as shown in Fig. 5. The opposite can be found in scenario four (Fig. 6). Similar substitutional relationships between electric and oil vehicles in the transport sector and between district heating and electrification in the building sector can be observed across all five scenarios.

The BNs further reveal intertemporal transitions, following the cross-year probabilistic relationships represented in Figs. 1 and 2. There are two types of transitions: (1) temporally persistent transition and (2) temporally versatile transition. The former includes technologies that have long lifetimes and require heavy investments in facilitating infrastructure, such as nuclear power plants (ELC-NUC) and district heating

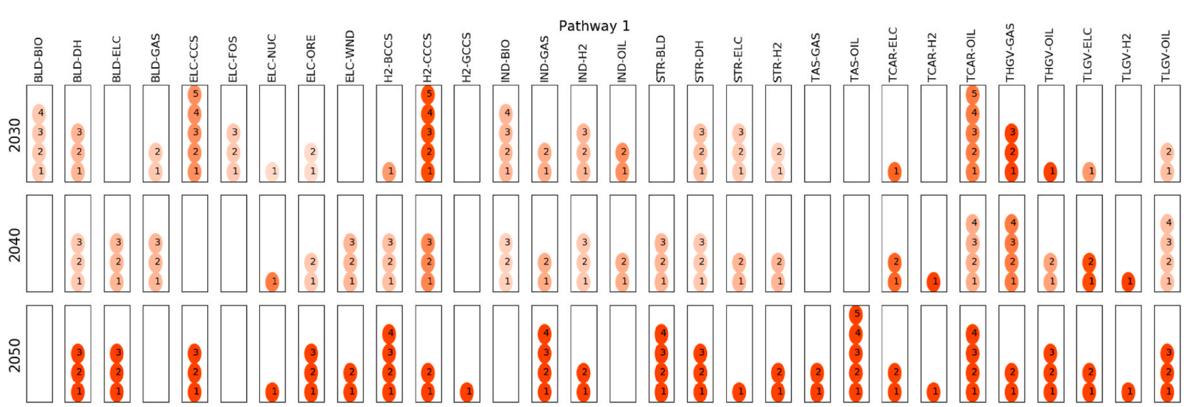


Fig. 3. Uncertain technological transition for scenario 1 (high BECCS). The number of stacked circles shows the discretized level of deployment (1–5). Darker colours represent higher certainty (lower variance). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

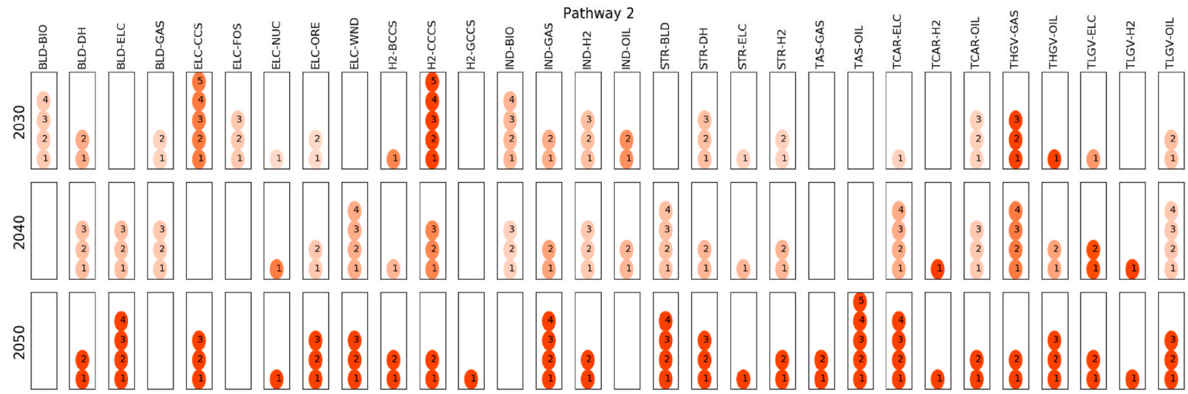


Fig. 4. Uncertain technological transition for scenario 2 (high electrification). The number of stacked circles shows the discretized level of deployment (1–5). Darker colours represent higher certainty (lower variance). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

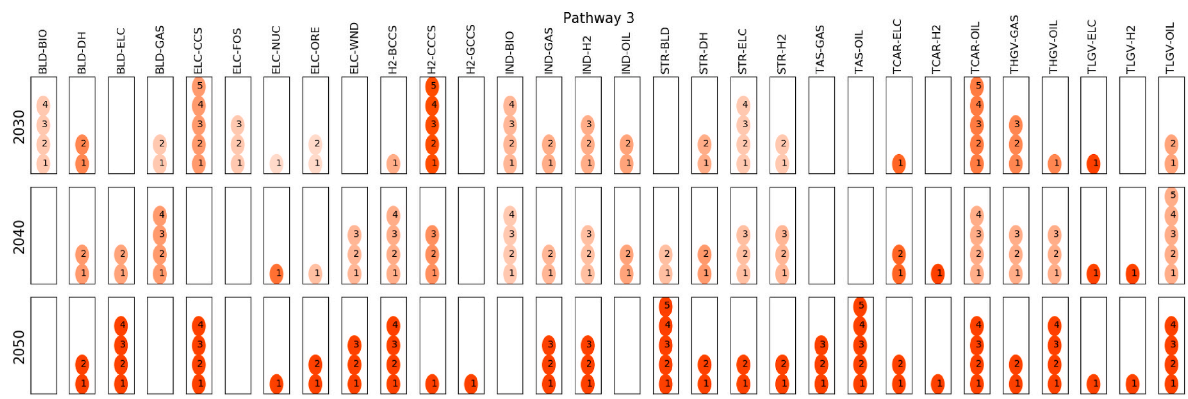


Fig. 5. Uncertain technological transition for scenario 3 (CCS-central). The number of stacked circles shows the discretized level of deployment (1–5). Darker colours represent higher certainty (lower variance). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

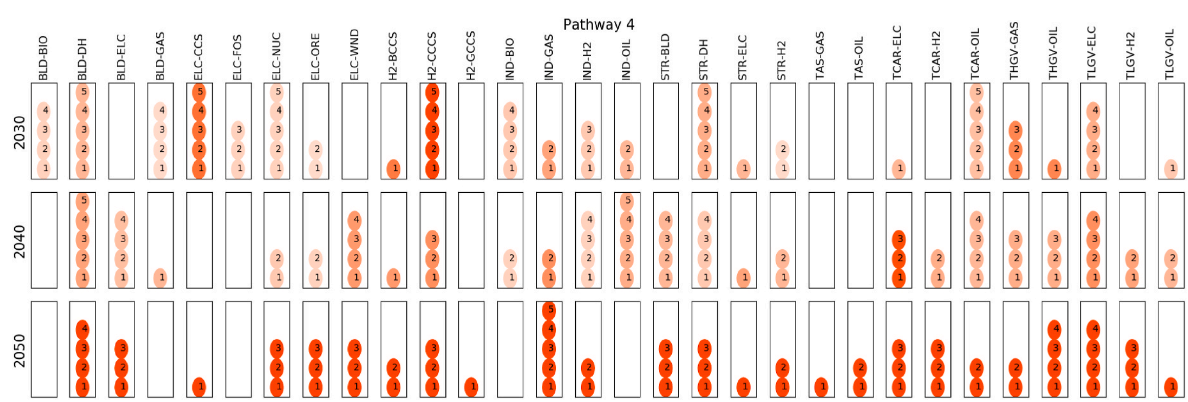


Fig. 6. Uncertain technological transition for scenario 4 (high hydrogen and electrification). The number of stacked circles shows the discretized level of deployment (1–5). Darker colours represent higher certainty (lower variance). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(BLD-DH). Levels of these technological metrics are usually consistent over time, which means that early actions are essential to reach high market shares for these technologies. For instance, in Fig. 6, nuclear power is constantly at a higher level across all three milestone years, compared with those in other scenarios. Similarly, a high deployment of district heating at an early stage is essential for a high share of district heating in the residential heating sector by 2050, as shown in Fig. 6. The temporal persistency of these technologies is due to the direct probabilistic relationships of the same technologies between different years in the BNs (Figs. 1 and 2).

On the other hand, when a technology has a shorter lifetime without strict deployment requirements, it is more probable to ramp up the technology deployment to replace carbon-intensive alternatives in a relatively short period of time. Consequently, metric levels of these technologies are more versatile across three milestone years, which means a high share of a technology in 2050 does not suggest the

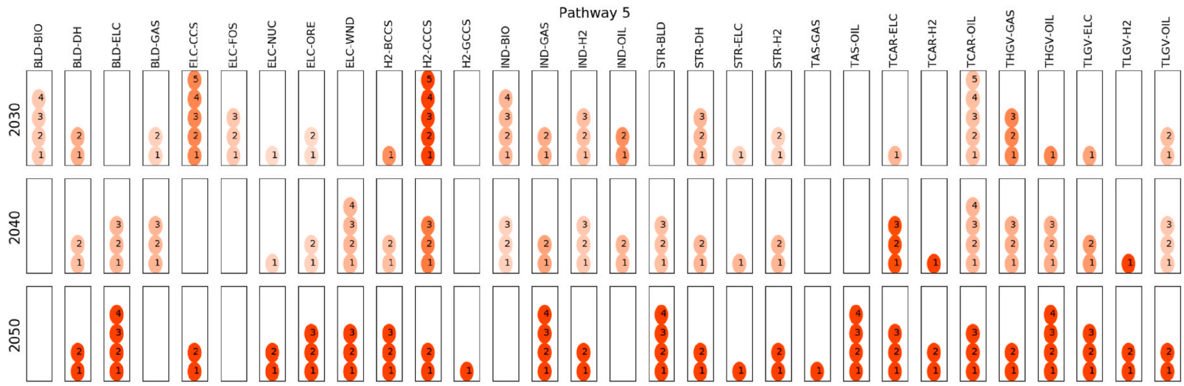


Fig. 7. Uncertain technological transition for scenario 5 (low CCS). The number of stacked circles shows the discretized level of deployment (1–5). Darker colours represent higher certainty (lower variance). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

technology should also be deployed at a high level in the prior years. There are many technologies falling into this category, such as wind power (ELC-WND) and building electrification (BLD-ELC). Cross-year connections are weak or even absent for these technologies in the BNs (Figs. 1 and 2). Hence, no clear level of the metrics is presented in 2030 in Figs. 3–7, suggesting that early actions are not essential for these technologies.

Furthermore, the strengthening of GHG emission targets also drives the energy system transitions in terms of technology replacements during the two decades before 2050. This is particularly evident in the transport sector. Low-carbon vehicles (e.g. electric and hydrogen vehicles, denoted as TCAR-ELC, TCAR-H2, TLGV-ELC, and TLGV-H2) are scaling-up over time while oil vehicles (e.g. TCAR-OIL) are phasing out. Positive correlations of these technologies across years in the BNs explicitly represent these trajectories. It is noteworthy that the cross-year correlation for hydrogen vehicles only exist in the BN for the later period (2040–2050) since those vehicles are unlikely to be adopted at scale at the early stage due to their relatively high prices. However, when oil vehicles remain at a high level in 2050, it is not sufficient to offset the GHG emissions by simply deploying more low-carbon vehicles, as shown in Figs. 3 and 5. Introducing more BECCS technologies (hydrogen production by BECCS, denoted as H2-BCCS) from the prior year (2040) becomes critical, indicated by higher metric levels in both 2040 and 2050. This intertemporal relationship can be found in Fig. 2. Oil vehicles in 2050 (2050_TCAR-OIL) has a positive correlation with hydrogen production by BECCS in 2040 (2040_H2-BCCS), which in turn directly influences the deployment of hydrogen production by BECCS in 2050 (2050_H2-BCCS).

Finally, across all five scenarios, fossil fuel power plants with CCS (ELC-CCS) and hydrogen production with coal gasification with CCS (H2-CCCS) in 2030 all remain at a high level, with a relatively high certainty, marked in a dark colour. This implies that these CCS-fitted technologies are likely to be essential measures to reduce GHG emissions at the early stage while other low-carbon technologies are still too costly to deploy at scale or their facilitating infrastructure is not in place yet. Interestingly, metric levels of these technologies are lower in the later years due to the strengthened reduction targets since operating these technologies still emits GHGs.

Nonetheless, technological transitions remain highly uncertain, given many metrics marked in a light colour in Figs. 3–7. This is because the uncertainty of metric levels is amplified while propagating towards the prior years in the BNs. The high uncertainty of technological transition might be due to two reasons. Firstly, there are multiple strategies to substitute carbon-intensive technologies over time. As a result, the transition of low-carbon technologies is not always in one single direction. This is in line with the fact that metrics can be influenced by other metrics both in the same year and those in different years. Secondly, the installation rate of low-carbon technologies might be high enough to ramp up the deployment of specific technologies in a relatively short period of time. The influences of metric levels are therefore not persistent across years.

3.4. Influential technologies in energy system transitions

Uncertainty of technological transitions can be reduced once policy targets of individual technologies are set. In other words, a specific decarbonisation scenario becomes more feasible in order to reach deep decarbonisation targets, given known deployment levels of individual technologies. However, due to the probabilistic relationships between technologies, individual technology metrics have various system-wide influences on the technology mixes. Policymakers should, hence, pay more attention to highly influential technologies so that they can have a clear idea about how to transform the system.

In this study, we used weighted importance (defined in Section 2.4) to determine the system-wide influences of individual technologies, as shown in Figs. 8 and 9. Basically, a technology metric has high system-wide influences if it is highly uncertain and has strong probabilistic relationships with other metrics with high uncertainties across pathways. Overall, technology metrics are majorly influenced by other metrics in the same year. Even though in fewer cases, evident influences from the prior years can still be found.

In the period between 2030 and 2040 (Fig. 8), technology metrics in the building and the power sectors have stronger impacts on the system transitions than the other sectors. The extremely high importance of the building sector is majorly due to two reasons: (1) multiple heating measures are still feasible, with relatively relaxed carbon limitations imposed, and hence the deployment of heating technologies is highly uncertain; (2) cross-sectoral technologies should be substituted in response to heating technologies adopted in the building sector. As suggested in Fig. 1 and Figure G. 1 in Appendix G, the increase in electricity demand in the building sector (2040_BLD-ELC) forces the significant adoption of wind power and lowers demands for hydrogen and gas in the building sector.

On the other hand, district heating and nuclear power are regarded as important due to their inter-temporal influences on the transitions, as shown in Fig. 1 and Figure G. 2. The higher adoption of district heating can reduce electricity demand; in turn, the demand profile becomes more stable, and hence nuclear power can ramp up for electricity provision. Moreover, strong inter-temporal relationships further expand these technologies' influences because of their long construction period and lifetime.

Unlike the previous period, in the later period between 2040 and 2050 (Fig. 9), the power sector is more influential to the system transitions. This is due to a sharp increase in electricity demand for high

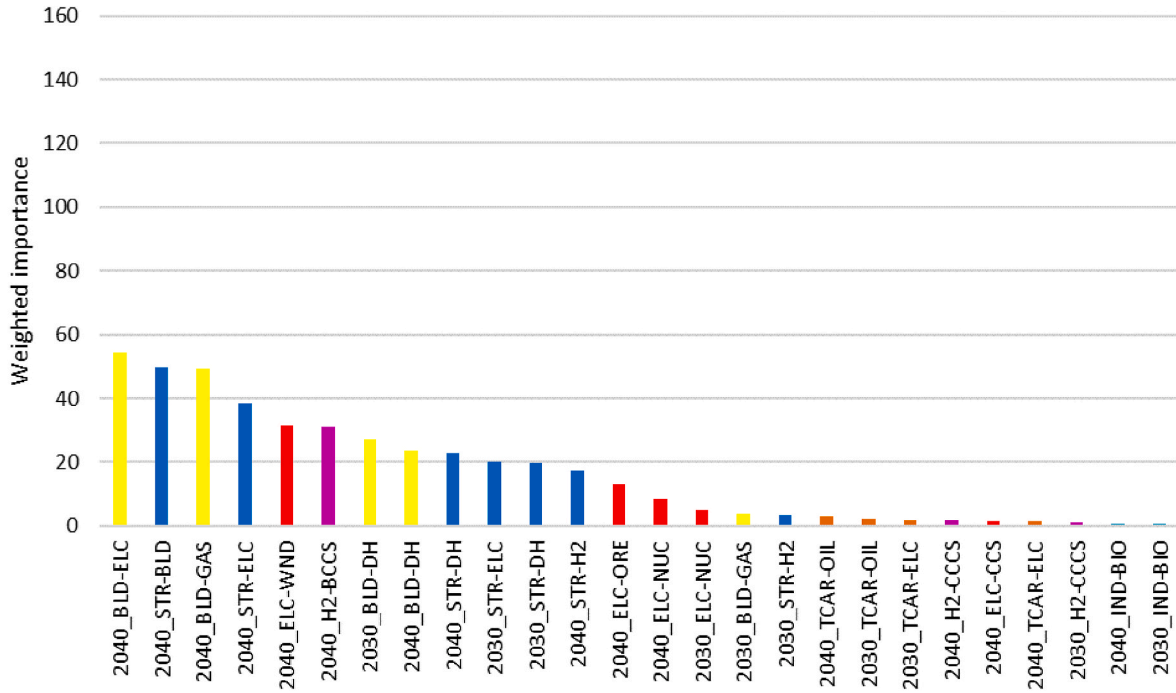


Fig. 8. Weighted importance of metrics in 2030 and 2040 in terms of the reduction in weighted uncertainty of metrics in 2040. The colours of bars represent different sectors. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

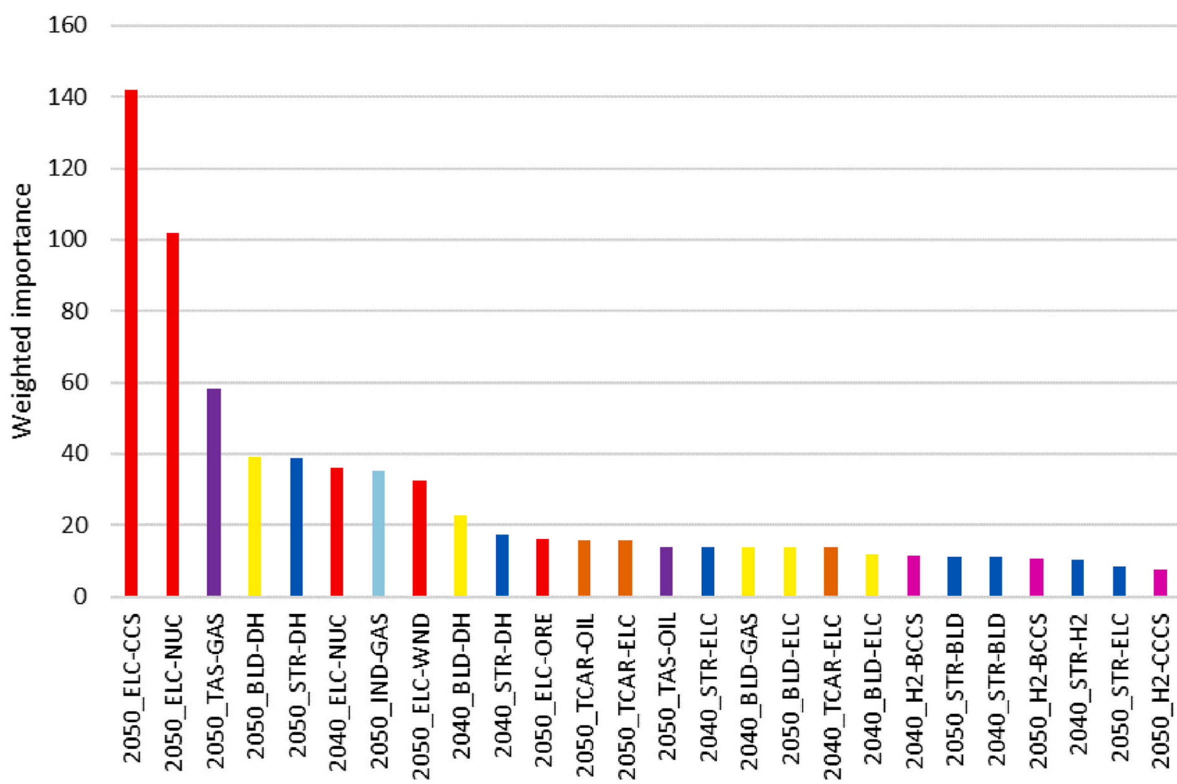


Fig. 9. Weighted importance of metrics in 2040 and 2050 in terms of the reduction in weighted uncertainty of metrics in 2050. The colours of bars represent different sectors. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

electrification levels across all sectors approaching 2050, along with substitutional technologies for low-carbon electricity provision, such as CCS-fitted, nuclear, and wind power plants.

In the building sector, only district heating remains highly influential as its long lifetime further propagates the uncertainty of district heating from the prior period to the later period. Eventually, its strong probabilistic relationships with technologies in other sectors (Figure G. 4) lead to significant impacts on the system transitions.

Finally, the importance of oil and electric vehicles (TCAR-OIL and TCAR-ELC) is more observable than in the previous period. Higher transport demand in the later period increases the influences of vehicle technology choices on other sectors, as illustrated in Fig. 2. Additionally, the replacement of oil vehicles remains highly uncertain since residual emissions from the transport sector can be offset by negative emissions technologies, such as hydrogen production by BECCS (H2-BCCS), whenever necessary. These two technology metrics hence also require more attention. For more detailed explanations on the importance of technology metrics, please refer to Appendix G.

4. Conclusion and policy implications

The proposed BN-based approach has been shown to be able to reflect strong correlation relationships between technologies hidden in uncertain pathway ensembles. Those relationships are represented in graphical networks that visualise complicated technology entanglements across sectors and years intuitively. Unlike other approaches, such as clustering algorithms, not only is technology interdependency (positive correlation) captured, but also technology substitution (negative correlation) can be represented in a single statistical structure. Moreover, technology deployments in different years can be taken into account within the same framework, without distorting major technology correlations in an energy system. This approach is thus, to our knowledge, the first framework that can successfully reveal temporal technology transitions under uncertain decarbonisation pathways. With the embedded joint probability distributions at metric points, the constructed BNs can be further applied to estimate probable impacts of specific technological deployment strategies dynamically. Especially, highly influential technologies to the technology portfolios can be determined based on the expected reduction in mutual information between metrics using the statistical information in BNs. This analysis approach hence provides a systematic way to investigate dynamic relationships between technologies in uncertain pathways, to facilitate energy system planning and related policymaking.

Nonetheless, the BN-based approach does not replace other advanced data analytics measures, such as clustering-based approaches, in assessing a large ensemble of scenarios. Instead, multiple approaches can be applied in parallel to have a more comprehensive view on the technological dynamics across the complex uncertainty space in which the pathways exist. For instance, as shown in this study, the k-means algorithm can identify a limited number of representative technological scenarios from a larger scenario ensemble, to enhance the understanding of the technology correlations.

Beyond the methodology, several insights relevant to the UK policy have also been found from this study. First, attention should be paid to the technology transition in the building sector at an early stage as the majority of the building stock is still fitted with carbon-intensive facilities. Compared with other sectors, it is more cost-effective to decarbonise the building sector, and thus consumers' participation plays a particularly critical role in the transition (Li et al., 2018). Moreover, the increase in electrification of the building sector could lead to major adjustments across sectors, such as an increase in wind power production and a reduction in hydrogen production via BECCS. Finally, the decarbonisation of the buildings sector relies heavily on replacing gas with district heating and electricity, but as the former is slow to diffuse, the lack of early district heat diffusion would suggest that it may be better to focus on directly electrifying the sector instead.

Second, close attention should also be paid to the power sector and related electrification of end-use demands, as the power sector was, next to the buildings sector, identified as particularly influential for defining the transition. It appears likely, based on our analysis, that low-carbon power plant capacity should, towards mid-century, be dramatically ramped up to meet the sharp increase in electricity consumption following the high level of electrification across sectors. However, the decarbonisation strategy for the power sector is to be determined, as nuclear power and CCS-fitted power plants can crowd each other out and fluctuate widely across pathways. Based on the analysis, policymakers may want to consider focusing R&D and other support activities to one of the two, but not necessarily both. With that said, there is more room for the "let all flowers bloom" approach in the power sector than in some other sectors, according to our analysis, at least for cases in which electrification of the full system is extensive. Conversely, policymakers can facilitate electrification of the building sector with higher deployment of wind power, as the supply profile of wind power may be similar to the demand profile of the building sector. For energy storage, there is a substitutional relationship between building-level energy storage for heat and that of stand-alone electricity storage technologies, such as lithium-ion batteries. In our analysis the former tends to be the better option, at least in the cases in which the buildings sector is heavily electrified. This suggests that cost-effective storage solutions may be closely linked to choices made for buildings' heating systems and policymakers should thus closely monitor the developments there and coordinate their support efforts to exploit the synergies and focus on the technologies that support each other.

Third, once policymakers decide, or not, on the deployment of district heating to be a major measure to decarbonise the building sector, the decision should be persistent over time. This is not only because its deployment rate is relatively slow, but also because of its high impacts across sectors. Furthermore, the scale-up of district heating can also benefit from economies of scale, and bring down the costs of district heating in the long run.

Finally, BECCS technology (hydrogen production by BECCS) has also been found critical to offset GHG emissions from fossil-fuel powered technologies, such as vehicles using oil products and gas consumption in the building sector. Therefore, in the situation that the more carbon-intensive technologies cannot be phased out quickly, policymakers should opt for promoting BECCS technologies. Conversely, if the technology makes a breakthrough, this greatly reduces the pressure from other mitigation measures, such as the speed with which cars need to be electrified. The flexibility BECCS could offer to the system suggests that the attention the technology has received so far is well deserved and the R&D efforts should continue to be pursued.

The BN-based approach can, however, only reveal the technology correlations embedded in pathway ensembles, which in turn are generated by energy system models. The revealed technological relationships are thus predetermined by the assumptions adopted in the energy system models used. For instance, the temporal evolution of technology deployment may be affected by the maximum installation rates assumed in the model for the low-carbon technologies. Lower deployment rates could lead to stronger temporal correlations since reaching a specific level of technology deployment in the future would require an earlier start for the transition. Moreover, this study adopted the decarbonisation pathways using 80% reduction targets. In pathways with net-zero emissions targets, technological transitions could be further accelerated and the role of certain low-carbon technologies, such as BECCS, might be strengthened. Consequently, the technology dynamics in those pathways could well be different from what we found in this study, with less flexibility in choosing alternative mitigation strategies (Keppo and van der Zwaan, 2012). In the future, the same approach should be applied to analyse uncertain pathways with different climate targets, including net-zero emissions targets, to further verify the findings of this study. Also, replicating the study using results from a different energy system model would be important, to better understand the implications of specific model structures for the conclusions.

The developed BN-based approach can analyse any decarbonisation pathway datasets generated by energy system models at multiple scales (from subnational to global). The complicated interactions between technologies hidden in the datasets are revealed and visualised with BNs intuitively. Policymakers can then have a better understanding of the influences of technology deployments across sectors and over time. Our results from this initial implementation suggest that policymakers should pay particular attention to the building sector (especially district heating and electrification), deployment of different power generation technologies, and diffusion of electric vehicles to transform the system. Systematic planning for low-carbon transitions can be incorporated into the policy-making process more easily to avoid unexpected impacts from promoting specific technologies.

CRediT authorship contribution statement

Pei-Hao Li: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Behzad Zamanipour:** Writing – review & editing, Visualization, Validation, Investigation. **Ilkka Keppo:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

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

Data will be made available on request.

Acknowledgements

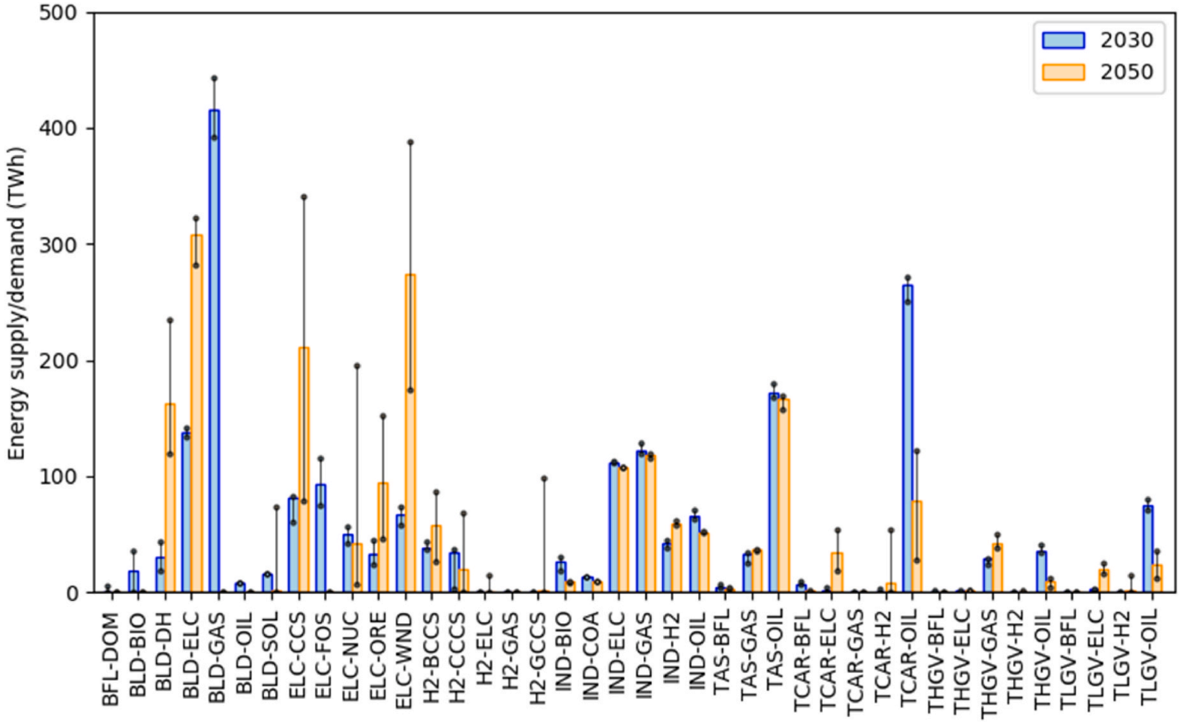
This work was supported by the REEEM project from the European Union's Horizon 2020 research and innovation programme under grant agreement 691739. PH.L. was also supported by UKERC 4 (EP/S029575/1) and EnergyREV Plus (EP/S031898/1) and I.K: also acknowledges the support from the Horizon Europe Project 'IAM COMPACT' (grant no. 101056306).

Appendix A. Definitions of technology metrics

Metric	Definition	Units
ELC-WND	Wind generation level	TWh
ELC-NUC	Nuclear generation level	TWh
ELC-CCS	CCS generation level	TWh
ELC-ORE	Other renewable generation level	TWh
ELC-FOS	Fossil generation level	TWh
BLD-BIO	Building bioenergy consumption	TWh
BLD-ELC	Building electricity consumption	TWh
BLD-GAS	Building gas consumption	TWh
BLD-OIL	Building oil consumption	TWh
BLD-DH	Building district heating consumption	TWh
BLD-SOL	Building solar energy consumption	TWh
H2-BCCS	H2 production by biomass gasification with CCS	TWh
H2-CCCS	H2 production by coal gasification with CCS	TWh
H2-ELC	H2 production by electrolysis	TWh
H2-GCCS	H2 production by gas (SMR) with CCS	TWh
H2-GAS	H2 production by gas (SMR)	TWh
IND-BIO	Industry bioenergy consumption	TWh
IND-COA	Industry coal consumption	TWh
IND-ELC	Industry electricity consumption	TWh
IND-GAS	Industry gas consumption	TWh
IND-H2	Industry hydrogen consumption	TWh
IND-OIL	Industry oil consumption	TWh
TAS-GAS	Aviation & shipping - gas	TWh
TAS-OIL	Aviation & shipping - oil	TWh
TAS-BFL	Aviation & shipping - biofuel	TWh
TCAR-ELC	Cars - electricity	TWh
TCAR-GAS	Cars - gas	TWh
TCAR-H2	Cars - H2	TWh
TCAR-OIL	Cars - oil	TWh
TCAR-BFL	Cars - biofuels	TWh
THGV-ELC	Heavy goods vehicles - electricity	TWh
THGV-GAS	Heavy goods vehicles - gas	TWh
THGV-H2	Heavy goods vehicles - H2	TWh
THGV-OIL	Heavy goods vehicles – oil	TWh
THGV-BFL	Heavy goods vehicles - biofuels	TWh
TLGV-ELC	Light goods vehicles - electricity	TWh
TLGV-H2	Light goods vehicles - H2	TWh
TLGV-OIL	Light goods vehicles - oil	TWh
TLGV-BFL	Light goods vehicles - biofuels	TWh

Note: Technology metrics in a specific year are marked by the year in the front of the metric names. For instance, 2050_ELC_WND means wind generation level in 2050.

Appendix B. Variance and average values of technology metrics in 2030 and 2050



Appendix C. Bayesian networks with distribution of metric levels

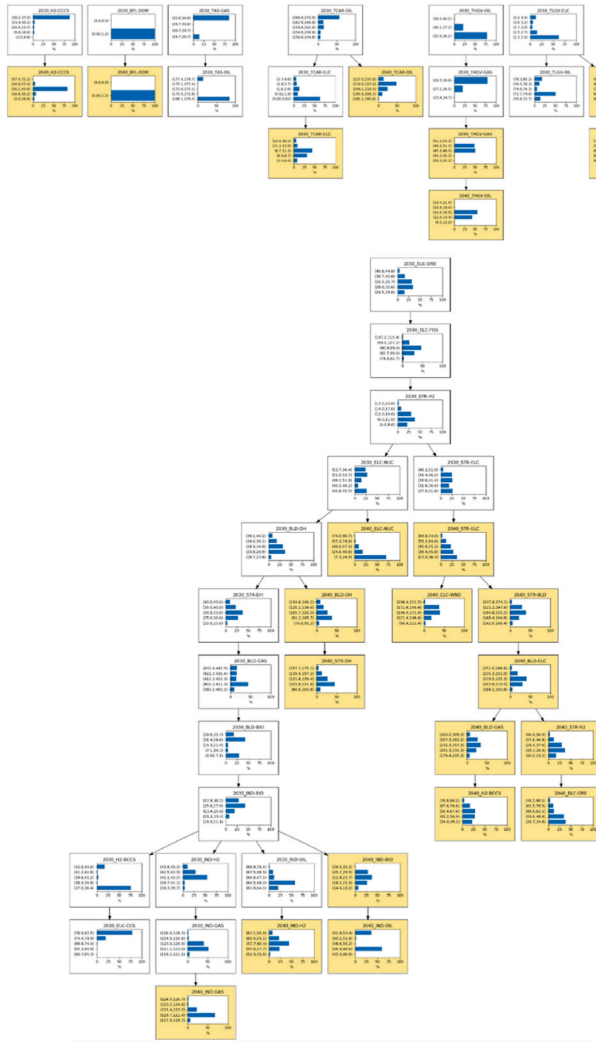


Fig. C. 1. Bayesian network for technology metrics in 2030 and 2040

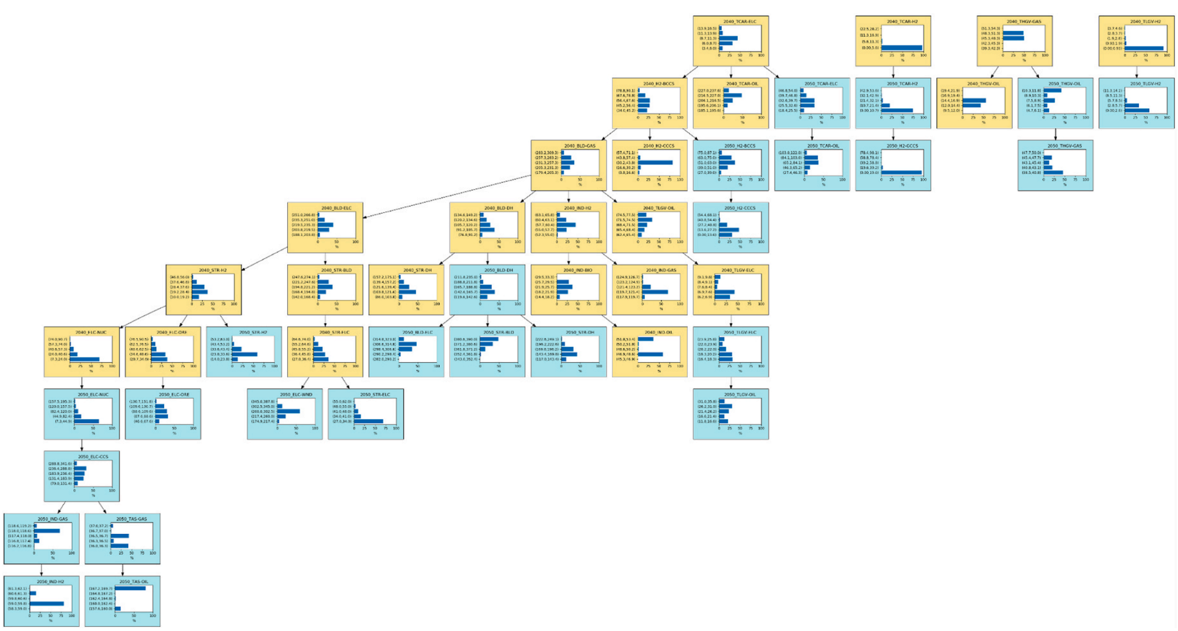


Fig. C. 2. Bayesian network for technology metrics in 2040 and 2050

Appendix D. Correlation matrix for technology metrics in different years

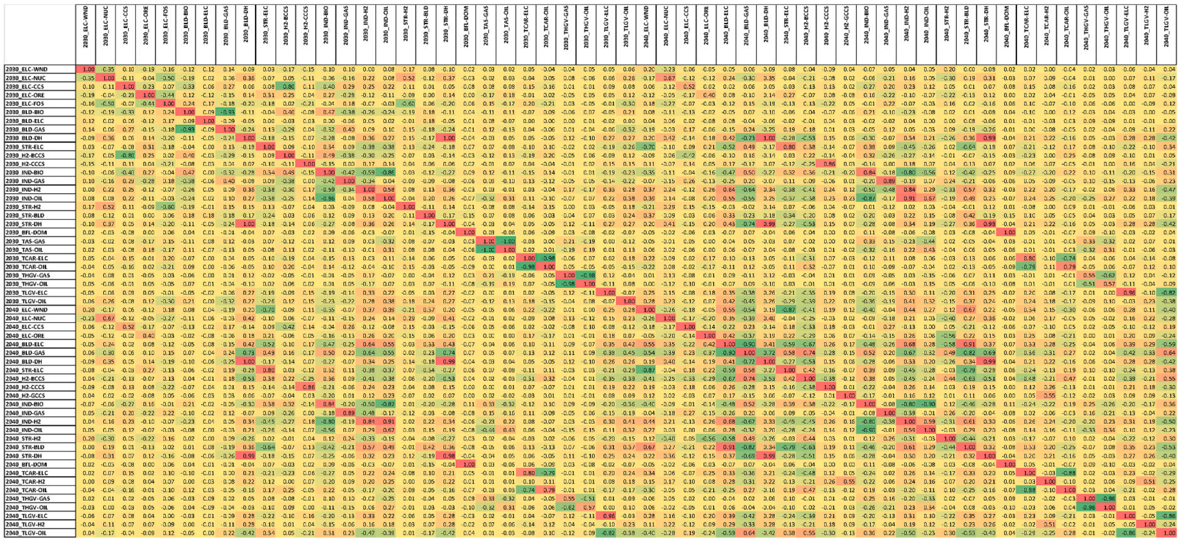


Fig. D. 1. Correlation between technology metrics in 2030 and 2040

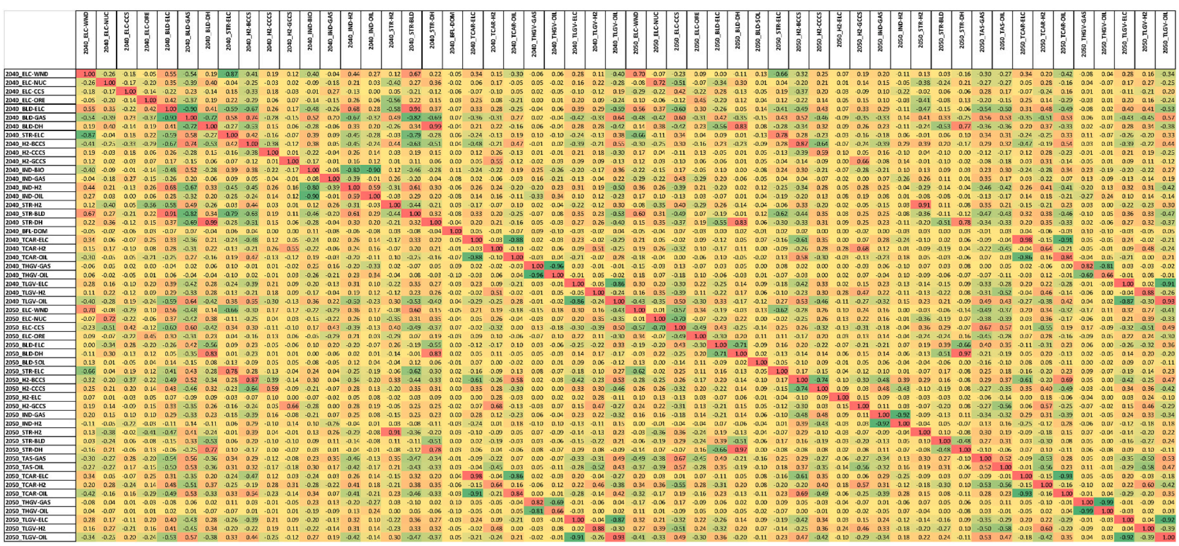


Fig. D. 2. Correlation between technology metrics in 2040 and 2050

Appendix E. Technology interdependency across pathways identified by Hierarchical Clustering

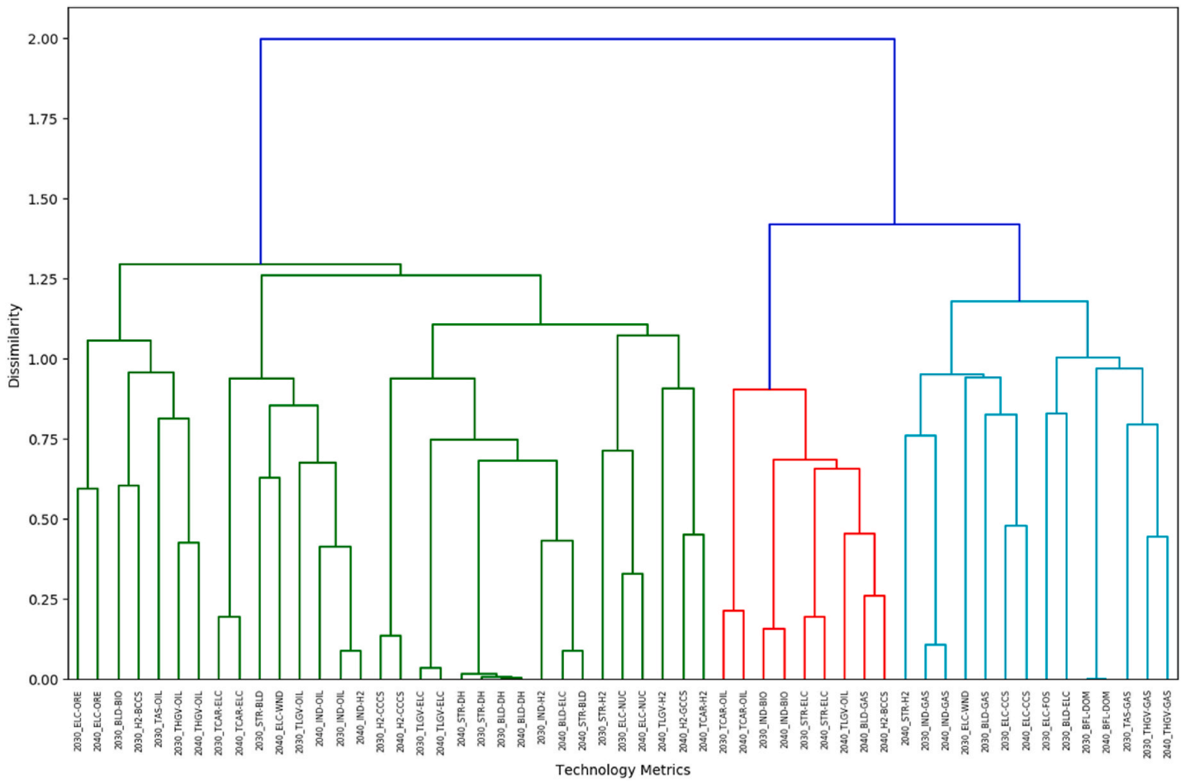


Fig. E. 1. Co-evolution of technologies in 2030 and 2040

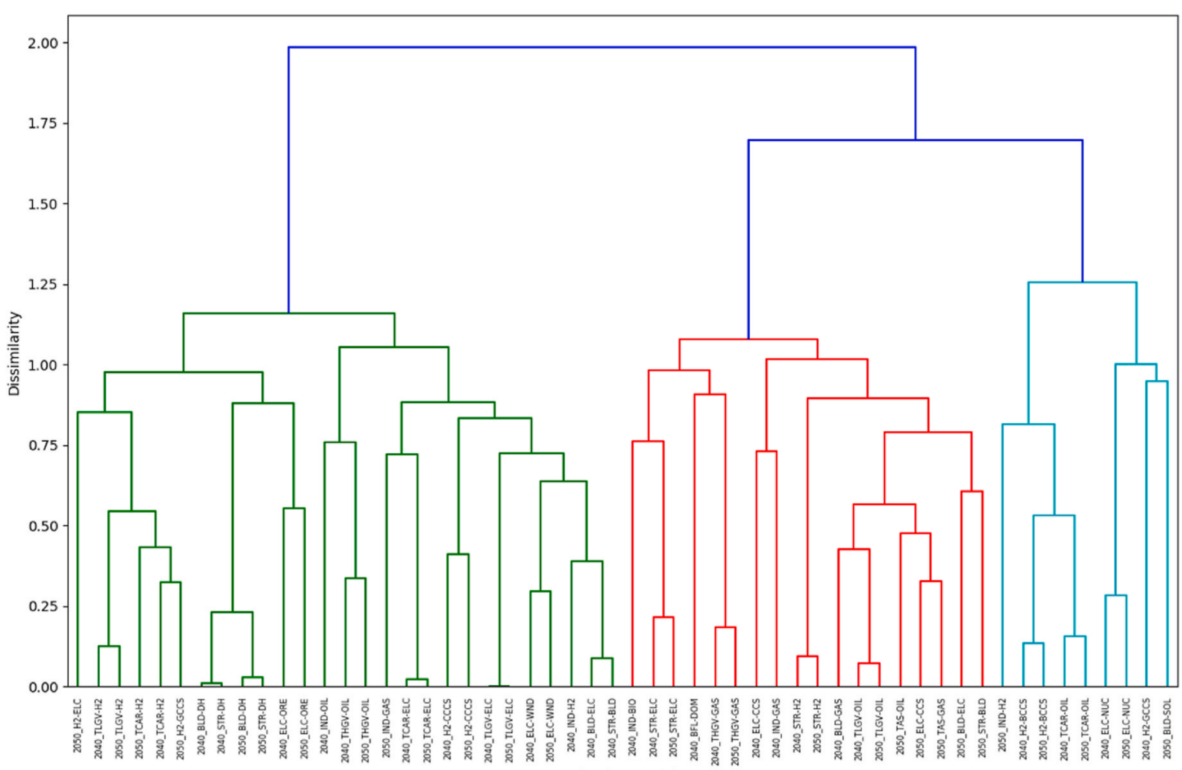


Fig. E. 2. Co-evolution of technologies in 2040 and 2050

Appendix F. Identification of representative scenarios in 2050 with K-means clustering algorithm

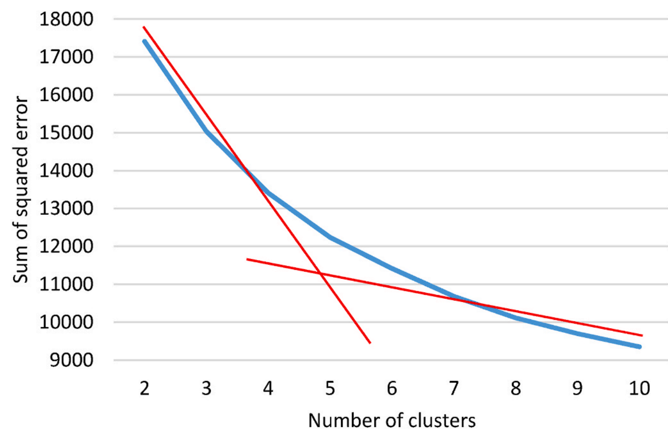


Fig. F. 1. Determination of a proper number of clusters with the elbow measure

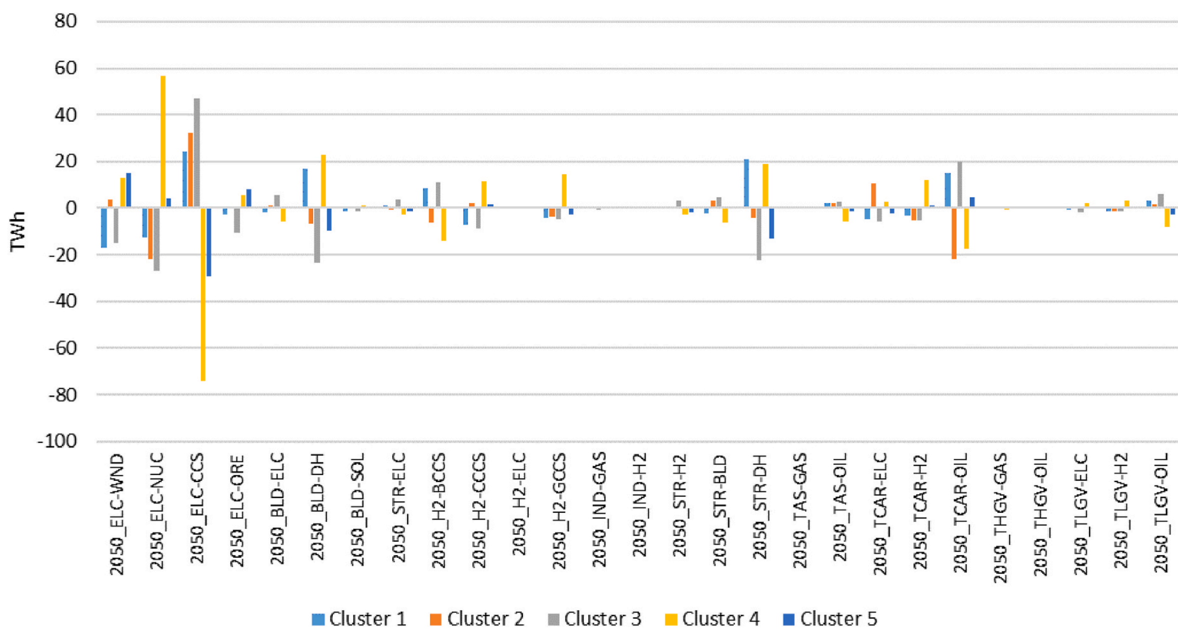


Fig. F. 2. Deviations from the average metrics for five representative technology portfolios in 2050

Appendix G. Influential technologies in energy system transition

The built BNs not only visualise the correlations between technological metrics, but also provide a framework to quantitatively determine the influences of individual metrics and to estimate possible distribution of metric states, given certain metrics being observed.

The proposed weighted importance analysis was applied to determine the importance of individual metrics in both periods (2030–2040 and 2040–2050) in terms of the reduction in weighted uncertainty of metrics as the corresponding metrics are given. A metric's weighted uncertainty is represented by its variance. The rankings of metrics by weighted importance index in both periods are illustrated in Figs. 8 and 9.

The results in Figs. 8 and 9 show that technology portfolios are majorly influenced by technology metrics in the same year, but significant impacts of some technologies from a previous year can also be found. As shown in Fig. 8, in 2040, metrics in the building sector are more influential than those in the other metrics, including electricity consumption (BLD-ELC), gas consumption (BLD-GAS), district heating (BLD-DH), and energy storage (STR-BLD) in the building sector. The deployment of district heating (2030_BLD-DH) in 2030 also has a considerable impact on technology portfolios in 2040. These influential metrics majorly have high correlations with other highly uncertain metrics, as shown in Figure G. 1 and Figure G. 2. The greyscale of nodes represents the reduction in weighted uncertainty of those nodes, given the observed metrics that are marked by red boundary.

As illustrated in Figure G. 1, influential technology metrics (e.g. BLD-ELC, BLD-GAS, and STR-BLD) are strongly interconnected in 2040; those metrics are usually highly uncertain across decarbonisation pathways. Likewise, the deployment of wind power (ELC-WND) is obviously influenced by the development of those technologies. On the other hand, as shown in Figure G. 2, the development of district heating in 2030 (2030_BLD-DH) can directly affect the deployment of districting heating infrastructure in 2040 (e.g. 2040_BLD-DH and 2040_STR-DH), which are highly uncertain. A higher share of district heating (2030_BLD-DH) might further imply that the demand profile of electricity is more stable so that low-carbon base-load power technologies, such as nuclear power (2030_ELC-NUC and 2040_ELC-NUC), can be deployed to provide more electricity.

Therefore, at the early stage, policymakers should pay more attention to determining how to decarbonise the building sector since those metrics are more influential and uncertain across decarbonisation pathways. As the scale of district heating remains limited at this stage, the substitutional relationships between district heating and other building metrics are thus not obvious. Nonetheless, the decarbonisation of the building sector can already have evident influences on the power sector.

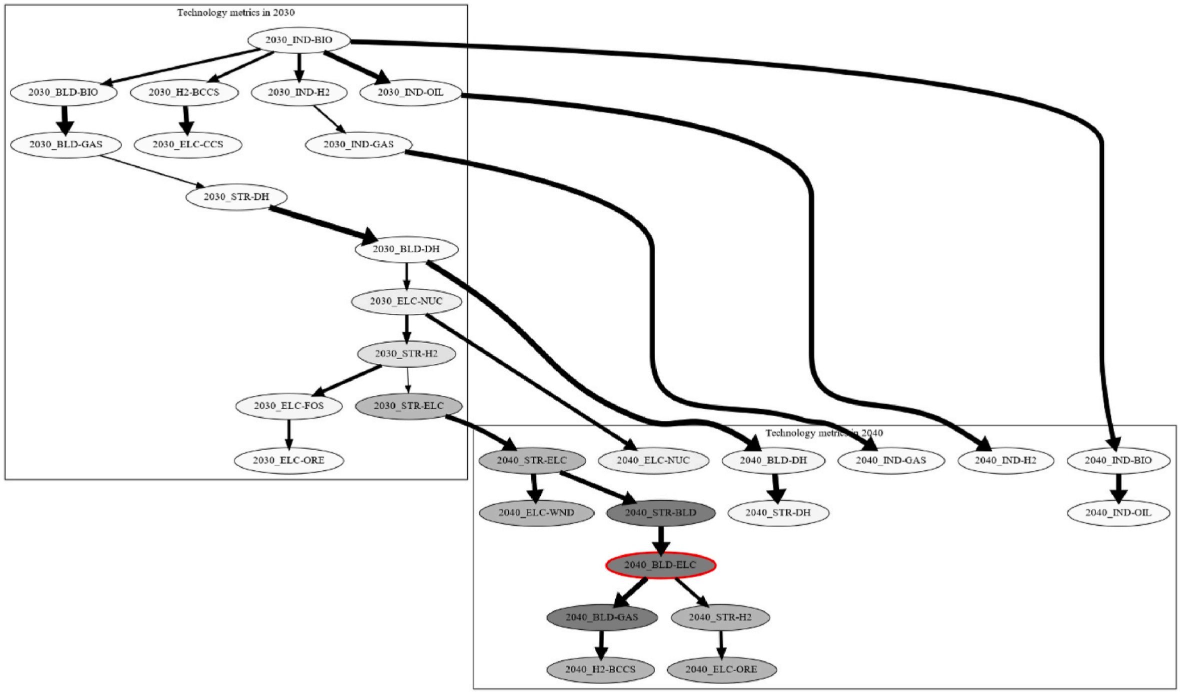


Fig. G. 1. Electrification of the building sector in 2040 (2040_BLD-ELC) is influential to many highly uncertain metrics in the same year (greyscale of nodes indicates the reduction in uncertainty of individual metrics once 2040_BLD-ELC is known)

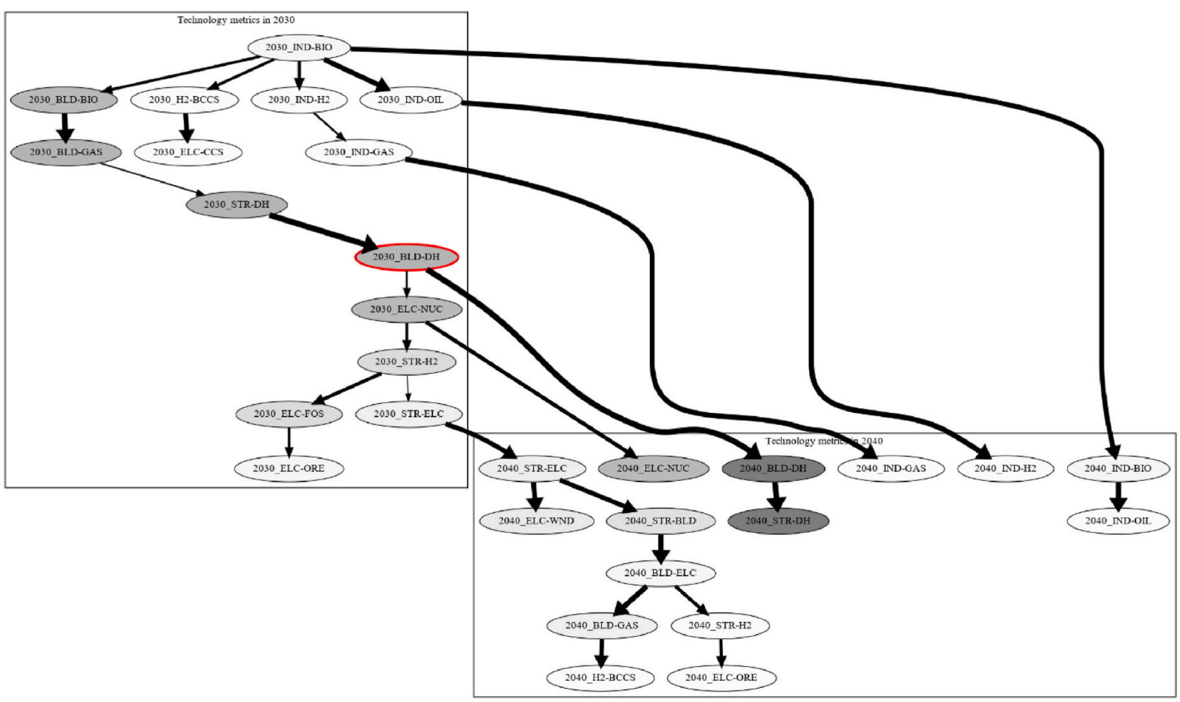


Fig. G. 2. District heating in 2030 (2030_BLD-DH) is influential to the uncertainty of some metrics in 2040 (greyscale of nodes indicates the reduction in uncertainty of individual metrics once 2030_BLD-DH is known)

Even though building metrics are still more influential than the most of other metrics to technology portfolios in 2050, there are some clear differences in the later period (2040–2050). Unlike the previous period, some power metrics (i.e. ELC-CCS and ELC-NUC) become the most influential in the whole system. In addition, more metrics (e.g. BLD-GAS) in the previous year (2040) have higher influences on the technology developments in 2050.

The high influence of two power metrics (ELC-CCS and ELC-NUC) is due to their high variance across pathways, as shown in Fig. 2. Therefore, even though the influences of these two metrics on other metrics are limited, as illustrated in Figure G. 3, the overall impacts of these two metrics are still significant. Especially, these two metrics have a strong substitutional relationship. Wind power (ELC-WND) in 2050 also has significant impacts due to its high variance and strong correlations with metrics in the previous year (2040). In turn, the changes to wind power in 2050 can thus also affect the deployment of other power technologies, such as 2050_ELC-ORE.

As for the significant influences of building technologies, those are due to their intrinsic high variances and strong cross-year correlations with building metrics in the previous year, as shown in Figure G. 4. The observation of the district heating metric in 2050 (2050_BLD-DH) can not only determine the state of the highly uncertain metric but can also further reduce the uncertainty of other metrics in 2050 via the strong correlations with district heating (2040_BLD-DH) and gas consumption (2040_BLD-GAS) in the building sector in 2040. Consequently, the reduction in uncertainty of the power sector (e.g. ELC-NUC and ELC-WND), transport sector (e.g. TCAR-ELC), and hydrogen production (e.g. H2-BCCS) can be observed, as indicated in Figure G. 4.

Moreover, metrics in the previous year have also been found influential, such as district heating (2040_BLD-DH) and gas consumption (2040_BLD-GAS) in the building sector. The influences of these metrics propagate through a similar route as the one just discussed above. Since the 2040_BLD-DH has more direct impacts on the district heating in 2050, the importance of 2040_BLD-DH is thus higher than 2040_BLD-GAS.

Finally, transport metrics (e.g. 2050_TCAR-ELC) become more influential in this period than in the previous period. This is majorly due to the high variance of vehicle-related metrics, including electric vehicles (TCAR-ELC) and oil vehicles (TCAR-OIL), and the strong correlations between them in this period. The stringent GHG budgets force the transport sector should be deeply decarbonised by either one of two decarbonisation strategies (i.e. electrification and oil vehicles with BECCS). The substitutional relationship has thus been significantly amplified.

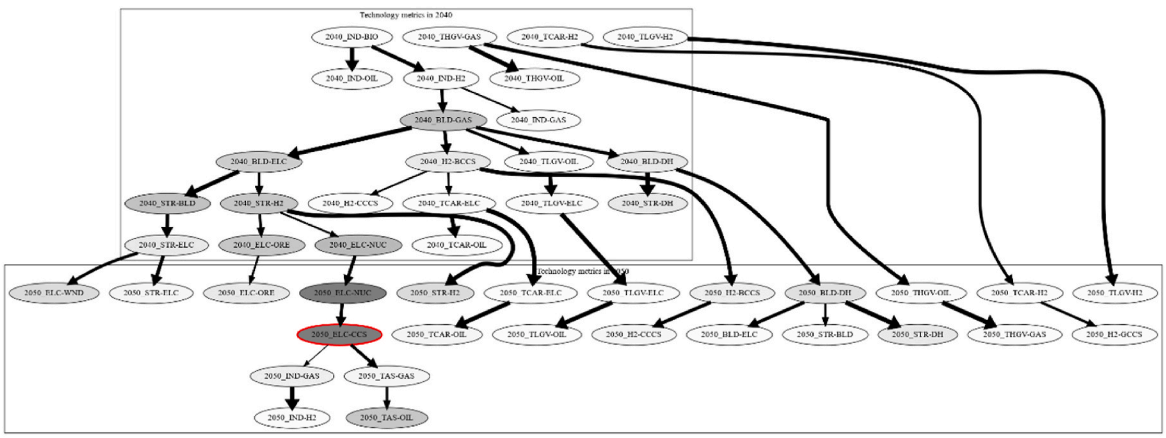


Fig. G. 3. Power metrics in 2050 are extremely influential as they have the highest uncertainties across decarbonisation pathways (greyscale of nodes indicates the reduction in uncertainty of individual metrics once 205_ELC-CCS is known)

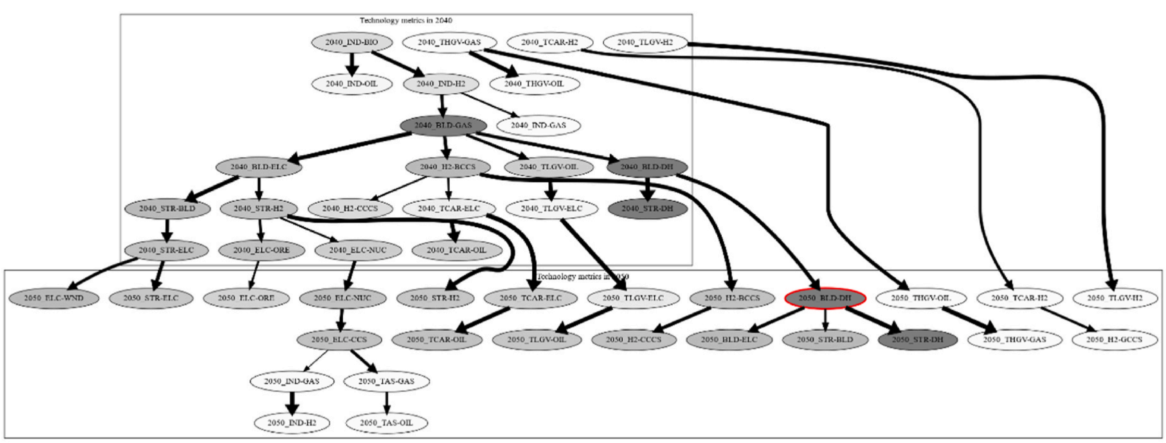


Fig. G. 4. District heating in 2050 is influential to the 2050 technology portfolios via its significant correlations with technology metrics in 2040 (greyscale of nodes indicates the reduction in uncertainty of individual metrics once 205_BLD-DH is known)

Hence, at the later stage, the developments of different technologies are more intertwined across sectors and years. More attention should be paid to the deployment of different power technologies due to the limited budgets available to the power sector. Moreover, decarbonisation measures in the building sector also have significant influences on technology portfolios in the transport sector. For instance, higher district heating might lead to less BECCS for hydrogen production. In turn, higher electrification level of passenger cars should be achieved to reduce GHG emissions from the transport sector by 2050. Furthermore, the increase in the electrification of the building sector requires more wind power to be deployed to balance the electricity supply and demand.

Overall, BNs can explicitly indicate the most influential technologies to technology portfolios across decarbonisation pathways. The correlations between deployments of technological metrics can be further revealed with the topology of BNs that can give policymakers clear ideas about which technologies should be paid more attention to and how to substitute and supplement technologies.

References

AGrUM/pyAgrum Team, 2023. Introduction to pyAgrum. https://pyagrum.readthedocs.io/en/latest/index.html. (Accessed 28 July 2023).

Barthelmes, Verena M., Heo, Yeonsook, Fabi, Valentina, Corgnati, Stefano P., 2017. Exploration of the Bayesian Network framework for modelling window control behaviour. Build. Environ. 126, 318–330.

Bassamzadeh, Nastaran, Ghanem, Roger, 2017. Multiscale stochastic prediction of electricity demand in smart grids using Bayesian networks. Appl. Energy 193, 369–380.

Basu, Kaustav, Hawarah, Lamis, Arghira, Nicoleta, Hussein, Joumaa, Ploix, Stephane, 2013. A prediction system for home appliance usage. Energy Build. 67, 668–679.

Bishop, Christopher M., Nasrabadi, Nasser M., 2006. Pattern Recognition and Machine Learning. Springer.

Carta, José A., Velázquez, Sergio, Matías, J.M., 2011. Use of Bayesian networks classifiers for long-term mean wind turbine energy output estimation at a potential wind energy conversion site. Energy Convers. Manag. 52, 1137–1149.

Chang, Yuanjiang, Zhang, Changshuai, Shi, Jihao, Li, Jiayi, Zhang, Shenyan, Chen, Guoming, 2019. Dynamic Bayesian network based approach for risk analysis of hydrogen generation unit leakage. Int. J. Hydrogen Energy 44, 26665–26678.

Chickering, Max, Heckerman, David, Meek, Chris, 2004. Large-sample learning of Bayesian networks is NP-hard. J. Mach. Learn. Res. 5, 1287–1330.

Csereklyei, Zsuzsanna, Thurner, Paul W., Langer, Johannes, Küchenhoff, Helmut, 2017. Energy paths in the European Union: a model-based clustering approach. Energy Econ. 65, 442–457.

Cui, Jitong, Kong, Yingying, Liu, Cuiwei, Cai, Baoping, Khan, Faisal, Li, Yuxing, 2024. Failure probability analysis of hydrogen doped pipelines based on the Bayesian network. Eng. Fail. Anal. 156, 107806.

Dao, Uyen, Sajid, Zaman, Khan, Faisal, Zhang, Yahui, 2024. Safety analysis of blended hydrogen pipelines using dynamic object-oriented bayesian network. Int. J. Hydrogen Energy 52, 841–856.

Düspohl, Meike, Döll, Petra, 2016. Causal networks and scenarios: participatory strategy development for promoting renewable electricity generation. J. Clean. Prod. 121, 218–230.

Fais, Birgit, Keppo, Ilkka, Zeyringer, Marianne, Usher, Will, Daly, Hannah, 2016. Impact of technology uncertainty on future low-carbon pathways in the UK. Energy Strategy Rev. 13, 154–168.

Friedman, Jerome H., Fisher, Nicholas I., 1999. Bump hunting in high-dimensional data. Stat. Comput. 9, 123–143.

Gordon, A.D., Breiman, L., Friedman, J.H., Olshen, R.A., Stone, Cj J., 1984. Classification and regression trees. Biometrics 40, 874.

Gross, Robert, Hanna, Richard, 2019. Path dependency in provision of domestic heating. Nat. Energy 4, 358–364.

Guivarch, Céline, Le Gallic, Thomas, Bauer, Nico, Fragkos, Panagiotis, Huppmann, Daniel, Jaxa-Rozen, Marc, Keppo, Ilkka, Kriegler, Elmar, Krisztin, Tamás, Marangoni, Giacomo, 2022. Using large ensembles of climate change mitigation scenarios for robust insights. Nat. Clim. Change 12, 428–435.

Guivarch, Céline, Rozenberg, Julie, Schweizer, Vanessa, 2016. The diversity of socio-economic pathways and CO2 emissions scenarios: insights from the investigation of a scenarios database. Environ. Model. Software 80, 336–353.

Hawarah, Lamis, Ploix, Stéphane, Jacomino, Mireille, 2010. User behavior prediction in energy consumption in housing using Bayesian networks. In: Artificial Intelligence and Soft Computing: 10th International Conference, ICAISC 2010, Zakopane, Poland, June 13-17, 2010, Part I 10. Springer, pp. 372–379.

Heaton, C., 2014. Modelling low-carbon energy system designs with the ETI ESME model. Energy Technologies Institute 1–28.

Herran, Diego Silva, Tachiiri, Kaoru, Matsumoto, Ken'ichi, 2019. Global energy system transformations in mitigation scenarios considering climate uncertainties. Appl. Energy 243, 119–131.

Højsgaard, Søren, 2012. Graphical independence networks with the gRain package for R. J. Stat. Software 46, 1–26.

Hosseini, Seyedmohsen, Sarder, M.D., 2019. Development of a Bayesian network model for optimal site selection of electric vehicle charging station. Int. J. Electr. Power Energy Syst. 105, 110–122.

Howells, Mark, Rogner, Holger, Strachan, Neil, Heaps, Charles, Huntington, Hillard, Kypreos, Socrates, Hughes, Alison, Silveira, Semida, DeCarolis, Joe, Morgan, Bazillian, 2011. OSeMOSYS: the open source energy modeling system: an introduction to its ethos, structure and development. Energy Pol. 39, 5850–5870.

Hu, Min, Chen, Huanxin, Shen, Limei, Li, Guannan, Guo, Yabin, Li, Haorong, Li, Jiong, Hu, Wenju, 2018. A machine learning bayesian network for refrigerant charge faults of variable refrigerant flow air conditioning system. Energy Build. 158, 668–676.

Hughes, Nick, Strachan, Neil, Gross, Robert, 2013. The structure of uncertainty in future low carbon pathways. Energy Pol. 52, 45–54.

Keppo, Ilkka, van der Zwaan, Bob, 2012. The impact of uncertainty in climate targets and CO 2 storage availability on long-term emissions abatement. Environ. Model. Assess. 17, 177–191.

Kwakkel, Jan H., 2017. The Exploratory Modeling Workbench: an open source toolkit for exploratory modeling, scenario discovery, and (multi-objective) robust decision making. Environ. Model. Software 96, 239–250.

Li, Francis GN., Evelina, Trutnevyte, 2017. Investment appraisal of cost-optimal and near-optimal pathways for the UK electricity sector transition to 2050. Appl. Energy 189, 89–109.

Li, Pei-Hao, Keppo, Ilkka, Strachan, Neil, 2018. Incorporating homeowners' preferences of heating technologies in the UK TIMES model. Energy 148, 716–727.

Li, Pei-Hao, Pye, Steve, Keppo, Ilkka, 2020. Using clustering algorithms to characterise uncertain long-term decarbonisation pathways. Appl. Energy 268, 114947.

Li, Pei-Hao, Pye, Steve, Keppo, Ilkka, Jaxa-Rozen, Marc, Trutnevyte, Evelina, 2023. Revealing effective regional decarbonisation measures to limit global temperature increase in uncertain transition scenarios with machine learning techniques. Climatic Change 176, 80.

Loulou, Richard, Remme, Uwe, Kanudia, Amit, Lehtila, Antti, Goldstein, Gary, 2005. Documentation for the times model part ii. Energy Technology Systems Analysis Programme.

Ma, Tai-Yu, Chow, Joseph YJ., Jia, Xu, 2017. Causal structure learning for travel mode choice using structural restrictions and model averaging algorithm. Transportmetrica: Transport. Sci. 13, 299–325.

Machado, Pedro Gerber, Ribeiro, Celma de Oliveira, Oller do Nascimento, Claudio Augusto, 2023. Risk analysis in energy projects using Bayesian networks: a systematic review. Energy Strategy Rev. 47, 101097.

Marcot, Bruce G., 2012. Metrics for evaluating performance and uncertainty of Bayesian network models. Ecol. Model. 230, 50–62.

Marcot, Bruce G., Douglas Steventon, J., Sutherland, Glenn D., McCann, Robert K., 2006. Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. Can. J. For. Res. 36, 3063–3074.

Masson-Delmotte, Valérie, 2018. Global Warming of 1.5° C: an IPCC Special Report on Impacts of Global Warming of 1.5° C above Pre-industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Contex of Strengthening the Global Response to the Thereat of Blimate Change, Sustainable Development, and Efforts to Eradicate Poverty *(No Title)*.

Moallemi, Enayat A., Fjalar, de Haan, Kwakkel, Jan, Lu, Aye, 2017. Narrative-informed exploratory analysis of energy transition pathways: a case study of India's electricity sector. Energy Pol. 110, 271–287.

Moglianesi, Andrea, Keppo, Ilkka, Lerede, Daniele, Savoldi, Laura, 2023. Role of technology learning in the decarbonization of the iron and steel sector: an energy system approach using a global-scale optimization model. Energy 274, 127339.

Moksnes, Nandi, Rozenberg, Julie, Broad, Oliver, Taliotis, Constantinos, Howells, Mark, Rogner, Holger, 2019. Determinants of energy futures—a scenario discovery method applied to cost and carbon emission futures for South American electricity infrastructure. Environmental Research Communications 1, 025001.

Morris, Peter, Vine, Desley, Buys, Laurie, 2015. Application of a Bayesian Network complex system model to a successful community electricity demand reduction program. Energy 84, 63–74.

Murphy, Kevin P., 2012. Machine Learning: a Probabilistic Perspective. MIT press.

Pizarro-Alonso, Amalia, Ravn, Hans, Münster, Marie, 2019. Uncertainties towards a fossil-free system with high integration of wind energy in long-term planning. Appl. Energy 253, 113528.

Price, James, Keppo, Ilkka, 2017. Modelling to generate alternatives: a technique to explore uncertainty in energy-environment-economy models. Appl. Energy 195, 356–369.

Pye, Steve, Li, Pei-Hao, Keppo, Ilkka, O'Gallachoir, Brian, 2019. Technology interdependency in the United Kingdom's low carbon energy transition. Energy Strategy Rev. 24, 314–330.

Pye, Steve, Sabio, Nagore, Strachan, Neil, 2015. An integrated systematic analysis of uncertainties in UK energy transition pathways. Energy Pol. 87, 673–684.

Quinn, Julianne D., Reed, Patrick M., Keller, Klaus, 2017. Direct policy search for robust multi-objective management of deeply uncertain socio-ecological tipping points. Environ. Model. Software 92, 125–141.

Riahi, Keywan, Van Vuuren, Detlef P., Kriegler, Elmar, Edmonds, Jae, O'neill, Brian C., Fujimori, Shinichiro, Bauer, Nico, Calvin, Katherine, Dellink, Rob, Fricko, Oliver, 2017. The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: an overview. Global Environ. Change 42, 153–168.

Scutari, Marco, 2009. Learning Bayesian networks with the bnlearn R package. arXiv preprint arXiv:0908.3817.

Small, Mitchell J., Wong-Parodi, Gabrielle, Kefford, Benjamin M., Stringer, Martin, Schmeda-Lopez, Diego R., Greig, Chris, Ballinger, Benjamin, Wilson, Stephen, Smart, Simon, 2019. Generating linked technology-socioeconomic scenarios for emerging energy transitions. Appl. Energy 239, 1402–1423.

Song, Guobao, Semakula, Henry Musoke, Fullana-i-Palmer, Pere, 2018. Chinese household food waste and its' climatic burden driven by urbanization: a Bayesian Belief Network modelling for reduction possibilities in the context of global efforts. J. Clean. Prod. 202, 916–924.

Unruh, Gregory C., 2000. Understanding carbon lock-in. Energy Pol. 28, 817–830.

Xie, Chi, Waller, S., 2010. Estimation and application of a Bayesian network model for discrete travel choice analysis. Transportation Letters 2, 125–144.

Yue, Xiufeng, Pye, Steve, DeCarolis, Joseph, Li, Francis GN., Rogan, Fionn, Gallachóir, Brian Ó., 2018. A review of approaches to uncertainty assessment in energy system optimization models. Energy Strategy Rev. 21, 204–217.