A Machine Learning Method to Extract Key Policy Decisions from Energy Transition Scenarios under Uncertainty

Stefano Moret,a,\* Florian Joseph Baader,a Wolfram Wiesemann,b Iain Staffellc and André Bardowa

aEnergy & Process Systems Engineering, Department of Mechanical and Process Engineering, ETH Zürich, 8092 Zürich, Switzerland

bCentrefor Environmental Policy, Imperial College London, London, UK

cImperial College Business School, Imperial College London, London, UK

\*morets@ethz.ch

Abstract

Uncertainties in input parameters, such as fuel prices and energy demands, often lead energy modelers to present large sets of scenarios that are difficult to interpret. Leveraging decision trees, a popular machine-learning technique, we translate the complexity of energy transition studies under uncertainty into a small set of key decisions. Application of our method to an Integrated Assessment Model under uncertainty shows that the global energy transition is primarily determined by choices on heating and industry electrification. Our replicable framework reduces the vast amount of plausible energy system scenarios to a few interpretable storylines and unveils the most important trade-offs in the energy transition.

**Keywords**: energy planning, uncertainty, machine learning, decision-making.

* 1. Introduction

Despite the stark reality of climate change as well as the evidence that immediate action is beneficial over a “wait-and-see” strategy, the Intergovernmental Panel on Climate Change’s (IPCC) latest report warns that “the pace and scale of what has been done so far, and current plans [to tackle climate change], are insufficient” (IPCC, 2023). This indecisiveness in policy deployment can, at least in part, be attributed to the extreme uncertainties affecting long-term energy planning and decision-making.

Energy system optimization models (ESOMs) can help unravel this complexity and assist policymakers in defining *quantitative* energy transition pathways. While, historically, ESOM studies have focused on generating a single optimal solution, evidence that models’ assumptions and input data – such as fuel prices, demands, interest rates, etc. – are highly uncertain (Moret et al., 2017) motivated modelers to consider uncertainty. However, uncertainty studies often generate hundreds of solutions that are challenging for decision-makers to interpret and act upon (Pickering et al., 2022).

We present a machine learning method to streamline hundreds or thousands of energy system scenarios from uncertainty studies to a few interpretable storylines. The storylines yield *qualitative* descriptions of energy system configurations and are described by urgent policy decisions. Specifically, we show that training decision trees on key outputs of interest of ESOMs allows translating many quantitative energy transition scenarios into a small number of storylines. These storylines are determined by a few policy decisions and, thus, are accessible and interpretable to a broader public.

* 1. Methods

First, different energy system scenarios are generated by solving a mathematical optimization problem under uncertainty. Given a probability distribution for each uncertain parameter $θ$, sampling – often Monte Carlo sampling – is performed to obtain *N* possible realizations of the uncertain parameter vector $θ$. The energy model is run for each combination of input parameters $θ\_{i }, i = \{1, ... , N\}$, resulting in *N* energy scenarios. Second, *k*-means clustering (Hastie et al., 2009) is performed to group these *N* different scenarios into *k* clusters (where $k\ll N$) with respect to *m* outputs of interest **y**, corresponding to high-level outputs needed to inform decision-makers. *k*-means clustering is performed using the Python package *sklearn.cluster.KMeans* with default settings, and the number of clusters *k* is selected using the “elbow method” (Ketchen and Shook, 1996).

Third, a decision tree (Hastie et al., 2009) is trained to identify the key decisions and their effects on the solution space. Decision trees are a supervised machine learning method and, hence, are trained on labeled data. In our case, the labeled data are the *N* energy system scenarios $y\_{i }$ and the labels $c$ are the cluster numbers assigned by the previous step. The decision tree learns to predict the label $c\_{i }$ given the energy system scenario $y\_{i }$. We limit the number of leaves of the tree to the number of clusters *k* as, in our numerical experiments, a tree with a number of leaves equal to the number of clusters has high prediction accuracy, i.e., 99% of the data points are correctly assigned to their clusters. Finally, we re-order the energy system scenarios to *k* clusters according to the rules of the decision tree, i.e., the few points that are predicted wrongly by the decision tree are re-assigned to the other clusters accordingly, making the clustering more interpretable.

* 1. Results

We apply our method to the global decarbonization pathways under uncertainty studied by Panos et al. (2023), who recently presented the first Monte-Carlo assessment of an integrated assessment model (IAM). Specifically, their study presents 1000 possible global decarbonization pathways generated by sampling 18 uncertain parameters. By applying our method to their published dataset, these 1000 scenarios can be summarized by two key policy decisions: (i) electrification of heating and (ii) electrification of industry (Figure 1). The resulting decision tree has only $k=3$ leaves, corresponding to three storylines for the global energy transition, illustrating the consequences of each decision. At the root node lies the entire decision space with 1000 scenarios, as presented by Panos et al. (2023), with the radar plots and the open locks indicating the large ranges of variations in the results.

The first decision differentiates between scenarios with low and high shares of electric heating and thereby splits the set of solutions: A low share of electric heating is found to also imply low non-fossil transport and minimal electrification of industry. Moreover, the deployment of renewables is low. On the other side of the tree with a high share of electric heating, the share of non-fossil transport must also be high. As the next key decision, the tree further differentiates between solutions with a low and a high electrification of industry. A high electrification of industry then implies a high deployment of renewables and maximum electrification of heating. The extent to which the resulting storylines rely on renewable energy and sector coupling increases from left to right in Figure 1. Overall, our method breaks down the 1000 Monte Carlo results into a decision- tree with only two key decisions, translating the quantitative output of the IAM study into three storylines corresponding to three actionable policies.



**Figure 1**: Decision tree translating 1000 scenarios for the global energy transition in 2100 from (Panos et al., 2023) into two key decisions (electrification of heating and industry) along five outputs of interest. The axes of the radar plots are normalized on the range of each output of interest, while the locks indicate the level of flexibility. The volume of the decision space at each node of the tree is expressed by Σ.

* 1. Conclusions

We present a machine learning method to extract key policy decisions from energy transition scenarios under uncertainty. Application to the global energy transition using an Integrated Assessment Model demonstrates that our method reduces the vast decision space to a small number of interpretable storylines and critical decisions that must be made to enable this transition. Additionally, we propose a new way to visualize these choices into decision trees, effectively prioritizing decisions and associating each choice with its implied consequences. This analysis unveils the most important interconnection and trade-offs between key policy decisions. These trade-offs (which can be thought of as pivotal policy decisions) are typically hidden both to the user and even to the creators of scenarios, as the volumes of results and high-dimensional solution space obscure ‘the wood for the trees’. By combining the strengths of energy systems models and their outputs with the simplicity and clarity of storylines, our tool empowers decision-makers to quickly uncover actionable insights.

**Acknowledgments**

S.M. acknowledges support from the Swiss National Science Foundation under Grant No. PZ00P2\_202117. F.B. and A.B. acknowledge funding by the Swiss Federal Office of Energy’s SWEET program as part of the project PATHFNDR.

References

T. Hastie, R. Tibshirani, J. Friedman. The elements of statistical learning: Data mining, inference, and prediction. Springer, 2009

Intergovernmental Panel on Climate Change (IPCC). AR6 Synthesis Report: Climate Change 2023. Technical report, Intergovernmental Panel on Climate Change (IPCC), 2023.

D. J. Ketchen, C. L. Shook. The application of cluster analysis in strategic management research: An analysis and critique. Strategic management journal 17, 441–458, 1996.

S. Moret, V. Codina Gironès, M. Bierlaire, F. Maréchal. Characterization of input uncertainties in strategic energy planning models. Applied Energy 202, 597–617, 2017.

E. Panos, J. Glynn, S. Kypreos, A. Lehtilä, X. Yue, B. Ó Gallachóir, D. Daniels, H. Dai. Deep decarbonisation pathways of the energy system in times of unprecedented uncertainty in the energy sector*.* Energy Policy 180, 113642, 2023.

B. Pickering, F. Lombardi, S. Pfenninger. Diversity of options to eliminate fossil fuels and reach carbon neutrality across the entire European energy system. Joule 6,1253–1276, 2022.