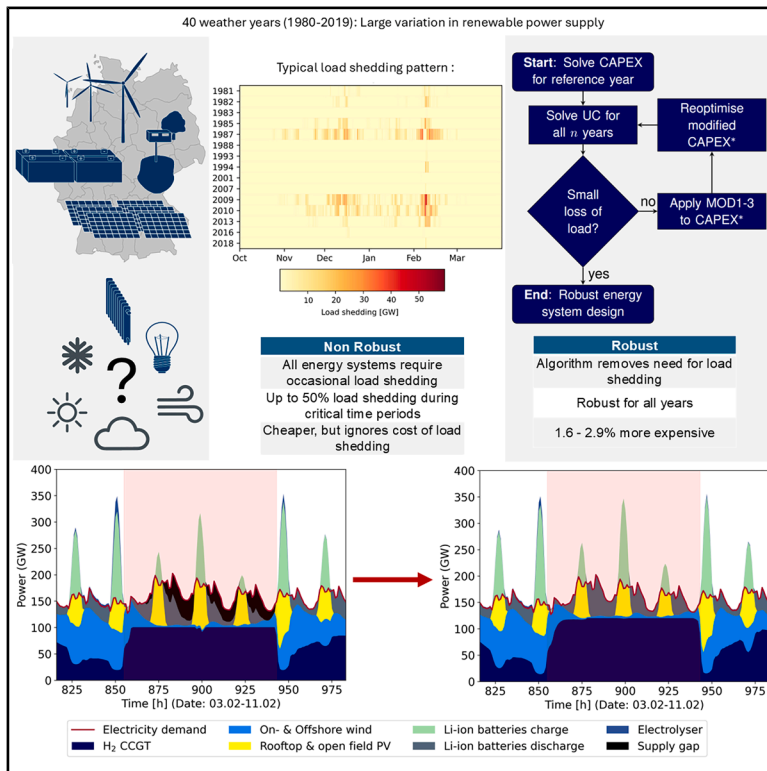


Robust capacity expansion modeling for renewable energy systems

Graphical abstract



Authors

Sebastian Kebrich, Felix Engelhardt, David Franzmann, Christina Büsing, Jochen Linßen, Heidi Heinrichs

Correspondence

s.kebrich@fz-juelich.de (S.K.), engelhardt@combi.rwth-aachen.de (F.E.)

In brief

Applied sciences

Highlights

- In energy system modeling, disregarding weather variability leads to misinvestment
- Average weather years fail to capture intra-/inter-year variability
- Best practice: Use average years, but validate and iteratively adjust solutions
- Protection against weather variability requires only 2%–3% additional annual costs



Article

Robust capacity expansion modeling for renewable energy systems

Sebastian Kebrich,^{1,4,5,*} Felix Engelhardt,^{2,4,*} David Franzmann,¹ Christina Büsing,² Jochen Linßen,¹ and Heidi Heinrichs^{1,3}

¹Forschungszentrum Jülich GmbH, Institute of Climate and Energy Research, Jülich Systems Analysis (ICE-2), Jülich, Germany

²RWTH Aachen University, Faculty of Computer Science, Teaching and Research Area Combinatorial Optimization, Aachen, Germany

³University of Siegen, Department of Mechanical Engineering, Chair for Energy Systems Analysis, Siegen, Germany

⁴These authors contributed equally

⁵Lead contact

*Correspondence: s.kebrich@fz-juelich.de (S.K.), engelhardt@combi.rwth-aachen.de (F.E.)

<https://doi.org/10.1016/j.isci.2026.114929>

SUMMARY

Future greenhouse gas neutral energy systems will be dominated by renewable energy technologies providing variable supply subject to uncertain weather conditions. For this setting, we propose an algorithm for capacity expansion planning: We evaluate solutions optimized on a single years' data under different input weather years, and iteratively modify solutions whenever supply gaps are detected. These modifications lead to solutions with sufficient capacities to overcome periods of cold dark lulls and seasonal demand/supply fluctuations. A computational study on a German energy system model for 40 operating years shows that preventing supply gaps, i.e., finding a robust system, increases the total annual cost by 1.6-2.9%. In comparison, non-robust systems display loss of load close to 50% of total demand during some periods. Results underline the importance of assessing the feasibility of energy system models using atypical time-series, combining dark lull and cold period effects.

INTRODUCTION

Modeling energy systems require access to different types of data. For future greenhouse gas neutral energy systems, time-series data are of special importance, as a vast expansion of wind power and *solar photovoltaic (PV)* that are dependent on the weather is perceived as indispensable.^{1,2} However, we cannot expect to have accurate long-term information on future weather conditions. Instead, historical time-series serve as a substitute in capacity expansion planning, and often a single “representative” sample year is selected for optimization.³ For example, the *International Renewable Energy Agency (IRENA)* recommends using 2018 as a reference year because it represents generation from renewable technologies well on average.⁴ In the following, we argue that using “representative” sample years may lead to *energy system models (ESMs)* that appear sound at first glance, but would fail to meet supply in reality. We then propose a tractable approach to counteract this effect and to make energy systems robust against uncertain conditions during operation, i.e., ensure security of supply.

There is ample evidence in literature that the results of energy system optimization are sensitive to changes in weather time-series data. Schyska et al. (2021) evaluated the sensitivity of capacity expansion models with regards to multiple sources of uncertainty. They conclude that some years are unsuited as reference years, as using them for optimization leads to significant misallocation of assets.⁵ Ruggles et al. (2024) assessed how many

years of weather data are needed to ensure ESMs are reliable even out of sample, i.e., if a different weather year were to realize. They conclude that between 15 and 40 years are required depending on amount of imports available/the acceptable level of loss of load.⁶ The works by Haddeland et al.,⁷ and Staffell et al.,⁸ which builds on earlier work by Pfenninger et al.,⁹ similarly find the choice of weather years significantly effects renewable generation and thus power output. Not only that, but also the effect of the weather increases with increasing share of renewable technologies.¹⁰ De Marco et al.¹¹ identified energy shortages across Europe and use those to optimize climate-resilient energy systems stochastically. Not all research agrees on this; Schlachtenberger et al.¹² optimized three weather years with hourly data both individually and as one time-series with a resolution of 3 h per time step, finding only small variations in *total annual cost (TAC)* and installed capacities. However, they note that aggregating multiple hours together introduces a smoothing effect that systematically favors PV and underestimate battery and wind generation requirements.

Additionally, there is a fundamental tension between identifying typical years, and build ESMs that are protected against extreme events, i.e., robust. On this, Hilbers et al.¹³ introduced a method of importance subsampling for time-series aggregation to explicitly preserve extreme events in the weather data as an alternative to established “representative days” clustering approaches. Ryberg's dissertation¹⁴ and Ryberg et al.¹⁵ investigated the impact of generation lulls in an energy system for a large part of Europe



Table 1. Investment problem: Chose number of solar power plants to build on locations A, B, C

Scenario	Location A	Location B	Location C
u_1	200	75	50
u_2	50	75	150

Values indicate expected energy supply in $\frac{\text{GWh}}{\text{time period}}$ per investment.

calculating backup capacities required to overcome these. Ruhau et al.¹⁶ looked into the storage requirements for a renewable-based ESM for Germany using 35 years of weather data taking consecutive extreme events into account. They concluded that consecutive extreme events increase storage requirements significantly compared to even the most extreme, but singular events. Thus, atypical weather time-series may be particularly well-suited for optimization because they capture important system behaviors, e.g., dark lulls, that significantly impact ESMs. In this context, Grochowicz et al.¹⁷ discussed optimizing sequential weather years. They use a geometry-based solution approach targeting the solution space. In follow-up work, they use electricity shadow prices to identify difficult weather periods.¹⁸ They observe that such difficult weather periods are not just meteorological events, but results of the interplay of meteorology and electricity storage and network structures.

In stochastic optimization, the notion of using typical, i.e., expected, behavior as a baseline for optimization is a well-known concept. Its usefulness is determined by the value of a stochastic solution, which represents the gap in expected performance of a solution obtained with expected data and one determined by solving a full stochastic optimization problem.¹⁹ In general, this gap may be arbitrarily large. However, research has also shown that we can still extract insights from solutions obtained with such data.²⁰ Specifically concerning weather robustness, Forghani et al. (2025) proposed an intermediate approach where not only a representative but also worst-/best-cost years are used as input data for a stochastic optimization. Their results show that this reduces loss-of-load to practically zero, at < 1% additional cost.²¹

An important modeling decision in this context is the value and availability of recourse, i.e., the ability to specify or change parts of the solution if uncertainty realizes. In capacity expansion planning, imports constitute such recourse. If arbitrary imports are allowed, any misplanning can be compensated for, if at a cost. We focus on a different setting, the one where imports are limited or energy systems are to be self-sufficient. Here, we use *adaptive robust optimization (ARO)* to model limited recourse while still enforcing strict guarantees on security of supply general solution approaches. To the best of our knowledge, Zeyringer et al.²² were the first to assess the effects of weather uncertainty in input data on capacity-expansion planning, looking at an ESM of Grein. They proposed using multiple historical time series, optimizing over each, and then evaluating the operational costs/supply costs incurred by the proposed energy systems under the assumption that the time-series of a different historical year realizes. Then, an ESM with lowest worst-case costs/lowest supply gap among the solutions is selected. They find that starting with the “wrong” reference year may lead to misallocation of re-

sources and that modeling with multiple years leads to more consistent results with lower worst-case cost, the latter of which had also been noted before.²³ For the United States, Dowling et al.²⁴ make the case for multi-year modeling to accurately capture long-term storage effects, noting that the cost of variable renewable power systems are especially sensitive to long-duration storage costs. A recent work by Gøtske et al.²⁵ also assessed energy systems based on different weather years. They employed CO₂ emitting backup technologies and analyzed structural elements of the respective solutions. Either approach allows to select more suitable ESMs, but it cannot assure solutions meet certain supply/demand across all years.

The two-stage robust setting

To illustrate the setting considered in this work, we begin with an example for an ARO ESM. ARO, alternatively called two-stage robustness, is characterized through a bilevel structure²⁶ in which a decision maker has to make a set of first stage decisions $x \in \mathcal{X}$, then one of multiple scenarios $u \in \mathcal{U}$ may realize and afterward the decision maker has the option to react to scenario $u \in \mathcal{U}$ with a scenario-dependent set of decisions $y(u) \in \mathcal{Y}(x, u)$. Equation (ARO) illustrates the structure of an ARO problem. Here, c and d are vectors of first and second stage costs, respectively. The matrix A encodes all information that is not affected by uncertainty, whereas the matrices $B(u)$ and the right-hand side vector $b(u)$ are dependent on the scenario u .

$$\begin{aligned} \min_{x \in \mathcal{X}} c^T x + \max_{u \in \mathcal{U}} \min_{y(u) \in \mathcal{Y}(x, u)} d^T y(u) \\ \text{s.t.} \quad Ax + B(u)y(u) \geq b(u) \quad \forall u \in \mathcal{U} \end{aligned} \quad (\text{ARO})$$

For the ESM(s) considered in this work, the first stage decisions correspond to the capacity expansion planning, e.g., investment in solar power plants, power lines, storage units, and the second stage corresponds to operational decisions, e.g., how much energy is to be stored in which battery at what time. If energy imports are allowed, these will also be encoded in \mathcal{Y} . The scenario u corresponds to a specific, continuous time period. Table 1 illustrates this setting. Here, we consider different locations A, B, C for a solar power plant, whose production is subject to changing weather patterns during the time periods u_1, u_2 .

The optimal solution depends on what type of recourse is available and how much supply is needed. For example, if we assume similar cost at all locations and no/expensive imports, investing in location B is the most cost effective for meeting a demand of $75 \frac{\text{GWh}}{\text{time period}}$. However, if the total demand were $200 \frac{\text{GWh}}{\text{time period}}$, it is more effective to invest in A and C instead of B.

Returning to our example in Table 1, the approach by Zeyringer et al. would not be able to find the optimal robust solution for a total demand of 75 MWh/year, which is to invest in location B. This is because if the future is known, i.e., we assume either u_1 or u_2 to realize, the optimal investment is always to either invest in location A or location B. An alternative approach was put forward by Caglayan et al. (2019),²⁷ who noted variations of up to 20% variations in TAC, and significant differences in energy system designs when optimizing an integrated electrical/hydrogen ESM for Western Europe with different weather years. They propose an iterative approach, where they begin with

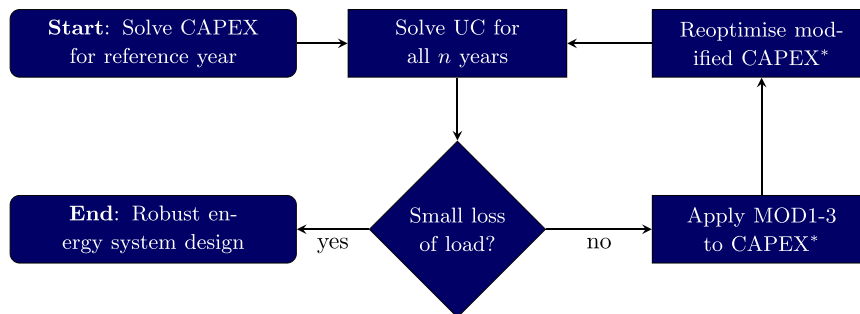


Figure 1. Flowchart depicting the proposed methodology for determining robust energy systems

In each main loop iteration, $n - 1$ unit commitment problems (UC) and one modified capacity expansion problem (CAPEX*) are solved until loss of load is sufficiently small.

optimizing individual years, but then average their designs and take the average as a lower bound for installed capacities in a new round of optimization.²⁸ Since the (average) capacities installed during sequential algorithm iterations will be non-decreasing and bounded, convergence and thus a feasible solution will be guaranteed in most cases. In our example, the algorithm would enforce capacity > 0 for location A and C, which would eventually lead to a feasible but unnecessarily expensive solution that uses both A and C.

A novel algorithm for computing robust ESMs

The proposed algorithm consists of four main steps, as shown in Figure 1. As its input, the algorithm takes a set of n times series, one of which is designated as a starting/reference year. We begin with solving the capacity expansion planning problem for this year (CAPEX), then we evaluate the performance of the proposed investment in all n years by solving the corresponding unit commitment problem (UC), which formalizes the validation proposed in previous works.^{22,25}

To ensure feasibility, existing approaches allow for energy imports. We consider the robust setting where imports are bounded/not available, i.e., the ESM needs to be self-sufficient. Based on the previous step, we identify supply gaps, which we denote by δ . The underlying ESM for Germany for 2045 utilizes renewable technologies only and includes PV, wind, Li-ion batteries, electrolyzer, hydrogen salt caverns, hydrogen combined cycle gas turbines as well as electricity transmission and hydrogen pipelines. The models are described in detail in the STAR Methods section for both *gurobipy* and the ETHOS.FINE. If the supply gaps are sufficiently small, we terminate. Otherwise, we use the information encoded in δ to iteratively make solutions more robust. For that, modifications (MOD) are applied to the optimization problem by either adding artificial demand MOD 1/Demand increase, substituting parts of the time series data MOD 2/Synthetic time-series or iteratively adding both demands and lazy valid inequalities MOD 3/Combine, before reoptimizing CAPEX to evaluate the loss of load. Detailed definition of the modifications can be found in the methods section.

The main contributions of our work include not only insights derived for the specific ESM, and an algorithm for solving CAPEX problems to find solutions that are robust to variations in operational conditions across the n years, but foremost a systematic comparison of different approaches for designing robust ESMs. Based on this, we recommend three working modifications that all lead to robust energy systems.

RESULTS AND DISCUSSION

In the following, we first compare solutions derived from individual years' data and assess their (lack of) robustness. Then, we modify them to become more robust, and evaluate the features of robust solutions. Note that we provide results for two different models, since some modifications require custom callbacks/lazy constraints which are hard to integrate into modeling frameworks. Thus, we use one model directly coded in *gurobipy* that allows for all modifications but is simplified, see reference linear program, and a more realistic model implemented in the ETHOS.FINE framework. Table 2 provides an overview over the technologies considered and Table 3 provides the techno-economic parameters used in this study.

Note that we also only use a single demand time-series from 2012 for comparison. This significantly simplifies evaluation of the algorithms as otherwise, loss of load and installed would need to be considered relative to each year's demands. However, in most practical settings where demands and weather are linked, e.g., when employing significant electric or hydrogen-based heating, the paired demand time-series ought to be used.

Optimal capacities strongly depend on yearly weather data

The TAC for energy systems for the 38 node model of Germany within ETHOS.FINE for the 40 different years of time-series data deviates around an average of 106bn€ with cost between 96.4bn€ and 113.6bn€ annually, which equals -9% to $+7\%$ compared to the default reference year 2018 recommended by IRENA.⁴ The results of optimizing each year independently are given by Figure 2. While the variations in overall TAC are limited, the energy system designs show substantial deviations. The cost shares of single technologies across the 40 different single years vary by 69% for hydrogen pipelines, 57% for hydrogen salt caverns, 53% for CCGT, 44% for Li-ion batteries, 40% for PV, 38% for electrolyzers, 22% for the electricity grid, and 20% for onshore wind, making it challenging to draw recommendations for planned capacity expansion for future energy systems.

In the simplified single node model, the cost variation for single technologies is on average slightly higher than in the 38-node model: onshore wind (45%), rooftop PV (57%), Li-ion batteries (51%), hydrogen salt caverns (54%), electrolyzers (54%), and CCGT (53%). Open field PV is installed up to its maximum capacity for all years, while offshore wind is never utilized. A plot of all individual years is given in Figure 3.

In summary, even small input data differences may lead to large variations in investment into different technologies. For example, offshore wind is not utilized at all in some years, but makes up over 13% of the TAC in 2014. Note that the single

Table 2. Energy system components considered for development of the proposed methodology for optimising ESMs for Germany

	Technology
Supply	Rooftop PV, Open field PV, Onshore wind, Offshore wind
Storage	Li-ion batteries, H ₂ salt caverns
Transmission	Electricity grid, H ₂ pipelines
Conversion	Electrolysers, H ₂ combined cycle gas turbines (CCGT)
Demand	Electricity demand

node model is not directly comparable to the ETHOS.FINE model, as it contains various simplifications. As such, solutions are significantly cheaper averaging 55.4bn€ with between 52.1 and 59.7bn€ annually, i.e., -6% to +7%.

Systems optimized on single years cause supply gaps

Feasibility testing shows that all 40 energy system designs solely based on one year's time-series lead to supply gaps in multiple other years. This means that none of the system designs are robust. However, we are not only interested in whether there is any loss of load but also in the magnitude of said loss of load. Figure 4 shows the amount of load shedding when testing the feasibility of one of five reference years used in ETHOS.FINE during the other 39 years.

Two trends can be observed: On the one hand, when the feasibility is tested for 1987 and 2010, load shedding occurs mostly in the first half of February. This is caused by demand spikes in early February due to the low temperatures in that time of the year, which the electricity demand was based on. At the same time, the energy systems optimized for 1987 and 2010 are characterized by the highest amounts of PV and

wind capacities installed while having insufficient backup capacities.

On the other hand, the energy system of the reference years 2009 has increased backup capacities, therefore suffering less during cold dark lull periods, but does not have enough PV and wind capacities to produce enough hydrogen and therefore has load shedding more evenly distributed over the months from October to March. The other two reference years 1996 and 2013 are in the middle of these two cases.

Figure 5 illustrates the reasons for that by showing three uncritical (top row) periods with sufficient supply as well as three critical (bottom row) time periods with significant supply gaps. During critical time periods electricity generation from PV and wind is low, providing less than 50% of electricity demand. The remaining electricity demand needs to be fulfilled by backup capacities, i.e., H₂ CCGT. The periods have varying duration and they are identified via clustering and feasibility testing. Uncritical time periods are characterized by high availability of PV or onshore wind or both, while critical ones are characterized by low availability of PV and low to negligible onshore wind combined with high demand. Offshore wind plays only a minor role due to its limited utilization. The most critical time period takes place in 1994 (see Figure 5F). Here, wind and solar supply indicate a dark lull. Combined with the cold period identified in the electricity data, this constitutes a cold dark lull.

For the *gurobipy* implementation, we get similar results: The minimal annual supply gap across all years is 1.170 GWh for 1987, which is about 0.13% of the total annual electricity demand. Additionally, during shorter time-periods the supply gap can reach up to half of the required electricity demand. In comparison, optimizing using 1990s weather data leads to a peak supply gap of 69.500 GWh, i.e., 8%, if the weather of 1987 were to realize.

Table 3. Techno-economic parameters considered in this work

Technology	CAPEX ₂₀₅₀	OPEX _{fix,2050}	Life time [a]	Source
PV (Rooftop)	474 $\frac{\text{€}}{\text{kW}}$	10 $\frac{\text{€}}{\text{kW a}}$	20	Tsiropoulos et al. ⁵³ and Kelm et al. ⁵⁴
PV (Open field)	320 $\frac{\text{€}}{\text{kW}}$	5.4 $\frac{\text{€}}{\text{kW a}}$	20	Tsiropoulos et al. ⁵³
Wind (Onshore)	1000 $\frac{\text{€}}{\text{kW}}$	25 $\frac{\text{€}}{\text{kW a}}$	20	Tsiropoulos et al. ⁵³ and Kreidelmeyer et al. ⁵⁵
Wind (Offshore)	2530 $\frac{\text{€}}{\text{kW}}$	63 $\frac{\text{€}}{\text{kW a}}$	20	Robinius et al. ⁵⁶
Li-ion batteries	131 $\frac{\text{€}}{\text{kWh}}$	3.3 $\frac{\text{€}}{\text{kWh a}}$	15	Stolten et al. ⁵⁷
H ₂ salt caverns	0.7 $\frac{\text{€}}{\text{kWh}}$	0.01 $\frac{\text{€}}{\text{kWh a}}$	40	Caglayan et al. ⁵⁸
Electricity grid	0.86 $\frac{\text{€}}{\text{kW km}}$	0.03 $\frac{\text{€}}{\text{kW km a}}$	40	Etri ⁵⁹
H ₂ pipelines	0.185 $\frac{\text{€}}{\text{kW km}}$	0.01 $\frac{\text{€}}{\text{kW km a}}$	40	Caglayan et al. ²⁸
Electrolysers	350 $\frac{\text{€}}{\text{kW}}$	11 $\frac{\text{€}}{\text{kW a}}$	10	Stolten et al. ⁵⁷
CCGT hydrogen gas	760 $\frac{\text{€}}{\text{kW}}$	23 $\frac{\text{€}}{\text{kW a}}$	20	Stolten et al. ⁵⁷

Additionally, in ETHOS.FINE we also assumed a self discharge of 0.004% per hour for Li-ion batteries.

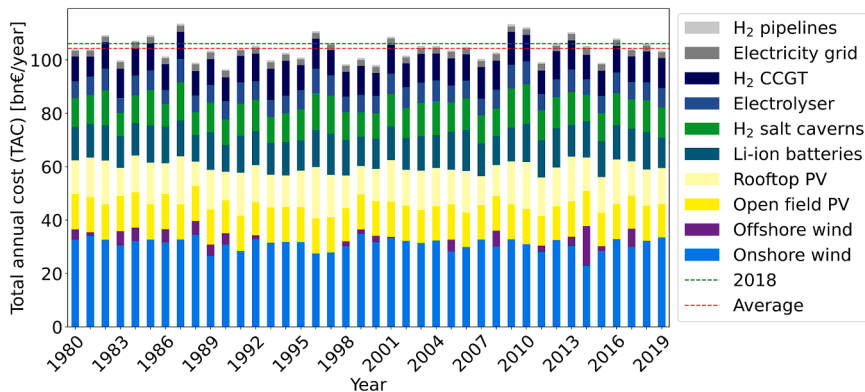


Figure 2. TAC comparison by technology for ESMs optimized from 1980 to 2019 aggregated for a 38 node Germany model set up in ETHOS.FINE

In summary, this strongly supports the results by Cagalayan et al.^{27,28}—weather patterns matter and ignoring in energy systems with significant renewables may lead to systems that have large supply gaps under anything but the most optimal conditions, which we cannot simply assume to be covered through imports.

Effect of modifications on installed capacities

All modifications lead to increased investment in CCGT, salt caverns and electrolysers, although the latter is less pronounced for MOD 1/Demand increase, as seen in Table 4. This is expected, as CCGT can provide electricity when the generation is low and demand is high and installing more CCGT also requires more H₂ infrastructure such as salt caverns for storage and electrolysers for H₂ conversion.

Note that the more detailed model implemented in ETHOS.FINE was used for evaluation of MOD 2/Synthetic time-series, since it outperformed the other algorithms when implemented in *gurobipy* and is therefore viewed as the most promising approach. An additional advantage of MOD 2/Synthetic time-series: It can be easily integrated into existing modeling frameworks.

MOD 1/Demand increase leads to robust energy systems regardless of the initial time-series chosen. On average, this in-

currurs additional cost of 7.8bn€ if no smoothing is performed. The average total cost of a robust system reaches 63.1bn€, with a range of [60.9,69.7]bn€. In comparison, using smoothing leads to slightly more expensive solutions. On average, making an ESM robust incurs additional cost of 9.7bn€(+17%).

The average total cost of a robust system reaches 65.2bn€, with a range of [61.0,72.7]bn€.

Figure 6 gives the results of optimizing each year independently for the smoothed MOD 1/Demand increase. While on average more expensive than non-smoothed MOD 1/Demand increase, the smoothed algorithm leads to on average lower investment increase in short-term battery storage (+ 25% non-smoothed vs. + 12% for smoothed) and a significantly higher investment in CCGT (+ 86% for non-smoothed vs. + 36% for smoothed). The higher costs may be due to the fact that additional artificial demand is added in time periods adjacent to those with previous supply gaps, which generates small supply demand gap time periods. The strong invest in CCGT compared to the non-smoothed modification suggests that CCGT power plants are used to offset those artificial demand gap time periods. For a visual comparison of the smoothing effect compared to non-smoothed we refer to Figure 12 in the appendix.

In summary, smoothed MOD 1/Demand increase leads to feasible solutions, but care should be taken that artificial energy demands do not lead to excessive building of CCGT power plants. Convergence of MOD 1/Demand increase can be slow, sometimes taking more than 20 iterations for a single pair of years. This number of iterations was used as a cut-off criterion, as no improvements were observed after that during testing.

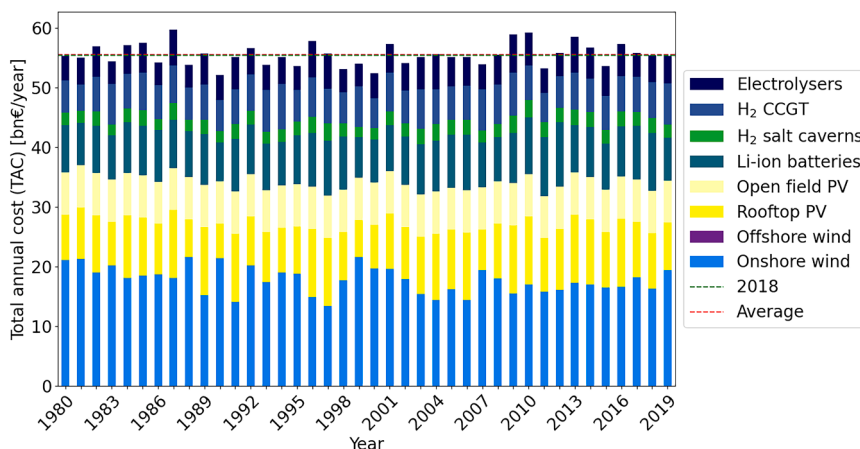


Figure 3. TAC comparison by technology for ESMs optimized from 1980 – 2019 on a single node Germany model set up using gurobipy

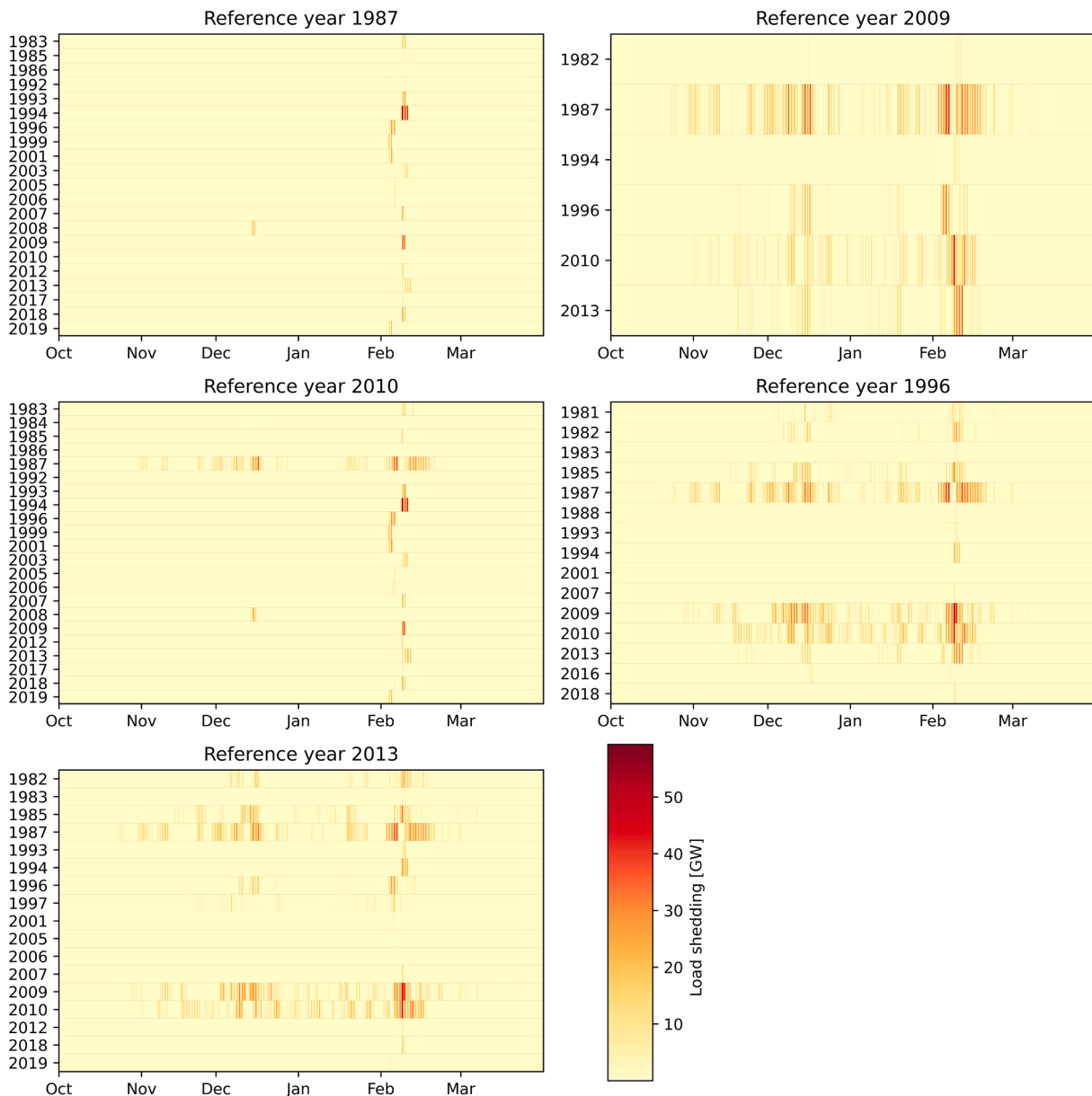


Figure 4. Load shedding when testing the feasibility of the 5 selected reference years in ETHOS.FINE

Only the years and the months from October to March are included, where load shedding occurs. For comparison, hourly load in the demand data, we use roughly 100 GW.

This appears to be due to very small residual supply gaps of a few GWh that get found and added to the model repeatedly. Given the small size of those supply gaps, the large overall production, and the fact that Gurobi²⁹ was used as a solver, which does not perform exact arithmetic, this may be caused by numerical instabilities. Using a suitable termination criterion (e.g., number of iterations or total supply gap less than some small number of GWh) counteracts this.

Notably, non-smoothed MOD 1/Demand increase incurs a bias toward installing more Li-ion battery storage. This is to be expected, as artificial short term demand peaks are added,

and Li-ion batteries are well-suited to compensate for those. Their capacity was increased by on average more than 25%, with a range of [7.7,19.1]bn€, compared to [6.5,10.8]bn€ in the reference years. Finally, non-smoothed MOD 1/Demand increase finds the overall cheapest robust solution. That solution is characterized by slightly more investment in onshore wind capacity (19.8bn€, + 12%) and roof top PV (10.2bn€, + 10%) than in an average single year solution. No additional batteries are installed, but more electrolyzers (5.6bn€, + 18%), CGGT (7.3bn€, + 23%), and salt caverns (2.7bn€, + 16%).

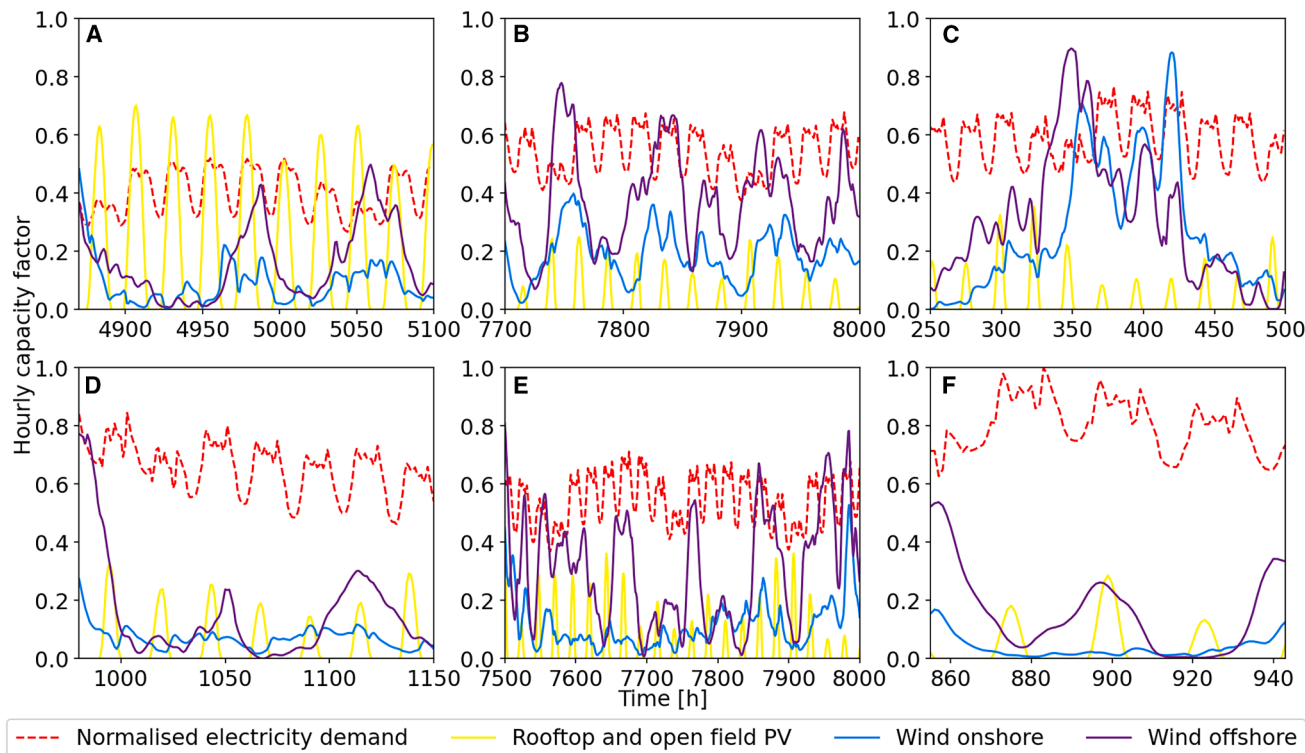


Figure 5. Six time periods from the 40 weather years and 1 year of future electricity demand data for Germany

The electricity demand is normalized to prevent overweighting; the weather data are aggregated. The upper three diagrams represent non-critical time periods, the lower three critical ones. Subfigure (A) is a typical summer period with high PV availability and low electricity demand due to low heating requirements. In (B) and (C), typical autumn and winter period are shown. They are characterized by low availability of PV, but ample wind power to supply sufficient electricity. Note the increased electricity demand due to increased heating required. In (D), a typical dark lull is characterized by low availability of PV and negligible amounts of wind, which coincides with high electricity demand due to increased heating. Subfigure (E) shows an elongated dark lull period. Low availability of both PV and wind combined with increased electricity demand lead to overall difficult period requiring large amounts of hydrogen to be burned in the energy system. The last graphic (F) shows the most critical period in the 40 years of data. Negligible wind combined with low availability of PV coincide with the highest electricity demand in the data due to high heating demand during an extreme cold spell in all of Germany.

MOD 2/Synthetic time-series is evaluated in Figures 7 and 8 for the model in *gurobipy* and the model in ETHOS.FINE, respectively, and Table 5, which give an overview of the results of generating robust ESMs. In the ETHOS.FINE model the five most expensive solutions based on the single years 1987, 1996, 2009, 2010, and 2013 were selected of the 40 energy system designs to make them robust using *MOD 2/Synthetic time-series*. These are hereon referred to as the five reference years.

On the one hand, applying *MOD 2/Synthetic time-series* to the *gurobipy* model leads to robust solutions with the lowest cost increase on average. As can be seen in Figure 7, the TAC is similar across all 40 years. On average, additional cost of 5.6bn€ or 10% with a range of [60.5, 63.15] are the result of this modification. The overall deviation from the average robust solution is in the range [-1%, 3%] showing the effectiveness of the algorithm independent of the selected reference year. The consistently lower costs indicate that *MOD 2/Synthetic time-series* is the best performing modification.

On the other hand, applying *MOD 2/Synthetic time-series* in ETHOS.FINE is shown in Figure 8. After modifying the five selected reference years with *MOD 2/Synthetic time-series* to

make the solutions robust, the share of TAC for wind onshore decreases (-2% to 27%). Similarly, a decrease in total investment in transmission (electricity grid and hydrogen pipeline) is observed (0% - 25%). A general increase in investment is seen for PV (+4% to +15%), Li-ion batteries (-1% to +26%) as well as the hydrogen sector (+9% to +18%) for the robust energy system designs. The increase of PV can be explained by its below average, but still relevant, availability during dark lulls combined with Li-ion batteries to cover daily fluctuations. As visible in Figure 9, PV is mainly utilized together with Li-ion batteries to cover the fluctuating part of the electricity demand, the CCGT cover the bulk of the electricity demand, while the generation from wind is negligible. Hydrogen is utilized for electricity generation to a higher degree, since it can provide flexible additional energy supply, especially during dark lulls. The overall increase in cost compared to the average cost of each of the five reference years is +12% to 13%, compared to the weather year 2018 it is 10% - 12% and compared to the most expensive single year, which is a lower (dual) bound on the objective, it is 2.9% - 5%. Thus, *MOD 2/Synthetic time-series* leads to robust and on average cheaper solutions in ETHOS.FINE similar to the results in the *gurobipy* model.

Table 4. Comparison of the different modifications in terms of convergence and performance

Modification	Increased Investment	Average	Least
MOD 1/Demand increase	Li-Ion, H ₂ -CCGT, Salt caverns	+5.7%	+2.0%
Smoothed MOD 1/Demand increase	H ₂ -CCGT, Salt caverns	+9.2%	+2.4%
MOD 2/Synthetic time-series gurobipy	PV, H ₂ -CCGT, Salt caverns, H ₂ -Electrolysers	+2.5%	+1.6%
MOD 2/Synthetic time-series ETHOS.FINE	PV, H ₂ -CCGT, Salt caverns, H ₂ -Electrolysers	+3.7%	+2.9%
MOD 3/Combine	PV, Wind, H ₂ -CCGT, Salt caverns, H ₂ -Electrolysers	+10.6%	+2.5%

Average and least cost increase compared to best lower (dual) bound, the highest TAC of a single unmodified year.

This also implies that optimization based on average or recommended reference years systematically underestimates the required cost for robust energy supply by > 10% in ETHOS.FINE.

Compared to the results of this work, Ryberg¹⁴ estimates a residual load of about 61 GW and additional backup capacity required of about 25 GW for Germany. The difference to the 118–125 GW found in this study can be attributed to the fact that in the integrated European setting that Ryberg¹⁴ used, dark lulls can be partly suppressed by electricity transmitted from regions not hit by that dark lull as well as differences in demand data.

In either case, the cold dark lull period is the most critical for CCGT—their installed capacity is mainly driven by a single dark lull period, as shown in Figure 9. Figure 9 also illustrates the effects of MOD 2/synthetic time-series. The left graphic shows the result of the ESM optimized for 1987 when testing its feasibility in 1994 reveals a supply gap. After applying MOD 2/synthetic time-series, the time period gets integrated into the optimization problem and after reoptimizing the supply can now be covered using existing capacities.

Modification MOD 3/Combine uses several types of cutting planes principles in one algorithm. It converges for all years, often only requiring one iteration of 2. Sometimes, multiple iterations of MOD 1/demand increase are necessary as well. The effect of MOD 3A/yearly balance is marginal: It does not effect model run times, nor results.

The total costs average out to 66.0bn€ per year, with a range of [61.2,72.1], which is slightly more than MOD 1/Demand increase. Thus, in terms of costs, the MOD 1/Demand increase is preferable. Figure 10 gives the results of optimizing each year independently for MOD 3/Combine. Notably, the results for each year are very similar to each other, suggesting robust solutions share some traits. Cheap solutions contain less onshore wind.

Comparison of modifications

All three featured approaches generate robust solutions, i.e., solutions with no supply gaps, subject to numerical tolerances. The modifications' performance is summarized in Table 4.

The results show a moderate increase in investment for robust models leading to TAC increases of 1.6to2.9% compared to the best available lower (dual) bound on the cost of a robust solution. The dual bound is given by the maximum total cost for a single year, since the cost of a robust solution cannot be lower than the cost of a single year. Thus, the highest cost year is also the only one that could be robust (Clearly, this can be shown via a proof by contradiction, since each solution is assumed to be optimal and thus there must not be a feasible solution with a lower objective value). For gurobipy, the lower bound is 59.7bn€ annually. For the model in ETHOS.FINE, it is 113.6bn€ annually.

An upper (primal) bound is given by the maximum capacity for each technology per region and year, since this will be feasible

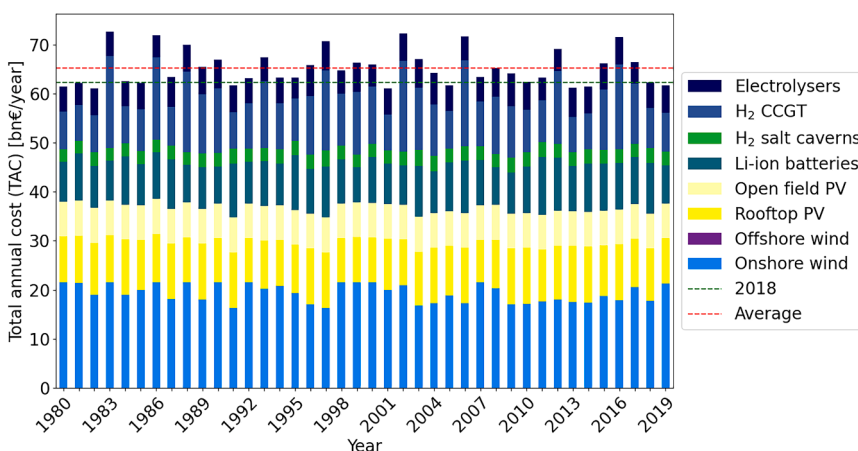


Figure 6. TAC comparison from 19802019 for robust solutions using modification smoothed MOD 1/Demand increase for the single node model in gurobipy, no temporal aggregation

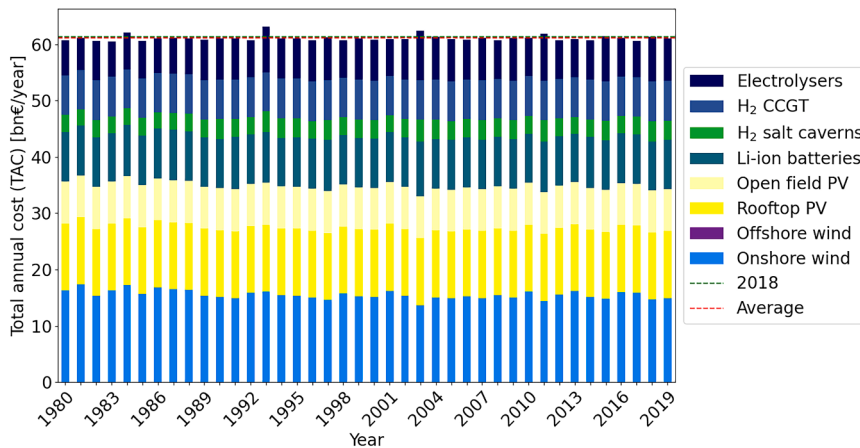


Figure 7. TAC comparison for optimizing for from 1980-2019 for robust solutions using modification MOD 2/Synthetic time-series gurobipy for the single node model in gurobipy, no temporal aggregation

for all years. This results in a range of [113.6,195.66]bn€ annually for an optimal robust energy system in ETHOS.FINE. For the *gurobipy* model, the range for optimal robust solutions is [59.7,67.6]bn€ annually.

Capacity changes in near-optimal robust solutions

It is important to acknowledge that any single robust solution may not accurately capture the properties of robust solutions as a whole, especially if considering the space of near-optimal solutions that may be relevant to decision makers, see the recommendations by Lombardi et al.³⁰ While we did not explicitly perform *modelling-to-generate-alternatives* on the solution space, computing 40 extremal solutions on single years and iteratively robustifying them gives us a sample of 40 near-optimal robust solutions that allow us to infer some properties of near-optimal robust solutions:

Compared to optimizing individual years, near-optimal solutions to the robust model systematically use less onshore wind. This is plausible, since most years contain no extended (dark) lull periods coinciding with peak demands. In years without extended dark lulls, wind power provides stable and cheap energy, compared to PV that might require more storage and conversion units. However, in years with dark lulls this advantage disappears. As such, if

costs for storage and conversion are priced in, optimizing a year without a dark lull may lead to more investment in onshore wind than would be efficient. Integrating appropriate dark lull periods, as suggested in this work, might help counteract that effect, leading to a more balanced energy mix. A functioning capacity market, especially for backup technologies, is essential to ensure the needed capacities are installed and ready to generate electricity during dark lull periods. Overall, robust solutions were only 23% more expensive compared to the lower (dual) bound given by the most expensive single year. Contrarily, a model based on average or recommended reference years systematically underestimates costs by over 10%.

Convergence

The *gurobipy* implementations of MOD 1/Demand increase and MOD 3/Combine tended to quickly converge. However, sometimes residual load differences in the order of MWh/kWh remained. In these cases, the algorithm was terminated after 20 iterations. MOD 2/Synthetic time-series needed between 1 and 54 iterations to converge in the *gurobipy* model. The number of iterations generally decreased for higher annual cost of individual years, but was mainly due to overwriting critical time periods where priority was given to critical time periods selected in earlier iteration steps. In ETHOS.FINE, this issue was prevented by evaluating critical time periods on the modified data largely preventing overwriting. The calculations then needed between 1 and 8 iterations for the 5 reference years. Here, one iteration incurs the same computational load of solving one ESM, or less if warm-starts help reducing computation times.

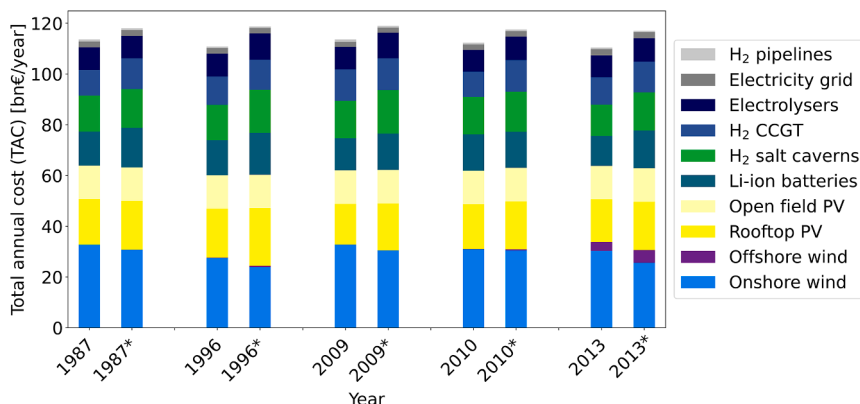


Figure 8. TAC comparison for optimizing for 5 selected reference years for Germany. Columns marked with an * indicate the robust system designed via MOD 2/Synthetic time-series ETHOS.FINE

Table 5. Capacity results of optimizing for 5 selected years for Germany. Columns marked with an * indicate the robust system designed with MOD 2/Synthetic time-series

Capacities	1987	1987*	1996	1996*	2009	2009*	2010	2010*	2013	2013*
Wind (onshore) [GW]	258	243	217	189	259	241	244	241	239	202
Wind (offshore) [GW]	0.0	0.0	0.2	1.8	0.0	0.0	0.4	1.4	12.4	17.7
PV (rooftop) [GW]	400	427	430	505	356	411	394	420	374	423
PV (open field) [GW]	348	348	348	348	348	348	348	348	348	348
Li-ion batteries [GWh]	722	840	742	888	684	770	770	765	635	802
H ₂ salt caverns [TWh]	195	211	191	234	202	237	204	217	169	206
CCGT hydrogen gas [GW]	101	121	112	118	124	125	99	125	108	121
Electrolysers [GW]	142	140	144	165	142	161	137	146	137	148
Electricity grid [GW]	444	444	429	405	386	386	405	425	483	463
Hydrogen pipelines [GW]	914	800	889	686	1029	686	800	686	686	571

Full load hours and system cost

Lower investment in onshore wind capacity was a reoccurring pattern in the effects of modifications. The left graphic in Figure 11 shows the annual full load hours (FLH) for wind on- and offshore as well as PV compared to TAC for the respective models. For wind on- and offshore, these are strongly correlated (Pearson correlation coefficients of -0.77 and -0.81 for wind on- and offshore, respectively). The FLH of PV and the TAC are nearly uncorrelated (Pearson correlation coefficient of -0.04). For wind, this mirrors earlier results of Gotske et al.,²⁵ who showed similar correlations for an European System.

The right graphic in Figure 11 shows the share of TAC of PV and wind, combined for on- and offshore, compared to the TAC. For wind, these are negatively correlated (Pearson correlation coefficient of -0.59) while for PV, these are strongly positively correlated (Pearson correlation coefficient of 0.59) indicating that as full load hours of wind drop, wind capacity gets replaced with PV capacity. The two diagrams in Figure 11 together show that the TACs depend strongly on the availability

of wind. In weather years with low full load hours for wind, an increase in PV capacities is observed indicating a higher reliance on PV in general which is also reflected in the robust years as investment increases for PV capacities for MOD 2/Synthetic time-series and MOD 3/Combine.

Application of results

The approaches outlined in this work can be used to make any ESM more robust to uncertain time-series data. As we showed, this serves to makes models more realistic and more suitable for real-world application preventing supply gaps for a moderate cost increase of 1.6% – 2.9% .

Since uncertainties in demands or costs, which appear in the right-hand side (demands) or the objective (cost) of a mixed-integer linear program (MILP) can be reformulated as constraint-wise uncertainty, the methods outlined in this work can be applied to cost uncertainty as well. Another natural usage for the modifications proposed in this work is as part of a Benders decomposition framework. However, this was not the

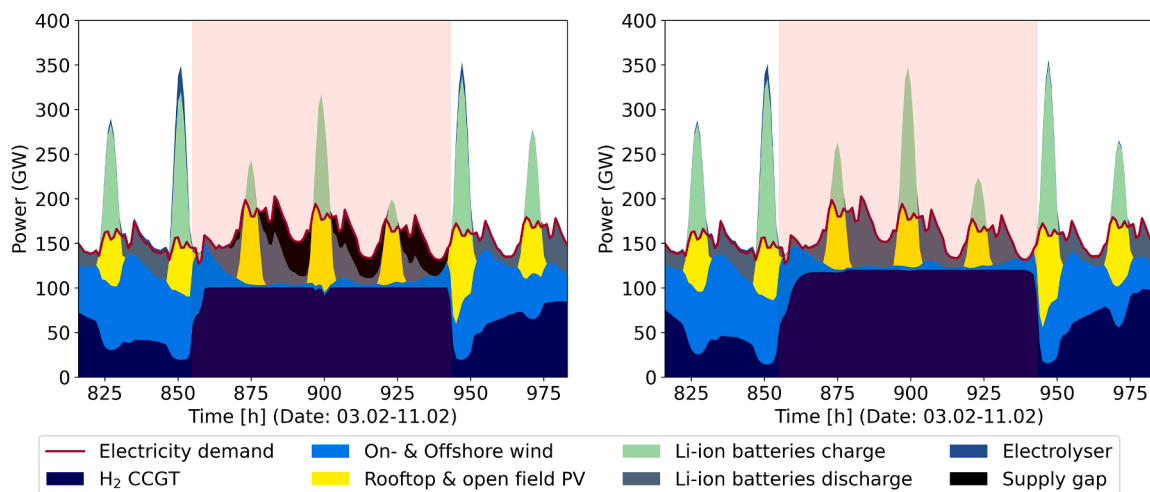


Figure 9. Feasibility testing of the energy system optimized for 1987 in 1994 before modification on the left and after modification on the right
The cold dark lull period in early February (see Figure 5F) is marked in red. In the left graphic, due to insufficient backup capacity the supply gap variable has to be utilized meaning the energy system is not robust, i.e., there are still supply gaps after optimization. In the right graphic, the energy is fully operational during the cold dark lull after applying MOD 2/Synthetic time-series to the original optimization problem.

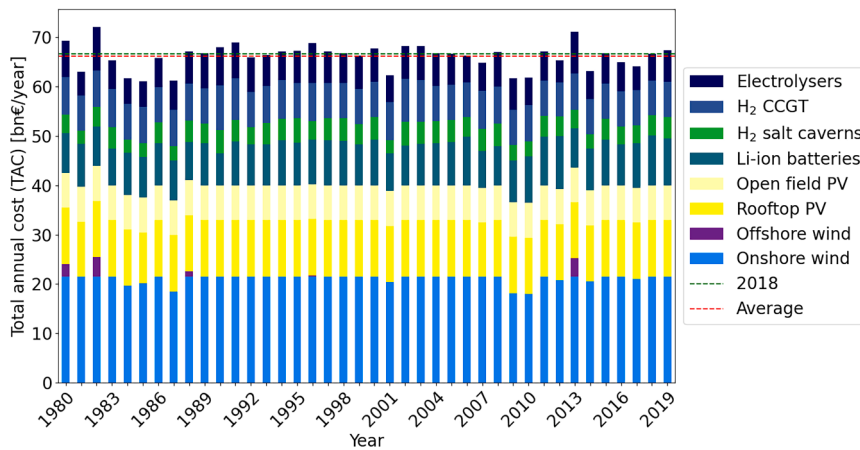


Figure 10. TAC comparison for robust solutions based on single years using MOD 3/Combine, a single node gurobipy model

and local demand. The weather data used are based on historical weather years leaving out the effects of climate change, which very likely will influence the design of energy systems in the future and should be included in studies using these methodologies. The model used to develop the presented method is a Germany model without the possibility of imports or exports. While this island model approach was necessary for

focus of this work, but the constraints that are added in each modification can equally well be used as feasibility/optimalty cuts, as they serve to invalidate significant parts of the solution space. Here, the algorithmic performance for achieving robustness can be seen as a proxy for their potential value as Benders cuts.

Limitations of the study

The main drawback in this case study is that only one time-series for electricity demand in 2050 was used. The availability of future demand data is limited and is difficult to compare if taken from different sources. The selected data from Forschungsstelle für Energiewirtschaft e. V. (FfE)^{31–34} was chosen since it includes a severe cold period rarely observed in Germany. In combination with the weather years, the resulting operational conditions revealed several periods pairing low electricity generation with high demand as can be seen in Figure 5. Therefore, a high degree of robustness can be assured. Further demand time-series might still provide additional insights. This is especially the case if using weather years with an overall cold winter time and therefore a high amount of heating degree days, and thus varying total

method development, the interconnectedness of the European energy system, as well as (cold) dark lulls covering large parts of Europe, is likely crucial, but has been neglected in this study. The Germany model proved sufficient for model development, but a more general model is necessary for more holistic recommendations. Assuming that years with disadvantageous distribution of sunny and wind hours as well as low full load hours are the exception, allowing the last time step to have a lower state of charge than the first would lower the conservatism of the system and reduce cost. This would require a measure of robustness that fully protects against a certain base uncertainty set, but allows using up some stored hydrogen for outlier events. One possible approach for that is outlined in Bärmann et al.³⁵ At the same time, the model only provides an operational schedule under the assumption of perfect foresight within one year. In a more complex model setting, computing an operational schedule dynamically throughout the year might lead to some efficiency losses in the usage of energy, requiring additional investment to counteract this. Perfect foresight is assumed for the optimization of energy systems. This provides the advantage of being able to plan capacities with full information about cold

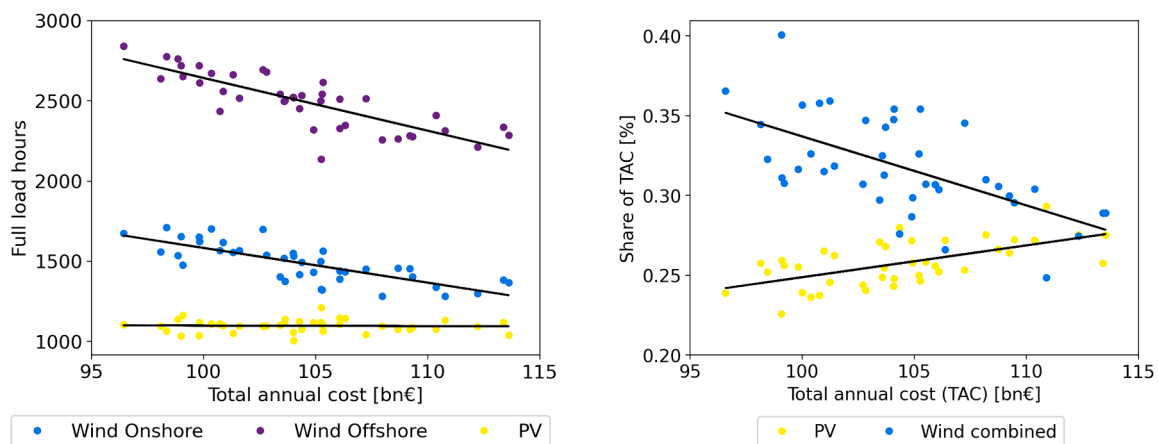


Figure 11. The left graphic shows wind on- and offshore as well as PV full load hours compared to TAC in for all years

The right graphic shows combined cost for wind and for PV as share of the TAC compared to TAC for all years Each dot represents 1 year. Lines indicate regression lines of the different energy generator types.

dark lulls and overall energy generation. In a real-world setting, a higher degree of conservatism leading to additional generation or storage would be needed.

Summary of implications for planning energy systems

Our work contributes to the mounting evidence that suggest that using a fixed reference year and planning a renewable energy system based on that is unsuitable for practice.^{5–7,9,16,18} We show that not only does this lead to misallocation of resources, but to a systematic underestimation of costs for and investment in storage and conversion, i.e., Li-Ion battery capacity and CCGT. Thus, if energy systems are planned based on single years, policy makers need to separately assess how much, not whether, additional storage and conversion capacities are needed and be aware of the fact that this will incur additional costs. For practitioners who model energy systems, we propose three workable approaches that can ensure energy system designs are robust against a range of weather realizations. We also note that the total amount of wind hours per year is strongly correlated with energy system costs. Thus, for a conservative estimate, low-wind reference years are better suited.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Sebastian Kebrich (s.kebrich@fz-juelich.de).

Materials availability

This study did not generate new materials.

Data and code availability

The full code, model, and data are publicly available via GitHub.³⁶ Any additional information required to recreate the results reported in this paper is available from the [lead contact](#) upon request.

ACKNOWLEDGMENTS

This work is co-funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—2236/2. This work was partly supported by the Helmholtz Association as part of the Platform for the Design of a Robust Energy System and Raw Material Supply (RESUR) and the program, “Energy System Design.” This work was partly funded by the European Union (ERC, MATERIALIZE, 101076649). Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

AUTHOR CONTRIBUTIONS

S.K., conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing – original draft, writing – review and editing, and visualization; F.E., conceptualization, methodology, software, validation, formal analysis, investigation, writing – original draft, writing – review and editing, and visualization; D.F., supervision, writing – original draft, and writing – review and editing; C.B., funding acquisition, resources, supervision, writing – review and editing; H.H., funding acquisition, resources, supervision, writing – original draft, and writing – review and editing; J.L., resources.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Appendix

Reference linear program

The following models gives a high-level explanation of the ESM we consider. For a detailed model, we refer to the implementation provided via GitHub.³⁶ Consider a single node model. Let $t \in T$ denote time and $p \in P = P^{prod} \cup P^{stor} \cup P^{conv}$ denote different technologies, i.e., production, conversion and storage. Then, define the following variables:

$x_p \in [0, P_p^{max}] \subseteq \mathbb{R}$	– amount of technology p built,
$s_t^{El} \in [0, S_t^{max}] \subseteq \mathbb{R}$	– electrical energy in storage at time t ,
$s_t^{H_2} \in [0, S_t^{H_2}] \subseteq \mathbb{R}$	– hydrogen in storage at time t ,
$\delta_t^+ \in [0, C_{H_2}^{max}] \subseteq \mathbb{R}$	– energy for H_2 conversion at time t ,
$\delta_t^- \in [0, C_{El}^{max}] \subseteq \mathbb{R}$	– energy from H_2 conversion at time t ,
$\Delta_t \in [0, d_t] \subseteq \mathbb{R}$	– load shedding at time t ,

where the upper bounds are given by technical system specifications, and by the respective demand per time for the load shedding. For simplicity, electrical energy storage is modeled without conversion losses, given that those are comparatively small. Contrary to that, H_2 conversion and storage is modeled explicitly. Furthermore, unit commitment is not explicitly modeled, instead, the modeled guarantees the possibility of sufficient production, which might require turning off some generators in practice.

Then, minimise over

$$\sum_{p \in P} c_p x_p + M \sum_{t \in T} \Delta_t,$$

where $M \in \mathbb{R}^+$ is chosen sufficiently large so that load shedding is never used if an alternative is possible. Note that this objective also implies that operational costs can be modeled as a flat percentage of investment cost. The minimisation above is subject to the following constraints.

First, energy must be conserved at each day:

$$\sum_{p \in P^{prod}} a_{tp} x_p + \delta_t^- - \delta_t^+ + s_{t-1}^{El} - s_t^{El} + \Delta_t \geq d(t) \quad \forall t \in T.$$

Here, a_{tp} are coefficients modeling the production of energy from technology $p \in P^{prod}$ at time $t \in T$ that in general are dependent on the weather scenario chosen. For ease of notation, $-1 \sim |T|$ is used for indexing time.

Second, hydrogen storage and conversion is modeled explicitly:

$$s_t^{H_2} = s_{t+1}^{H_2} + a_j^+ \delta_t^+ - a_j^- \delta_t^-.$$

Here, the $a_j^{+/-}$ coefficients are solely dependent on the conversion technologies used.

Third, both electricity and hydrogen storage must be bounded by the storage capacity built:

$$s_t^{H_2} \leq X_{H_2, storage},$$

$$s_t^{El} \leq X_{El, storage}.$$

Fourth, yearly net energy production must be zero over the course of a year:

$$s_0^{H_2} = s_{|T|}^{H_2},$$

$$s_0^{El} = s_{|T|}^{El}.$$

Techno-Economic parameters

Table 2 shows the techno-economic parameters used in this work.

Description of modifications

/ Add constraint to enforce security of supply for time periods with low supply and high demand

MOD4 / Local Constraints

Begin with an energy system based on a single year time-series data. After performing feasibility testing, consider the case where at least one time period $T \subseteq \mathcal{T}$ is found for which in some weather year j , the given energy system can not supply sufficient energy.

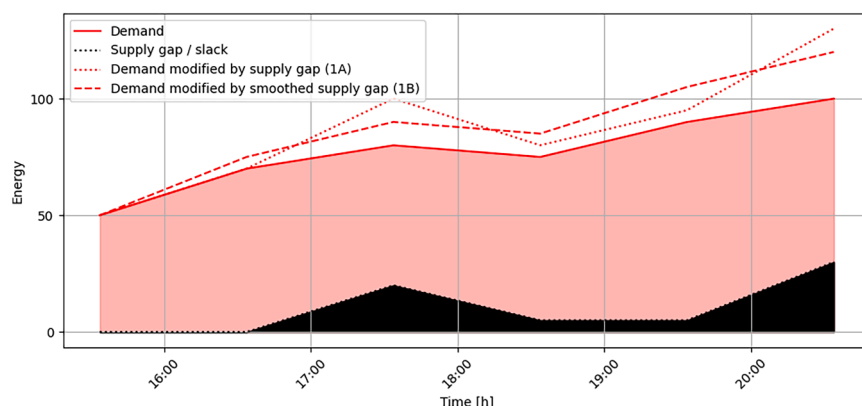


Figure 12. Exemplary visualization of the effect of smoothing for modification 1A and 1B on example data

The y axis has no units, since this is for illustrative purposes only.

For the energy system to supply sufficient energy, the energy supply during that period needs to be at least as large as the demand. Therefore, add a constraint that

$$\sum_{t \in T} \left(\sum_{p \in P_{\text{sub}}} a_{tp}^j x_p \right) + \underbrace{a_{H_2} x_{H_2} \text{ Gas CCGT}}_{\text{Potential energy supply via CCGT during hour } t} - d(t) \geq 0.$$

Renewable energy supply in year j during hour t

Here, x are the capacity variables and $d(t)$ is the energy demand at time t . Note that this contrary to the coefficients a_{tp}^j that model changing weather, the CCGT a_{H_2} conversion parameter is weather-independent.

Battery storage is also weather-independent, but generally has less capacity than gas caverns. Therefore, it is not included in the constraint. However, for a model with small time periods or large expected battery storage, those can be integrated analogously to the H_2 case by adding the maximum battery capacity to the left hand side of the constraint above.

MOD6 / Global-H2

/ Increase H2 storage level at the end of year based on global supply gap.

Both of the approaches mentioned before are designed to compensated for supply gaps in specific time periods. In comparison, Modification 6 aims to ensure sufficient energy supply for a full year as a whole.

For that, consider the vector of slack variables δ that encodes insufficient energy supply throughout the year. We increase the required hydrogen storage level at the end of the year by the total absolute value of the slack vector, i.e., $1^T \delta$. This ensures not only that more energy is supplied, but that the extra energy is supplied as H_2 , which means it can be used flexibly.

Comparison of modifications 1A and 1B

Figure 12 illustrates the differences between the two algorithms. For 1A, the supply gap is directly added to the demand time series. For 1B, smoothing of the supply gap lowers the peaks in the resulting demand profile, while introducing additional loads for nearby time periods.

Evaluation of Ineffective Modifications

None of the modifications that are outlined in Section C are able to generate robust solutions on their own. Reasons for that are outlined below:

For Modification MOD4/Local Constraints, the algorithm will ensure that each cluster period is feasible, if sufficient hydrogen is available. However, the total available hydrogen is restricted by the total available energy and electrolyser capacity, which are not part of Modification MOD4/Local Constraints. Therefore, no convergence can be guaranteed in practice, due to insufficient total energy supply and H_2 conversion capacity.

Compensating this supply gap through the dynamic addition of hydrogen demands in Modification 2 vastly improves performance. If gaps remain, they are in the range of hundreds of GWh, instead of tens of thousands of GWh in Modification MOD4/Local Constraints. However, this still does not

lead to overall feasibility, if no mechanism for continuously adapting either the required extra supply or the total artificial energy demand is provided.

For Modification MOD6/Global $-H_2$, the total available energy is increased eventually, but that does not imply that production capacities are sufficient during each time period. Convergence is slow, often not achieving a robust solution even after 20 iterations. Furthermore, costs tend to be high with large capacities for hydrogen production, but too few CCGT power plants to deal with extended dark lulls.

Finally, Modification MOD 3A/Yearly balance suffers from non-convergence as well, even if large additional energy demands are added. To give an example, in one instance adding $>50\%$ to Germany's total energy demand, using the production parameters from another reference year, was insufficient to enforce feasibility for that specific weather year.

Clearly, this can be shown via a proof by contradiction, since each solution is assumed to be optimal and thus there must not be a feasible solution with a lower objective value.

STAR+METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHOD DETAILS
 - Robustness in capacity expansion modeling
 - Feasibility testing
 - Optimisation problem modifications and clustering
 - Case study of a German ESM
 - Case study data
- QUANTIFICATION AND STATISTICAL ANALYSIS

Received: July 29, 2025

Revised: August 25, 2025

Accepted: February 3, 2026

Published: February 6, 2026

REFERENCES

1. Gielen, D., Boshell, F., Saygin, D., Bazilian, M.D., Wagner, N., and Gorini, R. (2019). The role of renewable energy in the global energy transformation. *Energy Strategy Rev.* 24, 38–50. <https://doi.org/10.1016/j.esr.2019.01.006>.
2. Victoria, M., Zeyen, E., and Brown, T. (2022). Speed of technological transformations required in Europe to achieve different climate goals. *Joule* 6, 1066–1086. <https://doi.org/10.1016/j.joule.2022.04.016>.
3. Ringkjøb, H.-K., Haugan, P.M., and Solbrekke, I.M. (2018). A review of modelling tools for energy and electricity systems with large shares of

- variable renewables. *Renew. Sustain. Energy Rev.* **96**, 440–459. <https://doi.org/10.1016/j.rser.2018.08.002>.
4. IRENA (2022). Global hydrogen trade to meet the 1.5°C climate goal: Part III – Green hydrogen cost and potential (International Renewable Energy Agency, Abu Dhabi), p. 16.
 5. Schyska, B.U., Kies, A., Schlott, M., Von Bremen, L., and Medjroubi, W. (2021). The sensitivity of power system expansion models. *Joule* **5**, 2606–2624.
 6. Ruggles, T.H., Virgúez, E., Reich, N., Dowling, J., Bloomfield, H., Antonini, E.G.A., Davis, S.J., Lewis, N.S., and Caldeira, K. (2024). Planning reliable wind- and solar-based electricity systems. *Adv. Appl. Energy* **15**, 100185. <https://doi.org/10.1016/j.adapen.2024.100185>.
 7. Haddeland, I., Hole, J., Holmqvist, E., Koestler, V., Sidelnikova, M., Veie, C.A., and Wold, M. (2022). Effects of climate on renewable energy sources and electricity supply in Norway. *Renew. Energy* **196**, 625–637.
 8. Staffell, I., and Pfenninger, S. (2018). The increasing impact of weather on electricity supply and demand. *Energy* **145**, 65–78.
 9. Pfenninger, S. (2017). Dealing with multiple decades of hourly wind and PV time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability. *Appl. Energy* **197**, 1–13. <https://doi.org/10.1016/j.apenergy.2017.03.051>.
 10. Collins, S., Deane, P., Ó Gallachóir, B., Pfenninger, S., and Staffell, I. (2018). Impacts of Inter-annual Wind and Solar Variations on the European Power System. *Joule* **2**, 2076–2090.
 11. De Marco, F., Mannhardt, J., Oneto, A., and Sansavini, G. (2025). Climate-resilient energy systems planning via system-informed identification of stressful events. *Adv. Appl. Energy* **19**, 100235. <https://doi.org/10.1016/j.adapen.2025.100235>.
 12. Schlachtberger, D.P., Brown, T., Schäfer, M., Schramm, S., and Greiner, M. (2018). Cost optimal scenarios of a future highly renewable European electricity system: Exploring the influence of weather data, cost parameters and policy constraints. *Energy* **163**, 100–114.
 13. Hilbers, A. P., Brayshaw, D. J., and Gandy, A. (2020). August). Importance subsampling for power system planning under multi-year demand and weather uncertainty. In *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)* (IEEE), pp. 1–6.
 14. Ryberg, D.S. (2020). Generation Lulls from the Future Potential of Wind and Solar Energy in Europe (Ph.D. thesis RWTH Aachen University).
 15. Ryberg, D.S., Heirichs, H., Stolten, D., and Robinius, M. (2019). Occurrence of variable renewable energy system generation lulls in a future european capacity scenario. *Energy Proceedings*.
 16. Ruhnau, O., and Qvist, S. (2022). Storage requirements in a 100% renewable electricity system: Extreme events and inter-annual variability. *Environ. Res. Lett.* **17**, 044018. <https://doi.org/10.1088/1748-9326/ac4dc8>.
 17. Grochowicz, A., van Greevenbroek, K., Benth, F.E., and Zeyringer, M. (2023). Intersecting near-optimal spaces: European power systems with more resilience to weather variability. *Energy Econ.* **118**, 106496.
 18. Grochowicz, A., van Greevenbroek, K., and Bloomfield, H.C. (2024). Using power system modelling outputs to identify weather-induced extreme events in highly renewable systems. *Environ. Res. Lett.* **19**, 54038.
 19. Birge, J.R. (1982). The value of the stochastic solution in stochastic linear programs with fixed recourse. *Math. Program.* **24**, 314–325.
 20. Maggioni, F., and Wallace, S.W. (2012). Analyzing the quality of the expected value solution in stochastic programming. *Ann. Oper. Res.* **200**, 37–54. <https://doi.org/10.1007/s10479-010-0807-x>.
 21. Forghani, K., Kan, X., Reichenberg, L., and Hedenus, F. (2025). Modeling Robust Energy Systems Considering Weather Uncertainty and Nuclear Power Failures: A Case Study in Northern Europe. Preprint at arXiv. <https://arxiv.org/abs/2504.08145>. arXiv:2504.08145.
 22. Zeyringer, M., Price, J., Fais, B., Li, P.-H., and Sharp, E. (2018). Designing low-carbon power systems for Great Britain in 2050 that are robust to the spatiotemporal and inter-annual variability of weather. *Nat. Energy* **3**, 395–403.
 23. Bloomfield, H.C., Brayshaw, D.J., Shaffrey, L.C., Coker, P.J., and Thornton, H.E. (2016). Quantifying the increasing sensitivity of power systems to climate variability. *Environ. Res. Lett.* **11**, 124025. <https://doi.org/10.1088/1748-9326/11/12/124025>.
 24. Dowling, J.A., Rinaldi, K.Z., Ruggles, T.H., Davis, S.J., Yuan, M., Tong, F., Lewis, N.S., and Caldeira, K. (2020). Role of long-duration energy storage in variable renewable electricity systems. *Joule* **4**, 1907–1928. <https://doi.org/10.1016/j.joule.2020.07.007>.
 25. Gøtske, E.K., Andresen, G.B., Neumann, F., and Victoria, M. (2024). Designing a sector-coupled European energy system robust to 60 years of historical weather data. *Nat. Commun.* **15**, 10680.
 26. Goerigk, M., Kurtz, J., Schmidt, M., and Thürauf, J. (2025). Connections between Robust and Bilevel Optimization. *Open J. Mathemat. Optim.* **6**, 1–17. <https://doi.org/10.5802/ojmo.38>. <https://ojmo.centre-mersenne.org/articles/10.5802/ojmo.38/>.
 27. Caglayan, D.G., Heinrichs, H.U., Linssen, J., Robinius, M., and Stolten, D. (2019). Impact of different weather years on the design of hydrogen supply pathways for transport needs. *Int. J. Hydrogen Energy* **44**, 25442–25456.
 28. Caglayan, D.G., Heinrichs, H.U., Robinius, M., and Stolten, D. (2021). Robust design of a future 100% renewable european energy supply system with hydrogen infrastructure. *Int. J. Hydrogen Energy* **46**, 29376–29390.
 29. Gurobi Optimization, LLC (2024). Gurobi Optimizer Reference Manual. <https://www.gurobi.com>.
 30. Lombardi, F., van Greevenbroek, K., Grochowicz, A., Lau, M., Neumann, F., Patankar, N., and Vågerö, O. (2025). Near-optimal energy planning strategies with modeling to generate alternatives to flexibly explore practically desirable options. *Joule* **9**, 102144. <https://doi.org/10.1016/j.joule.2025.102144>.
 31. für Energiewirtschaft, e.V., and FfE), F. (2021). Load curves of the tertiary sector – extremos solideu scenario (europe nuts-3). <https://opendata.ffe.de/dataset/load-curves-of-the-tertiary-sector-extremos-solideu-scenario-europe-nuts-3/>.
 32. Forschungsstelle für Energiewirtschaft e. V. (FfE) (2021). Load curves of the household sector – extremos solideu scenario (europe nuts-3). <https://opendata.ffe.de/dataset/load-curves-of-the-private-household-sector-extremos-solideu-scenario-europe-nuts-3/>.
 33. für Energiewirtschaft e. V. (FfE), F. (2021). Load curves of the transport sector – extremos solideu scenario (europe nuts-3). <https://opendata.ffe.de/dataset/load-curves-of-the-transport-sector-extremos-solideu-scenario-europe-nuts-3/>.
 34. Forschungsstelle für Energiewirtschaft e. V. (FfE) (2021). Load curves of the industry sector – extremos solideu scenario (europe nuts-3). <https://opendata.ffe.de/dataset/load-curves-of-the-industry-sector-extremos-solideu-scenario-europe-nuts-3/>.
 35. Bärmann, A., Büsing, C., and Liers, F. (2019). Globalized Robust Optimization with Gamma-Uncertainties. <https://optimization-online.org/?p=15853>.
 36. Kebrich, S., and Engelhardt, F. (2024). <https://github.com/FZJ-IEK3-VSA/Robust-Capacity-Expansion>. <https://github.com/FZJ-IEK3-VSA/Robust-Capacity-Expansion>.
 37. IRENA. Renewable power generation costs in 2017. Report. Tech. Rep. International Renewable Energy Agency (2018).
 38. Hooker, J. Logic-Based Benders Decomposition: Theory and Applications (2024). ISBN 978-3-031-45038-9. doi:10.1007/978-3-031-45038-9.
 39. Waite, M., and Modi, V. (2020). Electricity load implications of space heating decarbonization pathways. *Joule* **4**, 376–394. <https://doi.org/10.1016/j.joule.2019.11.011>.
 40. Jackson, R., Zhou, E., and Reyna, J. (2021). Building and grid system benefits of demand flexibility and energy efficiency. *Joule* **5**, 1927–1930. <https://doi.org/10.1016/j.joule.2021.08.001>.
 41. Hoffmann, M., Kotzur, L., and Stolten, D. (2022). The Pareto-Optimal Temporal Aggregation of Energy System Models. *Appl. Energy* **315**, 119029. <https://www.sciencedirect.com/science/article/pii/S0306261922004342>. <https://doi.org/10.1016/j.apenergy.2022.119029>.

42. Bahl, B., Kümpel, A., Lampe, M., and Bardow, A. (2016). Time-series aggregation for synthesis of distributed energy supply systems by bounding error in operational expenditure. *Computer Aided Chemical Engineering* 38, 793–798.
43. Hastie, T., Tibshirani, R., Friedman, J.H., and Friedman, J.H. (2009). *The Elements of Statistical Learning: Data mining, Inference, and Prediction*, 2 (Springer).
44. Reddy, C., and Vinzamuri, B. (2018). A Survey of Partitional and Hierarchical Clustering Algorithms, pp. 87–110. <https://doi.org/10.1201/9781315373515-4>.
45. Welder, L., Ryberg, D.S., Kotzur, L., Grube, T., Robinius, M., and Stolten, D. (2018). Spatio-temporal optimization of a future energy system for power-to-hydrogen applications in germany. *Energy* 158, 1130–1149.
46. Klütz, Theresa (2025). Ethos. fine: A framework for integrated energy system assessment. *J. Open Source Soft.* 10, 6274. <https://doi.org/10.21105/joss.06274>.
47. Pfenninger, S., and Staffell, I. (2016). Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy* 114, 1251–1265.
48. Staffell, I., and Pfenninger, S. (2016). Using bias-corrected reanalysis to simulate current and future wind power output. *Energy* 114, 1224–1239.
49. Risch, S., Maier, R., Du, J., Pflugradt, N., Stenzel, P., Kotzur, L., and Stolten, D. (2022). Tool for renewable energy potentials - database. Zenodo. <https://doi.org/10.5281/zenodo.6414018>.
50. Risch, S., Maier, R., Du, J., Pflugradt, N., Stenzel, P., Kotzur, L., and Stolten, D. (2022). Potentials of Renewable Energy Sources in Germany and the Influence of Land Use Datasets. *Energies* 15, 5536. <https://doi.org/10.3390/en15155536>. <https://www.mdpi.com/1996-1073/15/15/5536>.
51. Council of European Union (2003). Council regulation (EU) no 1059/2003. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R2066>.
52. (2024). Deutscher Wetterdienst (Wetterrückblick). <https://www.wetteronline.de/wetterueckblick/rueckblick-februar-2012-sehr-kalt-trocken-und-sonnig-2012-02-29-fr>.
53. Tsiropoulos, I., Tarvydas, D., and Zucker, A. (2018). Cost Development of Low Carbon Energy Technologies: Scenario-Based Cost Trajectories to 2050 (Publications Office of the European Union).
54. Kelm, T., Metzger, J., Jachmann, H., Günnewig, D., Püschel, M., Schick-etanz, S., Kinast, P., Thylmann, M., and Nazerian, V. (2019). Vorbereitung und Begleitung bei der Erstellung eines Erfahrungsberichts gemäß § 97 Erneuerbare-Energien-Gesetz: Teilvorhaben II c: Solare Strahlungsenergie (Abschlussbericht). https://www.erneuerbare-energien.de/EE/Redaktion/DE/Downloads/bmwi_de/zsv-boschundpartner-vorbereitung-begleitung-eeg.pdf.
55. Kreidelmeier, S., Dambeck, H., Kirchner, A., and Wünsch, M. (2020). Transformationspfade für strombasierte Energieträger (Prognos AG on behalf of the German Federal Ministry for Economic Affairs and Energy).
56. Robinius, M., Markewitz, P., Lopion, P., Kullmann, F., Heuser, P., Syranidis, K., Cerniauskas, S., Schöb, T., Reuß, M., Ryberg, S., et al. (2020). Wege für die Energiewende: Kosteneffiziente und klimagerechte Transformationsstrategien für das deutsche Energiesystem bis zum Jahr 2050 (Forschungszentrum Jülich GmbH Zentralbibliothek).
57. Stolten, D., Markewitz, P., Kraus, S., Franzmann, D., Schöb, T., Grube, T., Heinrichs, H., Gillessen, B., Linssen, J., Pflugradt, N., et al. (2022). Neue Ziele auf alten Wegen? Strategien für eine treibhausgasneutrale Energieversorgung bis zum Jahr 2045 (FZJ-2022-02577 Technoökonomische Systemanalyse).
58. Caglayan, D.G., Weber, N., Heinrichs, H.U., Linßen, J., Robinius, M., Kulkla, P.A., and Stolten, D. (2020). Technical potential of salt caverns for hydrogen storage in Europe. *Int. J. Hydrogen Energy* 45, 6793–6805.
59. Etri, E. (2014). Energy Technology Reference Indicator Projections for 2010–2050 (Joint research centre (JRC), European Commission (EC)).

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Load Curves of the Tertiary Sector – eXtremOS solidEU Scenario (Europe NUTS-3)	FfE, eXtremOS Project	https://opendata.ffe.de/dataset/load-curves-of-the-tertiary-sector-extremos-solidEU-scenario-europe-nuts-3/
Load Curves of the Household Sector – eXtremOS solidEU Scenario (Europe NUTS-3)	FfE, eXtremOS Project	https://opendata.ffe.de/dataset/load-curves-of-the-private-household-sector-extremos-solidEU-scenario-europe-nuts-3/
Load Curves of the Transport Sector – eXtremOS solidEU Scenario (Europe NUTS-3)	FfE, eXtremOS Project	https://opendata.ffe.de/dataset/load-curves-of-the-transport-sector-extremos-solidEU-scenario-europe-nuts-3/
Load Curves of the Industry Sector – eXtremOS solidEU Scenario (Europe NUTS-3)	FfE, eXtremOS Project	https://opendata.ffe.de/dataset/load-curves-of-the-industry-sector-extremos-solidEU-scenario-europe-nuts-3/
40 years of wind and PV data in hourly resolution	Pfenninger et al.	https://www.renewables.ninja/
Tool for Renewable Energy Potentials – Database	Stanley et al.	https://zenodo.org/records/6414018
Council Regulation (EU) No 1059/2003	Eur Lex, European Union	https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R2066
Software and algorithms		
Full code, model and data	This paper	https://github.com/FZJ-IEK3-VSA/Robust-Capacity-Expansion

METHOD DETAILS

After solving a initial CAPEX problem, the proposed algorithm consists of four main steps, as shown in Figure 1: 1. feasibility testing using data of other years (In robust optimisation, *scenarios* are commonly used as a term to denote different possible realisations of uncertainty. In this setting, each scenario consists of one year–long time–series of data for each type of renewable energy in combination with the demand affected by the weather. As such, this is a *discrete uncertainty set*. While we introduce these notions for weather patterns only, the theory applies to any finite scenario set and their convex combination $\mathcal{U} = \text{conv}(u_1, \dots, u_n)$), 2. assessing loss of load/supply gaps, 3. identification of reasons for infeasibility, specifically critical time periods to derive modifications and, 4. solving the modified CAPEX problem before repeating steps 1–4. Before describing these steps, we define robustness for ESMs.

Robustness in capacity expansion modeling

As noted before, general ARO has a bilevel structure: first stage decisions are made, some (worst–case) uncertainty realises, but then second stage decisions can be then made based on the then known uncertainty and the fixed first stage decisions. For energy system modeling, we let the capacity decision variables be the first stage, since new power plants can not be built ad hoc to react to weather variations. The second stage decisions consist of operational decisions that can be made based on known weather variations. Thus, all second stage costs are operational expenditures (OPEX). We assume that the OPEX can be given as a flat percentage of investment costs/capital expenditure (c_{CAPEX}^T), which is a reasonable assumption for many renewables such as wind power and solar PV.³⁷ This is assumed for all technologies included, but the fuel cost for hydrogen, which are endogenous, since hydrogen gas turbines can only use hydrogen produced by the system itself.

Via those simplifications, we remove costs from the second stage decisions. Mathematically, this is a crucial simplification as it allows us to collapse the “min–max–min” objective of general ARO into a simple minimisation problem. A second motivation for doing that is that our approach emulates the principle of a Bender’s decomposition (see the most recent book by Hooker³⁸ for an overview of Bender’s decomposition). By removing the second stage costs, we remove the need for optimality cuts, which significantly simplifies computations. This results in the following, simplified ARO model:

$$\begin{aligned}
 & \min_{x \in \mathcal{X}} c_{\text{CAPEX}}^T x \\
 & \text{s.t.} \quad Ax + B(u)y(u) \geq b(u) \quad \forall u \in \mathcal{U} \\
 & \quad \quad y(u) \in \mathcal{Y}(x, u) \quad \forall u \in \mathcal{U}.
 \end{aligned}
 \tag{ARO-ESM}$$

Here, we model supply and dispatch of electricity, (dis-)charge of storage systems as well as transmission and conversion of electricity and hydrogen for the technologies given in Table 2.

Reference linear program gives an overview of the simplified model used for implementing constraints and variable demands to allow for an easy and fast testing of the multiple solution modifications used below. The full code, model and data are publicly available via GitHub.³⁶

Feasibility testing

A solution to the mixed-integer program ARO-ESM that is optimal and feasible for one year can lead to supply gaps for other years. Furthermore, since *MOD 2 / Synthetic time-series* modifies the weather data resulting in artificial weather years, which are not based on atmospheric physics. The feasibility testing serves as a method to ensure that optimized energy systems can be operated in historic weather years. Therefore, feasibility testing has two purposes: First, to detect supply gaps. Second, to ensure that optimized energy systems are realistic. We call an ESM *robust* against a set of years \mathcal{U} , if and if only it can be operated without supply gaps for any year $u \in \mathcal{U}$:

Definition 1 (Robust energy systems) An energy system is robust against a set of time-series data \mathcal{U} if and if only for each time-series, there exists a operation schedule that supplies all demand in time, while ensuring that the total amount of energy in storage is at least as big at the end of a year, as it was at the beginning.

Remark that this definition explicitly defines robustness relative to a *known* uncertainty set \mathcal{U} . Furthermore, it requires non-decreasing storage, which implies long-term supply security. That means even multiple sequential years with low full load hours for renewables will not deplete storage levels. This is a risk-averse strategy, as increased energy supply in beneficial years can not be used to offset lack of supply in bad years.

Now consider a solution $x_{u_i}, y(u_i)$ for one year $u_i \in \mathcal{U}$. To evaluate how well a given system would have performed in a different year $u_j \in \mathcal{U}, i \neq j$ it is necessary to determine whether for an optimized operating schedule any supply gaps remain. This is equivalent to reoptimising

$$\begin{aligned} & \min_{y(u_j) \in \mathcal{Y}(x_{u_i}, u_j), \delta_{ij} \in \mathbb{R}_{\geq 0}^n} \mathbf{1}^\top \delta_{ij} \\ \text{s.t. } & B(u_j)y(u_j) + \delta_{ij} \geq b - Ax_{u_i} \end{aligned}$$

Here, δ is a vector of supply gaps. It is implied that δ_{ij} is only added to constraints that cover energy supply/demand. Note that in the mixed integer program above, x_{u_i} , the capacity expansion, is a fixed input parameter retrieved from the selected year. Now an energy system designed for year u_i is *robust* for any year u_j if and if only there exists an solution to the MILP $COMP_j^i$ above such that $\delta = 0$. If not, δ_{ij} encodes where supply gaps are and how large they are (Note that this is a relation between two x_{s_i}, x_{s_j} solutions, not their respective underlying time-series data, since there may be multiple (near-)optimal solutions for a given year. Computationally, this never proved to be an issue, so these notions are used interchangeably throughout this work. However, care should be taken if multiple capacity decisions have equal pricing).

Optimisation problem modifications and clustering

If an ESM displays supply gaps during feasibility testing, its CAPEX optimisation problem is modified to obtain a robust solution. Consider a solution δ to $COMP_j^i$. In this solution, $\mathbf{1}^\top \delta_j^*$ units of energy are missing. This information needs to be integrated into the original CAPEX problem to enable the ESM to provide the extra energy supply required to meet demand.

Multiple modifications were iteratively developed to capture different sources of uncertainty. Initial testing showed a subset of them to be ineffective. In order to also report null-results, these modifications are outlined in [description of modifications](#). The three promising modifications are inspired by three physical phenomena relevant to energy systems: 1) The importance of peak demands for energy system design,^{39,40} 2) The existence of extended critical time periods^{14,16-18} and 3) variations in total yearly energy supply and demand across different years.¹⁰

For *MOD 1 / Demand increase*, update the demand vector b_j of the starting year j by adding the demand-supply gap δ to it, i.e.,

$$b'_j = b_j + \delta. \quad (\text{MOD 1/Demand increase})$$

This adds the missing energy supply exactly when it is needed. One observation during initial testing was that this can lead to very large peaks in artificial demand, especially if done repeatedly over multiple iterations, leading to excessive battery installation. To counteract this, the artificial, additional demand can be divided up between neighboring time periods, i.e., update b and smooth δ , i.e., $b' = b + f(\delta)$ for some function $f: \mathbb{R}_{\geq 0}^n \leftarrow \mathbb{R}_{\geq 0}^n \setminus \emptyset$ that changes b locally. [Figure 12](#) in [description of modifications](#) illustrates the effects of smoothing. Throughout our work, we differentiate between smoothed and non-smoothed *MOD 1 / Demand increase*.

Since feasibility tests show that supply gaps usually occur in adjacent time steps, those are clustered to identify time periods with potential supply gaps, i.e., *critical time periods*. Critical time periods are usually characterised by a combination of low electricity generation from renewables together with increased electricity demand, so-called (*cold*) *dark lulls*. A hierarchical clustering was chosen, as this has been shown to be effective for time-series aggregation previously,^{41,42} and it allows for easy visual inspection and clustering along adjacent time-steps.^{43,44} For each time period, check whether the optimised capacities provide sufficient energy. For this, calculate the potential electricity generation of the selected ESM and compare this with the total demand in the time period (Note that a positive average supply gap means that the investigated energy systems lack dispatchable capacity, and thus the feasibility testing of the respective time periods will always find a supply gap. However, a time period with negative average hourly supply gaps

can still contain supply gaps that are only found during feasibility testing. This happens if either total annual electricity generation is too low to cover overall demand, or if there are some demand peaks that can not be covered). For calculating the potential electricity generation, full load is assumed for generation and backup capacities. Let $b_{t,j}$ be the demand at hour $t \in T$, and $a_{t,j}^k \cdot x_s^p$ the time-series data and the calculated capacities of the energy system component $p \in P$ in year u_j , respectively. $P_{\text{sup,con}} = P_{\text{sup}} \cup P_{\text{con}}$ is used to denote that summation only happens over supply and conversion technologies. Then, for time interval $T \subseteq \mathcal{T}$, the average hourly supply gap $\Delta(T)_j$ for some year u_j is given via

$$\Delta(T)_j := \frac{\sum_{t \in T} \sum_{p \in P_{\text{sup,con}}} a_{t,j}^p \cdot x_s^p - b_{t,j}}{\sum_{t \in T} 1} \quad (\Delta(T))$$

MOD 2/Synthetic time-series, iteratively inserts critical time periods into the original data to create synthetic time-series data for reoptimisation by first, selecting the ones with negative average hourly supply gap and second, when no critical time periods with negative hourly supply gap can be detected anymore, by reducing the sum of full load hours for PV and wind of the time period compared to the original data, i.e.,

$$a_{t,j}^p = a_{t,j'}^p \quad \forall t \in T, p \in P \quad (\text{MOD 2/Synthetic time-series})$$

for some critical time period T and years j, j' .

MOD 3/Combine consists of multiple parts. First, demand a weighted positive total energy balance for every year to ensure that there is sufficient energy overall. Then, if during feasibility testing a positive slack δ is detected for year j , add

$$\underbrace{\sum_{t \in T} \left(\sum_{p \in P_{\text{sup}}} a_{t,p}^j x_p \right)}_{\text{Renewable energy supply in year } j} - d(t) \geq \min\{0, 1^T \delta\}. \quad (\text{MOD 3A/Yearly balance})$$

to the model. Here, the total demand is increased by the total absolute value of the slack vector, i.e., $1^T \delta$. This is necessary since some energy will be lost during storage, conversion and transmission, which is not captured in the $a_{t,p}^j$ model parameters and the supply capacities P_{sup} .

If this does not suffice to achieve feasibility for some year j , we model long-term energy storage more explicitly. In our model, H_2 is used for that. The modeling is then done dynamically; for critical time periods T , auxiliary variables $\sigma_{i,T}$ are introduced that represent the use of CCGT power plants during periods of low renewable energy supply:

$$\underbrace{\sum_{t \in T} \left(\sum_{p \in P_{\text{sup}}} a_{t,p}^j x_p \right)}_{\text{Renewable energy supply in year } j \text{ during period } T} + \sigma_{i,T} - d(t) \geq 0. \quad (\text{Equation 1})$$

At the beginning, the hydrogen storage is at $s_0^{H_2}$. In every hour after that, depending on the net energy balance, either more energy is stored or energy is taken from storage via H_2 to electricity conversion. At any point during a year, the hydrogen stored so far must be sufficient to at least cover all hydrogen demands during past/present critical periods. This can again be encoded as a MILP constraint:

$$\underbrace{s_0^{H_2}}_{\text{Initial } H_2 \text{ storage}} + \alpha \underbrace{\sum_{t \in T, t < T} \left(\sum_{p \in P_{\text{sup}}} a_{t,p}^j x_p \right) - d(t)}_{\text{Net renewable energy balance in year } j \text{ to up time period } T} \geq \underbrace{\sum_{T' \subseteq T, T' < T} \sigma_{T'}}_{\text{Minimum } H_2 \text{ required for year } j \text{ to up time period } T}. \quad (\text{MOD 3/Combine})$$

The model parameter α encodes energy losses due to power to H_2 conversion. Note that this only is a lower bound on the energy required, since it is possible that more energy is used up during conversion in non-critical time periods or due to storage or transmission losses.

After reoptimisation only small supply gaps should remain. Subsequently **MOD 1 / Demand increase** that introduces additional demand is used to ensure leftover short-term supply gaps are covered, i.e., the remaining supply gaps are added as an artificial electricity demand in the time periods where they appear.

Case study of a German ESM

The algorithm and all modifications were implemented and evaluated based on a fully renewable ESM for Germany in 2045 that includes all technologies listed in [Table 2](#).

Two implementations were evaluated. First, a benchmark model was implemented in *gurobi* that allows for comparison of all algorithms. A description of this model can be found in [reference linear program](#). Then, a more detailed 38 node model was implemented in the open source Framework for Integrated Energy System Assessment (ETHOS.FINE)⁴⁵ within the ETHOS modeling

suite⁴⁶ to ensure validity of the results. This includes time-series aggregation via the TSAM package.⁴¹ Since ETHOS.FINE currently does not allow interacting with the solver during the solution process, which is necessary to add cutting planes, this was only done for *MOD 2 / Synthetic time-series*.

Case study data

40 years of wind and PV data in hourly resolution were taken from renewables.ninja, using data originally provided by Staffell and Pfenninger.^{47,48} They provide decades of hourly open data for both wind and PV allowing for easy integration into any kind of model. The regional maximum capacity potentials for wind and PV were taken from Risch et al.,^{49,50} as this is highly detailed and validated data for Germany that is freely available. The basis for geodata is the *Nomenclature des Unités territoriales statistiques* – NUTS, a classification of the European Union. Level 2 of this classification⁵¹ is used in this work. Electricity demand data in hourly resolution for the year 2050 for tertiary, household, transport and industry sectors were taken from the *Forschungsstelle für Energiewirtschaft e. V. (FfE)*.^{31–34} On one hand, this data was selected because it covers a broad range of future electricity demands including cooling and heating demands. On the other hand, the data was calculated using the weather year 2012 as a basis. Among the years considered in that study, 2012 was a year with an average amount of heating degree days, which fares well with our assumptions that storage levels at the first and last time steps need to be equal. The rare and extreme cold period with temperatures of -30°C occurring in early February of 2012 leads to a spike in heating demand making it ideal as a basis for highly robust ESMs. In combination with the dark lull in the weather data from 1994, this represents a severe, although short (< 4 days) and extremely rare event in Germany and a high level of robustness for the results is to be expected. According to the German weather service,⁵² the lowest temperature since the beginning of recordings is -37.8°C . The utilised data and the developed code for either of the following calculations are publicly available via GitHub.³⁶

CAESAR computing cluster specifications of utilised nodes for optimising the ESM for Germany

CPU	Intel(R) Xeon(R) Gold 6334
Cores per node	16
Threads per node	32
Threads used	up to 3
CPU max frequency	3.6 GHz
RAM	up to 50 GB (2 TB available)

QUANTIFICATION AND STATISTICAL ANALYSIS

There are no quantification or statistical analyses to include in this study.