Comparing expert elicitation and model-based probabilistic technology cost forecasts for the energy transition

Jing Meng[a,b], Rupert Way[c,d], Elena Verdolini[e,f], and Laura Diaz Anadon[b,g,1]

[a]The Bartlett School of Sustainable Construction, University College London, London WC1E 7HB, United Kingdom; [b]Cambridge Centre for Environment, Energy and Natural Resource Governance, Department of Land Economy, University of Cambridge, Cambridge CB3 9EP, United Kingdom; [c]Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, Oxford OX1 3UQ, United Kingdom; [d]Smith School of Enterprise and the Environment, University of Oxford, Oxford OX1 3QY, United Kingdom; [e]Department of Law, University of Brescia, 25121 Brescia, Italy; [f]RFF-CMCC European Institute of Economics and the Environment, Euro-Mediterranean Center on Climate Change, 73100 Lecce, Italy; and [g]Belfer Center for Science and International Affairs, Harvard Kennedy School, Harvard University, Cambridge, MA 02138

Edited by William C. Clark, Harvard University, Cambridge, MA, and approved May 7, 2021 (received for review October 5, 2019)

We conduct a systematic comparison of technology cost forecasts produced by expert elicitation methods and model-based methods. Our focus is on energy technologies due to their importance for energy and climate policy. We assess the performance of several forecasting methods by generating probabilistic technology cost forecasts rooted at various years in the past and then comparing these with observed costs in 2019. We do this for six technologies for which both observed and elicited data are available. The model-based methods use either deployment (Wright's law) or time (Moore's law) to forecast costs. We show that, overall, model-based forecasting methods outperformed elicitation methods. Their 2019 cost forecast ranges contained the observed values much more often than elicitations, and their forecast medians were closer to observed costs. However, all methods underestimated technological progress in almost all technologies, likely as a result of structural change across the energy sector due to widespread policies and social and market forces. We also produce forecasts of 2030 costs using the two types of methods for 10 energy technologies. We find that elicitations generally yield narrower uncertainty ranges than model-based methods. Model-based 2030 forecasts are lower for more modular technologies and higher for less modular ones. Future research should focus on further method development and validation to better reflect structural changes in the market and correlations across technologies.

expert elicitation | model-based technology forecasts | energy transition | energy technology costs | uncertainty

Designing robust and cost-effective policies and business plans to promote a carbon-neutral, sustainable economic system requires estimating the future cost of technologies that may play a significant role in the energy transition. The future of these and other technologies is notoriously hard to predict because the innovation process, that is, the "process by which technology is conceived, developed, codified, and deployed" (1), is part of a complex adaptive system (2) and is made up of interconnected actors and institutions (3, 4). Different frameworks for analyzing technology innovation and its determinants consider the innovation process at the level of nations (5), sectors (6), and technologies (2, 7, 8) and highlight different actors and relationships. And several relevant additional literatures contribute to a holistic understanding of different aspects of the innovation process. For example, the multi-level perspective is used to understand sociotechnical transitions from niches to regimes and shines a light on evolutionary, interpretive, and contextual processes and the role of agency (9). The economics of innovation literature focuses on mechanisms such as market demand (induced innovation) (10), knowledge spillovers from multiple areas of technology over time (11), and learning by doing during production or use (12). Finally, complexity economics emphasizes the organic and nonequilibrium nature of innovation (13). These approaches and literatures not only inform our understanding of the innovation process but also provide (sometimes implicitly and sometimes explicitly) a conceptual basis that guides how researchers, analysts, and policymakers forecast technological change.

A range of probabilistic forecasting methods have been developed and used to generate estimates of future technology costs. Two high-level types of approaches have been most often used to generate quantitative forecasts: expert-based and model-based approaches. Broadly speaking, expert-based approaches involve different ways of obtaining information from knowledgeable individuals who may have differing opinions and/or knowledge about the relative importance of various drivers of innovation and how they may evolve. Experts make implicit judgments about the underlying drivers of change when producing their forecasts and can take into account both public information about observed costs as well as information that may not yet be widely available or codified (14). Expert-based approaches are often the only source of information available to analysts when data on a given technology has not yet been collected—as is generally the case for emerging technologies. By contrast, model-based approaches explicitly use one

SUSTAINABILITY SCIENCE

ECONOMIC SCIENCES

Significance

Forecasting is essential to design efforts to address climate change. We conduct a systematic comparison of probabilistic technology cost forecasts produced by expert elicitation and model-based methods. We assess their performance by generating probabilistic cost forecasts of energy technologies rooted at various years in the past and then comparing these with observed costs in 2019. Model-based methods outperformed expert elicitations both in terms of capturing 2019 observed values and producing forecast medians that were closer to the observed values. However, all methods underestimated technological progress in almost all technologies. We also produce 2030 cost forecasts and find that elicitations generally yield narrower uncertainty ranges than model-based methods and that model-based forecasts are lower for more modular technologies.

Author contributions: E.V. and L.D.A. designed research; J.M., R.W., E.V., and L.D.A. performed research; J.M. and R.W. analyzed data; and J.M., R.W., E.V., and L.D.A. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

See online for related content such as Commentaries.

[1]To whom correspondence may be addressed. Email: lda24@cam.ac.uk.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1917165118/-/DCSupplemental.

Published June 28, 2021.

or more variables from available observed data to approximate the impact of the full set of drivers of innovation on technology costs, implicitly assuming that the rate of change in the past will be the best predictor of the rate of change in the future.

Expert- and model-based approaches can be used to produce deterministic or probabilistic cost forecasts. Notably, there are increasing calls to better understand and incorporate uncertainty in energy systems and policy modeling (15–20). Despite this, most integrated assessment models (IAMs) of economic activity and the climate (21–23), policy analyses, and industry perspectives rely on deterministic forecasts of energy technology costs or (in some cases) scenario analysis to inform public policy design or investment decisions. Even though expert- and model-based forecasting methods are increasingly used, we know very little about their relative performance when compared to observed costs.

This paper presents a systematic analysis of the relative performance of probabilistic cost forecasts from expert-based methods and model-based methods. We specifically focus on one expert-based method—expert elicitations (EEs)—and four model-based methods—two based on Wright's law (which model costs as a function of deployment) and two based on Moore's law (which model costs as a function of time). *SI Appendix*, Table S1 summarizes these methods and illustrates their links with the different innovation theories and concepts, including necessarily brief and stylized notes on the underlying intuition or mechanism behind each method.

We assess the performance of these forecasting methods by generating probabilistic technology cost forecasts rooted at various years in the past. We then compare these with observed costs in 2019 for six energy technologies for which both observed and elicited data are available. Although our analysis directly informs energy and climate technologies, which is an important area of research and policy, the general approach we develop and implement is applicable to other technology areas. We also compare probabilistic 2030 cost forecasts generated using these methods to each other for 10 technologies. The closest analysis to this part of our work is that by Neij (24), which compares experience curves (Wright's law), expert judgments, and "bottom-up analysis" for some technologies. We build on this analysis and significantly expand it by focusing on EEs (which were not yet available at the time) and including recent and more rigorous probabilistic model-based forecasting methods. Assessing the relative performance of expert-based and model-based forecasting methods in the energy space can help understand their suitability for important technology foresight tasks in research and practice, such as those needed in integrated assessment modeling and broader policy analysis.

To systematically compare the performance of expert elicitation forecasts and model-based forecasts using Wright's and Moore's laws, we undertake the following four analytical steps.

First, we collect, harmonize, and make available a large number of data points on the costs of 32 energy technologies relevant to support the energy transition. These data points include 25 sets of data from EEs conducted between 2007 and 2016 and cover a range of geographies (reference Dataset S1 for details) and 25 sets of observed technology data including the evolution of cost and deployment over different periods of time. This data collection effort was possible thanks to the growing literature on both observed data and, more markedly, energy technology EEs since 2007. *SI Appendix*, section 2 includes details on the observed data and indicates how to access the repository of data collected.

Second, we build on past work (25–29) to develop probabilistic forecasts using EEs and model-based approaches. We generate four different kinds of model-based forecast—two using Wright's law and two using Moore's law. For each of these two laws, we characterize uncertainty and project it forward using two different methods, which we call the Stochastic Shock method and the Stochastic Exponent method, respectively, as detailed below.

Third, we assess the performance of probabilistic forecasts made using these five methods from various dates in the past. We contrast their 2019 forecasts with observed costs for the subset of six energy technologies for which the necessary data are available. Given the probabilistic nature of these forecasts, we assess their performance by comparing both the median values of their 2019 forecasts and the forecast ranges with the average observed 2019 costs.

Fourth, we compare the five expert- and model-based probabilistic forecasts of technology costs in 2030—an important milestone in the energy transition (30)—to each other and reflect on their differences. This latter part of the analysis is possible for 10 energy technologies.

The urgency of developing policies for deep decarbonization (as outlined in the Intergovernmental Panel on Climate Change [IPCC] 1.5 °C report) (30) makes this systematic analysis timely and necessary. We conclude the paper with reflections about important avenues for future research on technology cost forecasting.

Technology Forecasting and Energy

Technological trajectories are complex and cannot be fully characterized by any single indicator (28, 31). Yet, for some specific purposes, knowledge of a few key technology characteristics can be extremely useful. In the context of climate change mitigation and energy policy, one of the most informative indicators is unit cost—the cost per unit of energy or functional capacity. Understanding the range of future costs of energy technologies is essential for the design of cost-effective and robust energy and decarbonization policies (20). As previously mentioned, the two most commonly used classes of methods to make technology cost forecasts in general—and in the energy sector in particular—are expert-based methods and model-based methods (16, 32–34).

Model-Based Methods. The use of model-based approaches to estimate future technology costs has a long history. The various approaches can be understood in terms of three elements: a) the "underlying model" of technological change—that is, the functional form or relationship between dependent and independent variables; b) the way uncertainty is characterized and used to project the underlying model forward in time to generate forecasts; and c) the way forecasting method performance is evaluated using forecast error models.

Functional forms. The first underlying model we consider describes the process of learning by doing and is known as Wright's law (i.e., learning curve or experience curve). Based on an empirical pattern originally identified in 1936 by Wright (35), it postulates that for a given technology, each doubling of experience is associated with a reduction of cost by some fixed, technology-specific percentage called the "learning rate" (36). "Experience" in this context means the sum total of humanity's experience with the technology in all aspects of its development, deployment, and use. It is often proxied by cumulative production or cumulative deployment. The intuition behind Wright's law model is that expending effort on a technology generally leads to updates and improvements. This pattern is observed across many technologies (37).

The second model we consider is known as Moore's law, since it was first observed in 1965 by Intel Corporation founder Gordon Moore (38). Emerging from the empirical observation that computer chips improved at roughly a constant rate, Moore's law (as it is currently interpreted) states that each year the cost of a given technology falls by some fixed, technology-specific percentage, which we call the "progress rate" here (i.e., costs decrease exponentially with time). This model represents the idea that as time passes, in most cases, technology improves and costs decrease.

Both Wright's law and Moore's law methods are widely used, despite their known limitations. These include the fact that they do not capture explicitly the various complex, interconnected innovation processes that ultimately result in technology cost reductions, including research and development, economies of scale, and knowledge spillovers, among others (39, 40). These model-based methods also rely on empirical relationships that capture correlations without implying causation (37, 41). For example, they are known to suffer from omitted variable bias since exogenous technological change from spillovers takes place alongside deployment, and learning by doing takes place alongside the passing of time. There is, of course, a dependency between time and deployment; in cases in which deployment increases exponentially over time, both relationships are to a large extent capturing the other (42). Nearly exponential deployment is very common for technologies, so it is usually hard to determine whether Wright or Moore's law is a better fit for observed data. Thus, for some technology areas, Wright's and Moore's laws yield very similar forecasts. However, given the history of energy policy and the strong role of governments, deployment rates often vary significantly over time, making Wright's law potentially more useful.

Several other model-based methods have been proposed and developed, though these are less well known and less used. They include the following: Goddard's law, which considers the rate at which costs change as a function of economies of scale (43); the Sinclair–Klepper–Cohen model, which considers costs as a joint function of deployment and scale (44); Nordhaus's model, which is a mixture of Moore's and Wright's laws (45); and the two-factor learning curve literature, which forecasts cost as a joint function of deployment and accumulated knowledge from research and development (46, 47) (reference *SI Appendix*, Table S1 for more detail). A few model-based studies have also included additional factors beyond learning by doing and research and development, for example, ref. 41.

Characterization of uncertainty. Model-based forecasting approaches also differ in how they characterize uncertainty. For a given functional form, the simplest way to generate a forecast is to estimate the value of the primary "technological change" parameter (e.g., the learning rate in Wright's law or progress rate in Moore's law). This is often done using a simple regression procedure applied to observed data for a given technology, with multifactor models having two or more such parameters. The relevant parameter is then used to project technological change forward in a deterministic manner, producing a point forecast for a given technology. In this case, the technological change parameter is assumed to be both known with certainty and constant throughout the technology's lifetime.

Probabilistic forecasts build upon this basic method by introducing uncertainty in different ways. In the context of technological change, uncertainty arises from three distinct sources: uncertainty because of measurement error (since we can never know the "true" value of any given parameter), uncertainty because of innovation being an intrinsically uncertain process (endogenous uncertainty), and uncertainty because of unforeseeable events elsewhere in the economy (exogenous uncertainty). These three sources can be incorporated into forecasts by, respectively, assuming that the technological change parameter is uncertain (e.g., ref. 48); assuming that, in addition, it may change over time (e.g., ref. 25); and by adding extra "noise" terms to the model, for example, periodic shocks representing unforeseeable fluctuations in the economy (e.g., refs. 27 and 49).

Once an uncertainty specification has been selected, probabilistic forecasts may be generated either analytically or by Monte Carlo methods (i.e., randomly generating a large number of deterministic forecasts and then aggregating these to form a probabilistic forecast). Typically, a central estimate of the technological change parameter is obtained, plus a few small variations around it. These variations can be informed by an error distribution produced by the regression procedure used to estimate the central parameter value and are used to create a small number of deterministic forecasts used for scenario analysis.

The literature estimating future energy technology costs in this way using Wright's law, Moore's law, or two-factor learning curves is vast. For example, Rubin et al. (47) review one-factor and some two-factor learning rates for 11 electricity generation technologies, Schmidt et al. (50) present experience curves for 11 electrical energy storage technologies, Weiss et al. (51) estimate average learning rates for 15 energy demand technologies and 13 energy supply technologies, and Malhotra and Schmidt (52) collect data on experience curves for 12 energy technologies. Many studies have focused on individual technologies, such as wind power (53, 54), solar power (25, 55), and electric vehicles (56). Wright's law has also been used to assess the costs and benefits of specific energy policies, including the cost effectiveness of a California residential solar photovoltaic (PV) subsidy (57), the expected costs of subsidizing solar or wind technologies until they are competitive with alternatives (25, 58), and the design of feed-in tariffs for renewable technologies (59). Another common use of projections based on Wright's law is in various IAMs, including many of those used in the IPCC reports and in other prominent analyses as a way of endogenizing technological change (60, 61). Cost reductions due to technological change have also been modeled as a function of time in IAMs (e.g., ref. 62).

Statistical testing of model-based methods. Understanding the reliability of each methods' forecasts is critical for several reasons. Natural variation in technologies' historical data series can lead to large uncertainty in estimated model parameters (63); statistical model identification is often difficult due to correlations between variables (45, 64), and using endogenous or exogenous technological change in energy models leads to very different results and policy implications (60, 65).

To shed light on these issues, a strategy of collecting data and systematically backtesting various forecasting models was proposed (66), and several studies advanced this approach (27–29). The underlying idea is that each technology's evolution over time is likely the result of sufficiently similar mechanisms that technologies can be treated as a set of nearly identical independent experiments whose data may therefore be pooled and analyzed to produce inferences about the future (i.e., out-of-sample forecasts).

The first such study used a set of 12 technologies (67) following which Nagy et al. (28) undertook a major effort to assemble data and conducted a more comprehensive analysis. They collected long, sequential time series data for 62 technologies and performed hindcasting experiments with several candidate models. This work led to a greater understanding of the role of error models in assessing forecast performance and to a focus on Wright's and Moore's laws in future forecasting efforts because of their superior performance in hindcasting tests. Farmer and Lafond (27) and Lafond et al. (29) extended this work by developing and testing one specific method for characterizing uncertainty, which we use in this paper and describe in more detail below. All the probabilistic forecasting studies doing statistical testing suggest that further development of these methods is both possible and desirable.

Forecasting Uncertainty Using Model-Based Methods in This Work. To make probabilistic cost forecasts, we use two functional forms—Wright's and Moore's laws—plus two methods of projecting uncertainty forward, yielding a total of four model-based methods (see *Methods* for more details). The first uncertainty projection method we use is the first-difference stochastic model developed in refs. 27 and 29, which we call the Stochastic Shock method. This model assumes there is a stable but uncertain technological change trend, on top of which periodic stochastic shocks occur, impacting costs repeatedly over time. The model is calibrated using differences between sequential observations in each technology's data series.

The second uncertainty projection method we use is a modification of the method proposed by ref. 25. The original method involves calculating the frequency with which learning rates (for Wright's law) or progress rates (for Moore's law) are observed in a given technology's data series, using a combinatorial approach, and then using the resulting distributions to project uncertainty forward. However, using all start/end year combinations to calculate these distributions pools together observations that are not entirely comparable since ranges overlap and vary in length. To avoid potential biases related to double counting, we use only single-period learning and progress rate observations; we call this modified method the Stochastic Exponent method. We fit normal distributions to the observed progress exponent distributions and generate forecast distributions in a Monte Carlo method style. Each single forecast is made by sequentially picking progress exponents from the relevant normal distribution and projecting costs forward, year by year, until the time horizon is reached; then, these are aggregated to form the final forecast distribution.

Expert-Based Methods. Expert-based approaches, which include EEs, are a different class of forecasting methods (*SI Appendix*, Table S1). They can be used to offer insights into the future of technologies for which historical deployment and cost data are not yet available or commercialized (14, 68), are designed to always provide probabilistic estimates of future cost developments (69), and do not assume that previous trajectories will continue and do not preclude the identification of technology surprises or discontinuities.

Among the expert-based approaches, EEs have been the most widely used method to estimate probabilistic future technology costs in the energy sector. Because of this relative richness of data, we focus our analysis of expert-based methods on EEs. EEs are structured surveys of experts who are asked to provide probabilistic estimates of future costs by using the best available information to them at the time of the elicitation (16, 32), which is likely to include different types of observed data. It is important to note that in some cases, this same observed data may serve as input to some of the model-based methods.

The EE method was pioneered in the 1960s and 1970s to support decision-making in the presence of extreme or unlikely events (70). EEs have often been applied to forecast technology cost and performance to support decision-making under uncertainty (16). In the past decade, an increasing number of expert elicitation studies were conducted on energy technologies (14) including (among others) nuclear power (71), wind energy (32), solar energy (72), water electrolysis (73), and energy storage (74).

EEs are of particular interest when data are sparse or missing or the technologies are emerging. However, they are time consuming and may differ in terms of the elicitation protocol and methods for administering the survey (Dataset S1). Research has suggested that in-person interviews have the advantage of allowing experts and researchers to have in-depth discussions and ask clarification questions but may be subject to small sample sizes and cognitive biases (72). Other research has shown that although online instruments are often pilot tested and can enable a broader representation of experts, which can reduce biases (33), challenges can remain since questions cannot be tailored in real time to adapt to respondent preferences and understanding. Some effort has also been devoted to understanding quantitatively the EE design factors that lead to systematic differences in uncertainty estimates from energy technology expert elicitations. For example, ref. 33 explores the role of elicitation mode, expert characteristics, and research and development (R&D) scenario variables in determining the uncertainty range surrounding energy technology cost forecasts using the largest set of EE studies available as of 2015. In the same vein, other work has investigated the extent to which elicitations focused on aggregate- versus component-level costs are associated with differences in elicited cost forecasts (75, 76). Another important consideration related to EE design is that, given that experts are busy and have limited time available, they are necessarily asked to provide estimates about a limited number of points in time. This means that using EE forecasts for specific policy or investment decisions and comparisons (such as the ones in this paper) often involves interpolating or extrapolating elicitation data over time.

Beyond questions about elicitation design, using EE outputs often requires making difficult decisions about whether or not to aggregate expert elicitations and, if so, how. While offering insights that are more easily interpreted by policymakers and other analysts, aggregating across experts presents limitations and drawbacks that must be considered when it comes to the communication of results as discussed in refs. 76 and 14. The aggregation, interpolation, and extrapolation methods used in this work are discussed in *Methods*.

Despite the significant and growing research on model-based and expert elicitation forecasts in energy technologies and other technological areas, we are not aware of any research collecting and systematically comparing the performance of those types of approaches to observed values or each other in the future.

Results

Data Collection and Harmonization. The first step and contribution of our analysis is a large collection of data on energy technology costs, which we make available as detailed in *SI Appendix*, section 3. These include 25 sets of data from EEs on energy technologies conducted between 2007 and 2016 and covering a range of geographies and 25 sets of observed energy technology data including cost, deployment, and time. Dataset S1 lists all the EE data sources by technology, including information about how the data points were described in the original sources and the cost metrics available for each technology. Dataset S2 lists all the references and links for both the EE and observed data. Note that due to the need for both elicitation and sufficient prior observed data on the same technology, forecast comparison analyses were only possible for much smaller subsets of technologies than those in the complete data set.

Probabilistic Forecasts of Energy Technology Costs. The second step of our analysis is to generate probability distribution functions of estimated costs from model-based and expert-based methods as detailed in *Methods*. This allows us to compare the probabilistic 2019 cost forecasts from both methods to observed average 2019 costs as well as to compare the 2030 probabilistic cost forecasts generated using the different methods to each other. Here, we highlight a few key aspects of the forecast generation process.

To build the expert elicitation forecasts for the comparison we a) fit one of three types of continuous probability distributions to the discontinuous probabilistic data points provided by each expert, b) aggregate the resulting continuous individual expert distributions for each technology assigning equal weights to each expert using the method in ref. 26, and c) when the time points provided by the experts did not coincide with our 2019 and 2030 forecasting horizons, we use an exponential fit between the two time horizons provided by the experts to either interpolate or extrapolate as required (*SI Appendix*, section 3).

To build the model-based forecasts, we generate probabilistic estimates from Wright's law using the Stochastic Shock (W1) and the Stochastic Exponent (W2) methods and from Moore's law using the Stochastic Shock (M1) and the Stochastic Exponent (M2) methods (Table 1). Three important features of the model-based forecasting process should be mentioned.

First, to apply Wright's law model, one needs to make assumptions about the future level of deployment. In this work, we assume recent historical growth trends continue into the future and persist for the duration of the forecasts: we extrapolate deployment using the compound average annual growth rate (CAAGR) observed over the 10-y period prior to each forecast. The W1 and W2 forecasts are therefore conditional upon these specified levels of deployment. We refer to this as a "continuation of past trends deployment scenario," which is tantamount to saying that deployment, R&D funding, and other variables continue the historical trajectory of the previous 10 y. We depart from this general rule in the case of two technologies because the calculated average rate of deployment over the 10-y time period may be too high compared to even the most ambitious deployment scenarios for those two technologies (see *Methods*).

Second, in the computation of the 2019 model-based forecasts, we only use observed data up to the year of the elicitation with which the forecasts are to be compared. In other words, if the elicitation for a given technology was conducted in (let us say) 2010, we produce the model-based forecast of 2019 costs using observed data up to 2010 only, even if observed data are available beyond 2010. This allows us to compare the EEs with model-based estimates that do not enjoy the benefit of additional years of data that were not available to the experts when they made their forecasts.

And third, although in many cases we found that relevant observed data were not available to generate model-based forecasts, in other cases we had to make choices about which data to include. The details on the data selection for nuclear power, offshore wind, and two electrolysis technologies can be found in *SI Appendix*, section 3.

Comparison of the Methods' Performance Forecasting 2019 Costs. Fig. 1 presents the systematic comparison of expert-based and model-based probabilistic forecasts with observed 2019 costs. The figure shows average 2019 observed costs and the probabilistic 2019 EE, W1, M1, W2, and M2 forecasts for the six energy technologies for which both elicitation and sufficient prior observed data are available. The horizontal axis shows the cost and the cost units, and the different colors on the vertical axis denote the forecast method. The whiskers represent the fifth to 95th percentile range, the short sides of the rectangular box the 25th to 75th percentile range, and the line dividing the rectangular box the 50th percentile (the median). The corresponding average observed 2019 cost is shown with a dashed gray vertical line in each of the six plots.

From left to right and from top to bottom, the technology forecasts are listed starting with the oldest forecast. The oldest forecasts are for nuclear power, which use an expert elicitation from 2009 and, correspondingly, observed data up to 2009 to produce the model-based forecasts. The most recent forecasts are for alkaline electrolysis cells (AEC) and proton exchange membrane (PEM) electrolysis cells, which are based on an expert elicitation from, and observed data up to, 2016.

Fig. 2 shows the log of the ratio of the 50th percentile of the cost forecast and the observed cost value. This is a dimensionless metric for characterizing the performance of the different methods, which allows us to systematically compare the size of the uncertainty ranges and the distance between the forecast medians and the observed average values. The line encompasses the log of the ratio of the 95th and fifth forecast percentiles and the observed value.

There are three main takeaways from the performance comparison shown in Figs. 1 and 2.

First, we find that, for this set of energy technologies, model-based forecasts outperform EE-based forecasts in terms of their ability to capture the 2019 observed value within the forecasted 2019 uncertainty range. Concretely, almost all model-based approaches produced fifth to 95th percentile ranges for 2019 that included the observed values for all six technologies. The only exception was the M2 method on solar PV, which failed to include the observed value, although only by very little (Fig. 2). Notably, the Stochastic Shock method (W1, M1) produced 25th to 75th percentile ranges that captured the average 2019 value in all technologies but PV. In contrast, the fifth to 95th percentile range of the EE 2019 forecasts only contained the observed cost value for the two wind technologies (i.e., onshore and offshore wind). Model-based methods also outperformed elicitations in that they produced forecast medians that were closer to the observed 2019 values. Specifically, model-based medians were closer to the observed costs than the EE medians for all technologies except offshore wind and the W2 model-based forecast for PEM electrolysis.

Second, in almost all cases, both the EE and the model-based forecast medians are higher than the observed 2019 costs—they underestimated technological progress over this forecasting period. The notable exception is nuclear power, in which the finding is reversed—all forecast medians are lower than the observed cost. The W2 forecast medians for the two electrolysis technologies are outliers; reference *SI Appendix*, section 4 for a detailed discussion of the uncertainty ranges generated by W2 and M2 in this case.

Third, for four out of the six technologies, the fifth to 95th percentile range of the 2019 forecast was larger for the model-based methods than for EEs (Figs. 1 and 2). This was not the case for solar PV and onshore wind. Note that the fifth to 95th percentile ranges for EE and model-based forecasts for onshore wind were very close. The fact that model-based forecasts tend to generate fifth to 95th percentile uncertainty ranges that are generally larger or equal to those of EEs may partly explain the first finding discussed above. It is important to point out that the EEs all included steps in the elicitation protocol aimed at reducing expert overconfidence.

Determining why all methods underestimated technological progress in all technologies except nuclear when generating 2019 forecast medians is hampered by the fact that our analysis was possible only for a small number of technologies. However, we can hypothesize a few possibilities. These include unforeseen changes in the industry structure and/or policy and/or an underestimation of the level of deployment. The five energy technologies for which forecast medians were higher than observed 2019 costs are likely correlated, that is, subject to a common set

Table 1. Overview of the five forecasting methods in this study

Forecasting approach	Method name	Summary: Explanatory variable	Summary: characterization of uncertainty
Model based	Wright 1 (W1) Moore 1 (M1)	Cumulative experience Time	Stochastic Shocks: an uncertain yet stable progress trend (Wright or Moore) is estimated, in addition to which random shocks occur each period; shocks accumulate over time.
	Wright 2 (W2) Moore 2 (M2)	Cumulative experience Time	Stochastic Exponents: each period a new progress exponent (Wright or Moore) is randomly selected from a fixed probability distribution.
Expert based	EE	Several experts estimate cost distributions at (at least) two different points in time. Individual expert distributions are aggregated to form a single distribution at each date, and a smooth exponential trend is assumed to connect the two.	

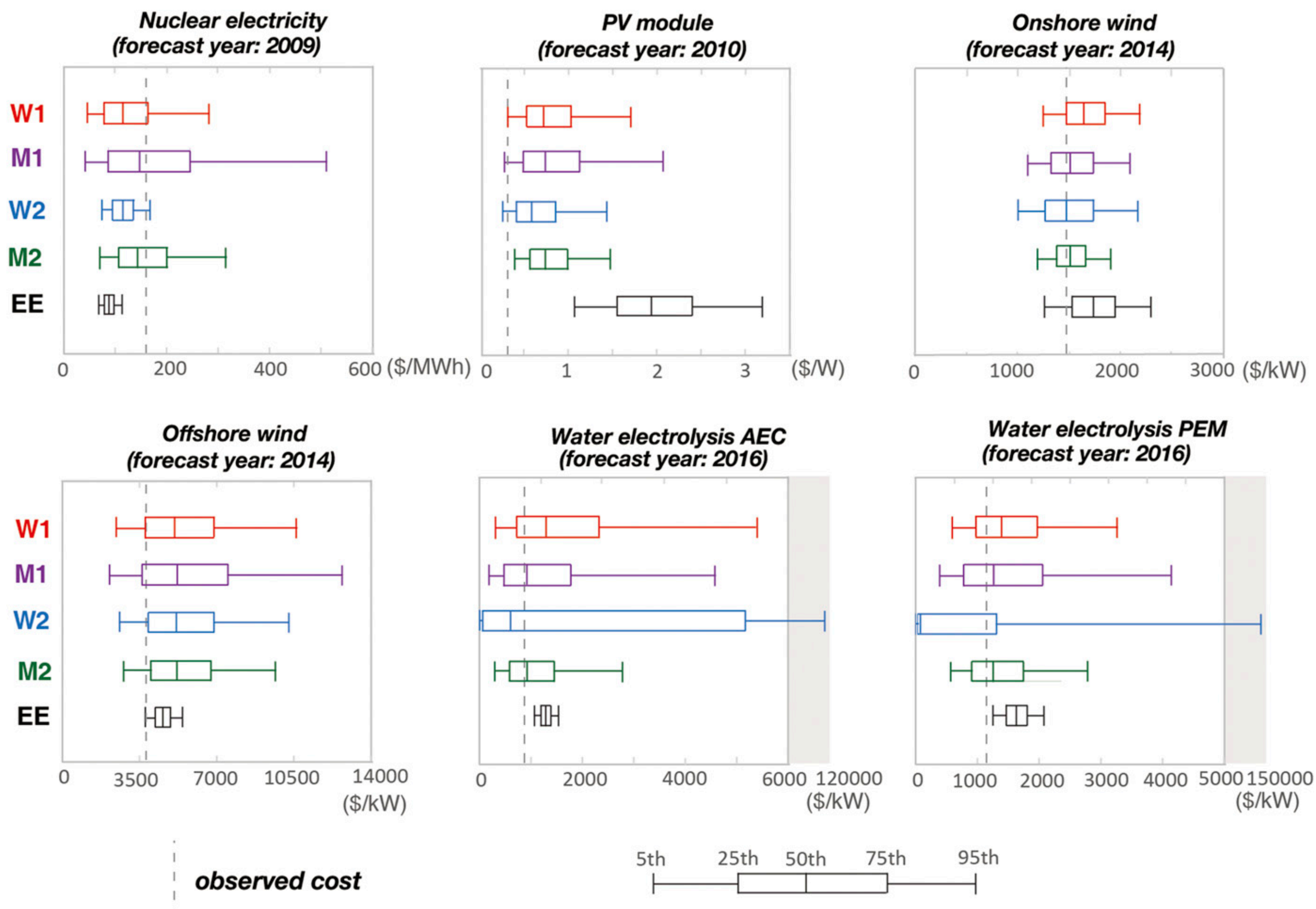
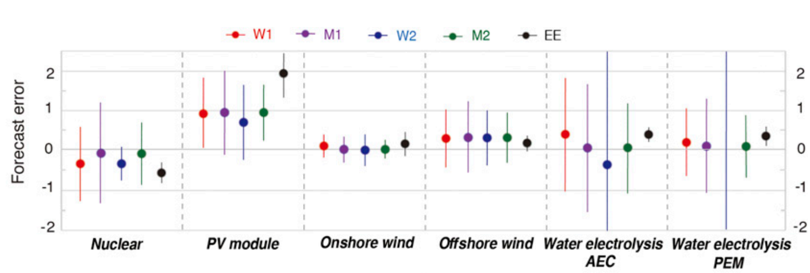


Fig. 1. Comparison of the expert elicitation and model-based 2019 forecasts for six energy technologies with the corresponding average 2019 realized values. The year listed in brackets below the name of the technology indicates the year in which the EE was conducted and, consequently, the latest observed data included as input to the model-based forecasts. The far-right whisker (the 95th percentile) of the W2 distribution for water electrolysis AEC and PEM is shown inside a gray band; this indicates that the *x*-axis was extended to include said forecast.

of underlying drivers related to decarbonization policies, investor perceptions and preferences, firm expectations, and societal pressures. It is therefore possible that the forecasting period used for our systematic comparison was one in which a range of underlying drivers of innovation changed to result in faster innovation across those five technologies. Along similar lines, for elicitations, under- or overestimation of technological change can occur if experts cannot foresee increases or decreases in policy support, deployment, or regulation for this particular set of technologies. As previous research suggests, it is also not surprising to see nuclear



Forecast year	2009	2010	2014	2014	2016	2016
Forecast horizon	10 years	9 years	5 years	5 years	3 years	3 years
Observed data	1971-2009	1975-2010	1983-2014	1983-2014	2003-2016	2003-2016
Number of observations	23	35	31	29	9	7

Fig. 2. Forecast errors of the five methods in 2019. The forecast error is a dimensionless performance metric. The dots represent the log of the ratio of the forecasted 50th percentile and the corresponding observed average 2019 value. The top of each line represents the log of the ratio of the 95th percentile with the realized value, while the bottom represents the log of the ratio of the fifth percentile with the realized value; the lines span the distances between these two forecast error extremities. The error of the forecast median for PEM using method W2 is −3.1, which is excluded for readability. Forecast horizon means the time horizon from the elicited year to 2019.

evolve differently (52, 77, 78) from other technologies. In other words, this faster pace of innovation in most of the energy technologies covered compared to the forecasts is likely the result of structural change across the energy sector due to widespread policies and social and market forces.

Importantly, our results do not necessarily imply that EE forecasts will be less accurate than model-based forecasts for all technologies. Indeed, *SI Appendix*, section 5 and Fig. S3) shows three additional EE technology cost forecasts (bioelectricity, lithium-ion batteries, and biodiesel) whose fifth to 95th percentile ranges included the 2019 observed values. These are not shown in the main body of the paper because no observed data were available to generate comparable model-based forecasts. It is of course possible that, had model-based forecasts been available, they would have been more accurate than these EE forecasts, or vice versa. In *SI Appendix*, section 6, we include an additional robustness check with different observed data inputs for the model-based forecasts for offshore wind—not including experience from onshore wind. The results are consistent with those shown in Figs. 1 and 2.

This comparison of different methods' performance represents only a first step, yet an important one. As more data becomes available, future research could further test whether model-based approaches outperform EEs in other technologies building on the approach presented here. We note that the ability to compare the performance of forecasts from different methods is directly determined by data collection efforts, pointing to the need for continued data collection to enable comparisons for a larger group of technologies. Two other important subjects of future research are, first, how to develop model-based forecasts for technologies that may be correlated due to structural changes in the underlying markets, and second, how to further reduce overconfidence in EEs.

Model-Based and Elicitation Forecasts of 2030 Energy Technology Costs. We also compare probabilistic EE and model-based 2030 cost forecasts with each other (as 2030 costs are not yet known) for the 10 energy technologies for which the necessary data are available. Fig. 3 shows the median and the fifth to 95th percentile ranges of the W1, M1, and EE forecasts to the year 2030 for onshore wind, offshore wind, crystalline silicon PV modules, thin-film PV modules, concentrating solar power (CSP), all PV modules, bioelectricity, nuclear power, AEC, and PEM electrolysis cells. 2030 forecasts using W2 and M2 forecasts are not presented here to increase readability, but they are shown in *SI Appendix*, section 7 and Fig. S5.

To generate the model-based forecasts for 2030, we rely on all the data available as of late 2020—that is, we do not shorten the time series to match the year of the expert elicitation as we did in Figs. 1 and 2. The objective of the comparison to 2030 is to summarize what is known today about the possible range of 2030 costs and to identify differences between EE and model-based forecasts going forward. Fig. 3 includes two different estimates for crystalline silicon (c-Si) PV costs, one relying on the full set of observed data for c-Si PV (Fig. 3*C*) and one relying on observed data since 2006 only (Fig. 3*D*, *Left*). We do this because the data available for thin-film PV covers only the period 2006 to 2019, so the most meaningful comparison with c-Si may be one that uses data beginning in 2006 also. This illustrates the importance and nuance of data choices underpinning the model-based forecasts. And as previously noted, *SI Appendix*, section 3 contains details regarding the data used as input for model-based methods in the case of offshore wind. We also conducted a sensitivity analysis with two additional deployment scenarios: one in which we use a deployment scenario consistent with the International Energy Agency's (IEA) Stated Policies Scenario and one that is consistent with the IEA's Sustainable Development Scenario (reference *SI Appendix*, section 8 for more information). Fig. 4 compares 2030 cost estimates of those subtechnologies that are more mature and established (i.e., "dominant") with those that are typically emerging, or with a much lower market share (i.e., "novel").

Three main insights emerge.

First, the uncertainty ranges (i.e., the 5th to 95th percentile range) for the 2030 forecasts generated using W1 and M1 are generally larger than those for the EE forecasts. This is similar to what was observed in Figs. 1 and 2 for the 2019 forecasts. The results for W2 and M2 are largely consistent (*SI Appendix*, section 8 and Fig. S5 and related discussion). While we cannot yet determine the accuracy of these forecasts when compared to observed 2030 costs, the smaller 2030 EE forecast ranges may be less likely to include them. Given the focus in the expert elicitation literature on addressing overconfidence, the fact that EE forecasts generally have smaller uncertainty ranges compared to those from model-based methods again suggests that additional research to reduce overconfidence would be useful.

Second, for nine out of the 10 technologies, the model-based 2030 cost forecasts have lower medians than the EE forecasts; nuclear power is the one exception in which EE forecasts are lower than model-based forecasts. For all technologies except bioelectricity, the EE forecast medians in 2030 are lower than the observed cost of the technology in the year when the EE was carried out (listed in the gray bands in Fig. 3), reflecting a general expectation from experts of cost reductions, although with substantial differences across technologies. The W1 and M1 forecast medians in 2030 are lower than the 2019 average costs for all technologies except nuclear power (and, in the case of W1, AEC also, though the difference is small). Reference *SI Appendix*, section 8 for more discussion on the uncertainty ranges generated using W2 and M2. Importantly, the W1 (and W2) findings presented here are robust to the different deployment assumptions used in the sensitivity analysis as discussed in *SI Appendix*, section 8.

Third, as illustrated in Fig. 4, "novel" subtechnologies generally have higher 2030 EE forecast medians and are characterized by larger uncertainty ranges than the EE forecasts for the corresponding dominant technologies. The comparison across all subtechnologies using all five forecasting methods was possible only for wind power, for which the comparatively novel subtechnology (offshore) has achieved significant diffusion already. This fact points to the particular value of EEs as a source of information to get an initial understanding of the future costs of technologies for which cost data are not available. Only as time progresses will we get a better sense of whether or not these comparatively novel technologies will "catch up" with their currently dominant technology counterparts by 2030 in terms of the costs. See more discussion in *SI Appendix*, section 9.

The results discussed suggest that EE 2030 forecasts should not be used for those technologies that have experienced significant cost reductions between the time of the expert elicitation and 2019. For several of the technologies covered in Fig. 3 (specifically Fig. 3 *A*, *C*, *D*, and *F*), the 2019 observed costs are already lower than the 2030 EE median forecasts. In these cases, model-based forecasts, which reflect information from the last 6 to 10 y, are preferable.

One possible explanation for the fact that most model-based forecasts were lower than EE forecasts emerges from the literature on the role of technology modularity or complexity (52, 78, 79) as determinants of technology innovation trajectories. Large-scale nuclear power plants and (to some extent) bioelectricity and CSP plants would fall into the category of less modular technologies compared to solar panels and batteries, for instance. Model-based 2030 cost forecast medians are lower than those from elicitations for more "modular" technologies and higher for the least modular technology (nuclear). They are closer to each other for the two technologies that are arguably in a midrange of modularity (CSP and bioelectricity). However, statistically testing this hypothesis is not within the scope of this paper. The fact that experts

SUSTAINABILITY SCIENCE

ECONOMIC SCIENCES

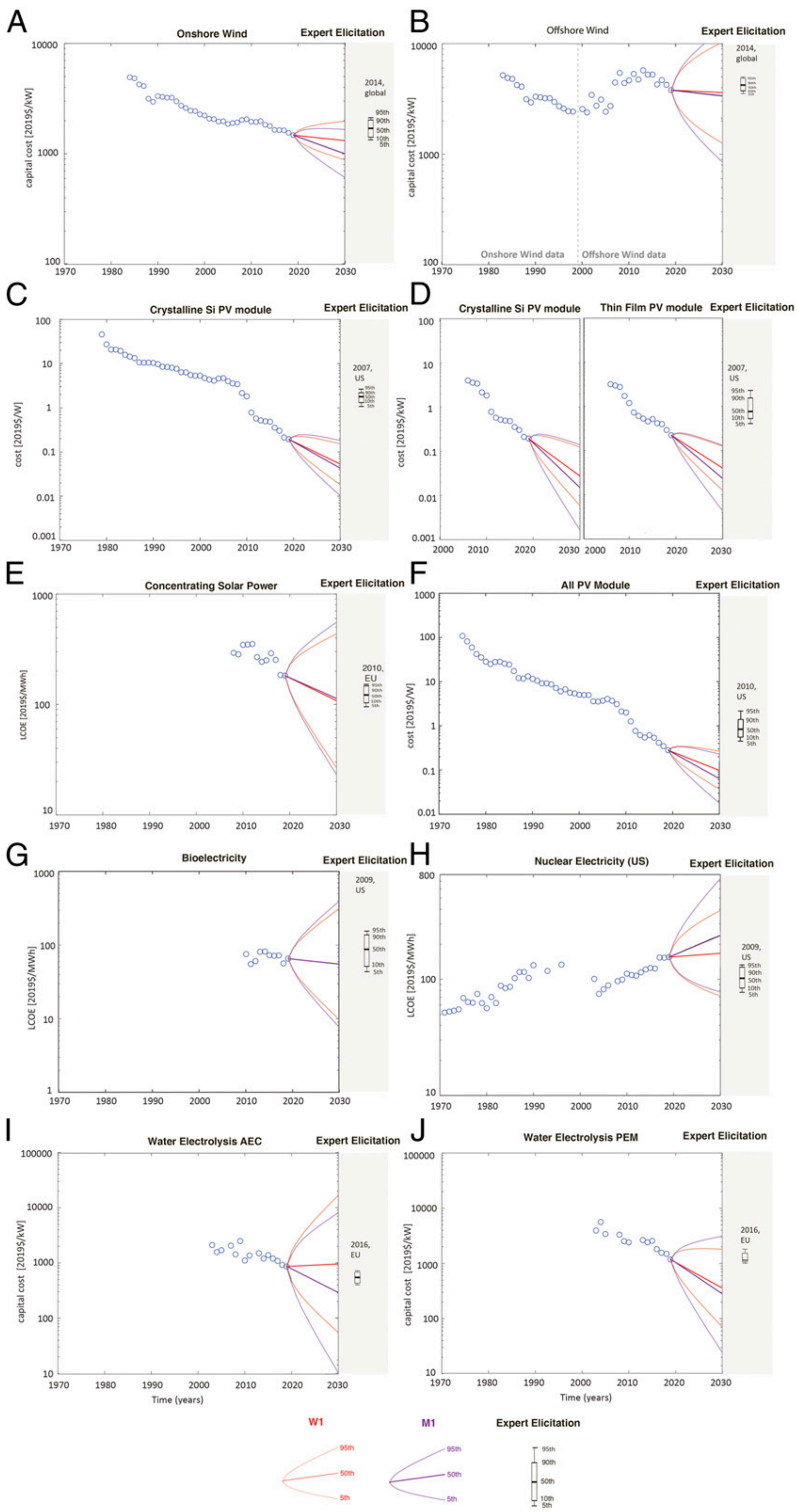


Fig. 3. Comparison of probabilistic 2030 cost forecasts using EEs and model-based methods. For each of the 10 technologies, (*A*) onshore wind, (*B*) offshore wind, (*C*) crystalline Si PV, (*D*) crystalline Si PV and thin-film PV, (*E*) concentrating solar power, (*F*) all PV module, (*G*) bioelectricity, (*H*) nuclear electricity, (*I*) water electrolysis AEC, and (*J*) water electrolysis PEM, the lines from 2019 to 2030 show the 5th, 50th, and 95th percentile forecast using the W1 method (the red and orange lines) and the M1 method (the purple and light purple lines), with the underlying observed data used to make them shown in blue circles. For each of the 10 technologies, the gray band on the right-hand side shows the EE forecast and the year in which the EE was conducted. The box with black borders with whiskers in this gray area indicates the 5th, 10th, 50th, 90th, and 95th percentiles (from the bottom to the top). The data sources for all forecasts are included in *SI Appendix*, section 3. For nuclear, the elicited data were overnight capital cost (Dataset S1); this was first converted into levelized capital cost, then augmented with operations and maintenance cost data in order to produce meaningful comparisons with the model-based forecasts (which rely on observed levelized cost of electricity data). Nuclear power here includes both light water reactors and Gen IV designs. We include two different c-SI PV forecasts. *C* uses the full observed time series, and *D* uses a time series that matches the length of the thin-film observed data. For more information on the sources of data, reference Dataset S2.

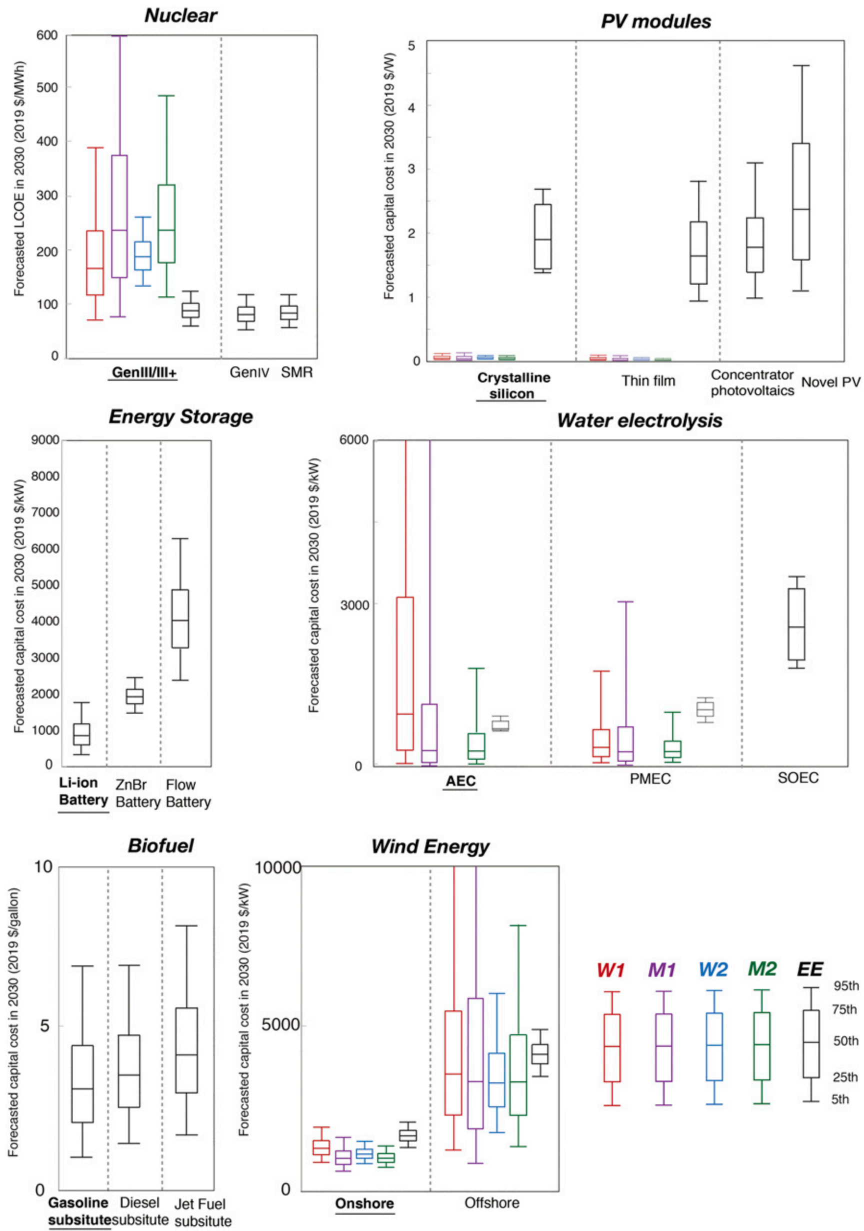


Fig. 4. Comparison of the forecasted 2030 costs of "dominant" and "novel" technologies for different technology classes using expert elicitation forecasts and model-based forecasts. The 2030 probabilistic forecasts using model-based methods rely on all the data available (not just up to the year of the elicitation). (*Bottom*, *Right*) Indicates the colors that correspond to the five different forecasting methods.

expect faster cost reductions for less modular technologies compared to model-based methods may also be due to the following: additional information they may possess about scientific breakthroughs, the industry, and/or policies; unawareness about the correlations between modularity and technological change; and/or bias and overconfidence. This indicates the need for future work to further test this finding using different metrics to account for modularity and/or complexity (e.g., ref. 79) as more data becomes available.

Discussion, Areas for Future Research, and Policy Implications

The increased availability of information on future energy technology costs, both in terms of forecasts made using observed data and data collected through EEs, provides a unique opportunity to introduce considerations on uncertainty around technical change into energy and climate change mitigation policies. It also enables us to conduct a systematic comparison of the relative performance of probabilistic technology cost forecasts generated by different expert-based and model-based methodologies with 2019 observed costs. Such a comparison is essential to ensure researchers and analysts have empirically grounded evidence to support assumptions in IAMs, cost benefit analyses, and broader policy design efforts. Undertaking this type of comparison to assess the performance of different forecasting methods should become much more common among modelers and forecasting practitioners. We also compare different longer-term (2030) model-based and expert-based forecasts with each other.

To perform this analysis, we undertook a major effort to collect up-to-date data on observed energy technology costs and elicited data on the future costs of energy technologies. We used these data to generate comparable forecasts from EEs and four types of model-based forecasting methods. We make this dataset available to modelers, researchers, and policy makers as detailed in *SI Appendix*, section 3.

Four key findings emerge from our research.

First, the comparison of EE and model-based forecasts with observed 2019 costs over a short time frame (a maximum of 10 y) shows that model-based approaches outperformed EEs. More specifically, the fifth to 95th percentile range of the four model-based approaches were much more likely to contain the observed value than that of EE forecasts. Among the model-based methods, the ones that more often captured 2019 observed costs were those using the Stochastic Shock method for characterizing uncertainty, with both Wright's and Moore's laws. We also find that the 2019 medians of model-based forecasts were closer to the average observed 2019 cost for five out of the six technologies. This comparison was possible only for a small number of technologies; furthermore, some of the EE forecasts included the observed value. For these reasons, additional research is needed to further validate these results on model-based methods outperforming EEs.

Second, both the EE and model-based methods underestimated technological progress in most of the energy technologies analyzed in this paper. For five of the six technologies, the methods produced 2019 cost forecast medians that were higher than the observed 2019 costs. This could be because these specific energy technologies are likely correlated. Indeed, it can be argued that in the period between 2009 and 2019, the energy sector underwent a structural change because of widespread policies and social and market forces common to all technologies (with nuclear being an exception). Given that our analysis is focused on this specific period and on correlated energy technologies, the extent to which this faster pace of progress compared to forecasts will continue (or not) in the future remains to be seen.

Third, in the majority of cases, EEs yielded fifth to 95th percentile uncertainty ranges that were significantly smaller than the uncertainty ranges produced with model-based methods, for both the short-term (2019) and longer-term (2030) forecasts. This can perhaps be attributed to the documented issue of expert overconfidence (80). Combined with the finding that EE forecasts were less likely to contain the average observed 2019 costs when compared to model-based forecasts, this result has at least two implications: a) additional research investigating how to reduce expert overconfidence would be useful, and b) when sufficient and reliable observed data are available, there would have to be very compelling reasons to select the EE method to generate forecasts. For emerging technologies, there is often little observed data that can be used to generate model-based forecasts, which means that expert-based methods may be the only option. We also find that for nine of 10 technologies, the medians of the 2030 EE forecasts are higher than those generated by the model-based methods. As suggested by recent literature, this could be due to a possible empirical relationship between technology characteristics (i.e., higher modularity) and (faster) innovation.

Fourth, our analysis highlights the value of testing different methods systematically to improve our understanding of the future of energy systems. Despite the relatively short forecasting horizon, the fifth to 95th ranges for all model-based PV module forecasts barely captured the 2019 realized values. This notwithstanding the fact that about four decades of observed costs of PV modules is available.

Taken together, these insights point to various worthwhile avenues for future research. With respect to EEs, previous research has shown that involving experts with diverse backgrounds and experiences and using different elicitation methods (i.e., face to face, mail, online survey) with carefully designed protocols matching the method (81) can help reduce overconfidence. This paper highlights the need to continue methodological improvements to reduce overconfidence. In addition, given the fast cost reductions in many important energy technologies, relying on older expert elicitation data would ignore relevant information. More broadly, and as previously mentioned, the paper raises questions about the value of conducting elicitations when reliable observed data are available to generate model-based forecasts when there are no specific reasons to do so.

With respect to model-based methods, this work highlights the challenge of finding (and collecting) data for many key energy technologies. It also calls for increased efforts in data collection and publication by international organizations and other entities. In addition, the underestimation of technological progress points to the value of further method development to reflect structural changes and technology correlations.

Lastly, given the large uncertainty ranges and major policy decisions associated with the energy transition and with addressing climate change, we hope this paper will stimulate much more research in this area. As more data becomes available and more time passes, additional research comparing the performance of different probabilistic forecasting approaches with observed values across a wider range of technologies building on the approach proposed in this analysis will be possible and valuable.

Methods

Data Collection. To produce probabilistic technology cost forecasts, we collected two types of data. First, we collected data on the evolution of cost and performance of energy technologies over time along with installed capacity or cumulative production. These are required to generate the model-based forecasts of future costs. We obtained this data from a range of databases (e.g., International Renewable Energy Agency [IRENA]), research articles [e.g., Nagy et al. (28)], and research organization reports [e.g., Fraunhofer (82) and Lawrence Berkeley National Laboratory (83)]. Second, we collected forecasts of future costs and performance (typically around 2030) using EEs from academic publications e.g., Verdolini et al. (14)]. Dataset S2 includes an overview of the technologies for which observed data were available (a total of 32 observed datasets in energy technologies), the time period covered, and the source.

Producing Probabilistic Forecasts. We first describe the two underlying technological change models used in this paper, and then we describe the two uncertainty characterization methods and then finally, the expert elicitation forecast generation method.

Wright's and Moore's Laws and Energy Technologies. Wright's law postulates that cost decreases at a rate that depends on the cumulative production as described by the following:

$$y_t = AX_t^{-w}, \quad [1]$$

in which y_t is the technology cost in year t, A is a constant, X_t is cumulative production, and w is the Wright exponent (or learning exponent). This exponent is then used to calculate the "learning rate," defined as $r = 1 - 2^{-w}$. This rate is interpreted as the percentage reduction in costs associated with each doubling of cumulative production.

Moore's law describes the exponential decrease in cost y_t of a technology as a function of time according to the following:

$$y_t = Be^{-\mu t}, \quad [2]$$

in which B is a constant, t is the year, and μ is the Moore exponent (or progress rate).

Developing Model-Based Forecasts with the Stochastic Exponent Method. This method uses the difference equation form of Wright's law (where t stands for time in years):

$$y_{t+1} = y_t\left(\frac{X_{t+1}}{X_t}\right)^{-W_{t+1}} \quad [3]$$

and of Moore's law:

$$y_{t+1} = y_t e^{-\mu t_{t+1}}. \quad [4]$$

To implement the Stochastic Exponent method, we first calculate all inter-annual exponents observed in the data for a particular technology (Wright exponents for W2 and Moore exponents for M2). For example, if we have 11 y of observed data for a technology, we can generate 10 exponents for both Wright's and Moore's law. Then, for each law, we fit a normal distribution to the sample of exponents obtained to create a Wright exponent distribution and Moore exponent distribution. These calculated exponent distributions are then used to generate cost forecast distributions for the technology by simulating large numbers of cost sample paths (we used 10,000 per forecast). Each sample path is generated by sequentially picking exponents from the corresponding exponent distribution for each year of the forecast and applying either the Wright's law or Moore's law difference equation, as shown above, until the forecast horizon is reached. The sample paths are then aggregated to approximate a single probabilistic forecast in the target year.

For the application of W2, it is necessary to assume a deployment scenario. In most cases, we used the CAAGR of cumulative experience (either installed capacity or electricity generation) observed over the 10 y before the forecast year to specify a future deployment scenario. There are two exceptions though: the PV and wind 2030 forecasts; in these cases, we used the CAAGR observed over the most recent 5 y instead (*SI Appendix*, section 8). This method implicitly assumes that deployment, R&D funding, and other variables continue on their recent historical trajectories for the entire duration of the forecasting period. We conducted a sensitivity analysis for the Wright's law forecasts using two additional deployment scenarios as detailed in *SI Appendix*, section 8.

To generate the model-based forecasts for the 2019 comparisons, we only use observed data up to the year of the elicitation with which the model-based forecast is to be compared, but for the 2030 forecasts, we use all available data.

Importantly, the model-based forecasts using both the Stochastic Shock and Stochastic Exponent methods can be produced only in those cases for which we have at least ~6 y of sequential annual technology data. This is because these methods infer cost trends from interannual cost differences, and at least around five samples are required for model calibration.

Developing Model-Based Forecasts with the Stochastic Shock Method. The Stochastic Shock method was developed and statistically tested by Farmer and Lafond (27) and by Lafond et al. (29) for Moore's and Wright's laws, respectively. The model represents the idea that in each year, technological progress occurs according to a stable underlying trend (either Moore's or Wright's law), and, in addition to this cost development, there is an exogenous shock that also impacts the cost. The magnitude of both the underlying trend and the stochastic shocks are specific to each technology and are determined by calibration using observed data. The periodic random shocks accumulate over time, giving rise to a probability distribution of forecast costs that grows wider over time (in log space, though not necessarily nonlog space, depending on the magnitude of the progress trend). In the simplest version of the model, the periodic shocks are independent and identically distributed (I.I.D.), but we use an augmented version in which they are correlated from one period to the next.

For Moore's law, the model is as follows:

$$y_{t+1} = y_t e^{-\mu} e^{\upsilon_{t+1}+\theta\upsilon_t}, \quad [5]$$

in which μ is the Moore exponent, $\upsilon_t \sim N(O, \sigma^2)$ are I.I.D. noise terms, and θ is an autocorrelation parameter. Following (27), we set $\theta = 0.63$.

For Wright's law, the corresponding model is as follows:

$$y_{t+1} = y_t \left(\frac{X_{t+1}}{X_t}\right)^{-w} e^{u_{t+1}+\rho u_t}, \quad [6]$$

in which w is the Wright exponent, $u_t \sim N\ (0,{\sigma_u}^2)$ are I.I.D. noise terms, X is the future deployment scenario (which is the same as that used for W2 as described in the previous section), and ρ is the autocorrelation parameter. Following ref. 29, we set $\rho = 0.19$. For a given technology, the parameters are estimated by using the observed data to perform an ordinary least squares regression through the origin. As such, the calibration relies on differences between sequential data points, which limits the data sources available for analysis (since many sources provide nonsequential data).

A hindcasting procedure was implemented in refs. 27 and 29 to statistically test the ability of these models to forecast observed progress trends. For this, a dataset of more than 50 technologies was used, spanning a variety of forecasting periods. For each model, many subsamples of data were used to make many forecasts for all technologies, and the resulting forecast errors were pooled to form an aggregate forecast error distribution. This was then compared to the forecast error distribution expected to arise from the model. In these two papers, the empirical and theoretical error distributions were a close match, indicating that the models did a good job at forecasting out-of-sample progress trends. Global values of the autocorrelation parameters (θ and ρ) for all technologies were estimated using the pooled sets of forecast errors.

Finally, refs. 27 and 29 derived analytical expressions for the forecast error distributions implied by each model in terms of all the estimated parameters. In this paper, we use the observed data for each technology to estimate all relevant parameters and then calculate the required percentiles of the probabilistic forecasts using these analytical expressions. The resulting forecasts thus take account of the underlying progress trends and observed volatility in different technologies' historical records as well as the future deployment scenario (for Wright's law) or time horizon (for Moore's law).

Developing Expert Elicitation Forecasts: Forecasting 2019 and 2030 Costs by Aggregating, Interpolating, and Extrapolating EE Estimates. It is important to note, first, that there are not large numbers of EEs in energy technologies [the first one we are aware of was published in 2008 (84)]. Second, many EEs only present cost estimates for one point in time, making it impossible to observe how cost distributions are expected to change over time. However, this is required in order to infer a distribution of costs at some intermediate year. Thus, we can only generate probabilistic 2019 cost forecasts for the subset of the expert elicitation studies that included information on technology costs in at least 2 y (Dataset S1).

In order to obtain estimated expert elicitation forecasts in 2019, we interpolate costs between 2010 and 2030 (nuclear and PV) or extrapolate costs between 2020 and 2030 (wind and water electrolysis) using an exponential functional form. This step is necessary for the comparison with model-based forecasts and realized costs in 2019 (reference *SI Appendix*, section 3 for further details).

Data Availability. The dataset containing all the expert elicitation and observed data collected and harmonized have been deposited in Technology Matrix Tool (https://tm.innopaths.eu/). The MATLAB code developed and applied for the Stochastic Exponent forecasts and the EE forecasts (using onshore wind as an example) is available on GitHub (https://github.com/jmeng-env/forecast-technological-change). All other study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. The majority of this research work was funded by the European Union's Horizon 2020 research and innovation program under Grant Agreements No. 730403 (INNOPATHS). We also acknowledge funding from the European Union's Horizon 2020 research and innovation program under Grant Agreements No. 730427 (COP21RIPPLES) and No. 853487 (ERCStg 2D4D); Partners for a New Economy; the Economics of Energy Innovation and System Transition (EEIST) program funded by UK Aid through the UK's Department for Business, Energy and Industrial Strategy (BEIS) and the Children's Investment Fund Foundation; the Natural Environment Research Council (NE/V002414/1); and The Royal Institute of International Affairs (Chatham House) project on the United Kingdom–China Cooperation on Climate Change Risk Assessment Phase 3. We would also like to acknowledge the help of the IRENA staff who helped to interpret some data, the participants of the 38th International Energy Workshop at the Organization for Economic Cooperation and Development in Paris, and the researchers of the project INNOPATHS (Innovation pathways strategies and policies for the low-carbon transition in Europe) who provided valuable comments throughout the project.

1. H. Brooks, Technology, evolution, and purpose. *Daedalus* **109**, 65–81 (1980).
2. L. D. Anadon *et al.*, Making technological innovation work for sustainable development. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9682–9690 (2016).
3. B.-Å. Lundvall, *National Systems of Innovation: Toward a Theory of Innovation and Interactive Learning* (Anthem Press, 2010).
4. R. R. Nelson, *National Innovation Systems: A Comparative Analysis* (Oxford University Press on Demand, 1993).
5. Organisation for Economic C-o & Development, *National Innovation Systems* (OECD Publications, Paris, 1997).
6. F. Malerba, Sectoral systems of innovation and production. *Res. Policy* **31**, 247–264 (2002).
7. C. Binz, B. Truffer, L. Coenen, Why space matters in technological innovation systems—Mapping global knowledge dynamics of membrane bioreactor technology. *Res. Policy* **43**, 138–155 (2014).
8. M. P. Hekkert, R. A. Suurs, S. O. Negro, S. Kuhlmann, R. E. Smits, Functions of innovation systems: A new approach for analysing technological change. *Technol. Forecast. Soc. Change* **74**, 413–432 (2007).
9. F. W. Geels, Micro-foundations of the multi-level perspective on socio-technical transitions: Developing a multi-dimensional model of agency through crossovers between social constructivism, evolutionary economics and neo-institutional theory. *Technol. Forecast. Soc. Change* **152**, 119894 (2020).

10. W. D. Nordhaus, "Modeling induced innovation in climate-change policy" in *Technological Change and the Environment*, A. Grubler, N. Nakicenovic, W.D. Nordhaus, Eds. (Routledge, 2002), vol. 9, pp. 259–290.
11. P. Aghion, X. Jaravel, Knowledge spillovers, innovation and growth. *Econ. J. (Lond.)* **125**, 533–573 (2015).
12. K. J. Arrow, The economic implications of learning by doing. *Rev. Econ. Stud.* **29**, 155–173 (1962).
13. K. Frenken, Technological innovation and complexity theory. *Econ. Innov. New Technol.* **15**, 137–155 (2006).
14. E. Verdolini, L. D. Anadón, E. Baker, V. Bosetti, L. Aleluia Reis, Future prospects for energy technologies: Insights from expert elicitations. *Rev. Environ. Econ. Policy* **12**, 133–153 (2018).
15. A. Kann, J. P. Weyant, Approaches for performing uncertainty analysis in large-scale energy/economic policy models. *Environ. Model. Assess.* **5**, 29–46 (2000).
16. L. D. Anadon, E. Baker, V. Bosetti, Integrating uncertainty into public energy research and development decisions. *Nat. Energy* **2**, 17071 (2017).
17. E. Baker, V. Bosetti, A. Salo, Robust portfolio decision analysis: An application to the energy research and development portfolio problem. *Eur. J. Oper. Res.* **284**, 1107–1120 (2020).
18. S. Pye, N. Sabio, N. Strachan, An integrated systematic analysis of uncertainties in UK energy transition pathways. *Energy Policy* **87**, 673–684 (2015).
19. J. DeCarolis *et al*., Formalizing best practice for energy system optimization modelling. *Appl. Energy* **194**, 184–198 (2017).
20. V. Bosetti *et al*., Sensitivity to energy technology costs: A multi-model comparison analysis. *Energy Policy* **80**, 244–263 (2015).
21. H. Dowlatabadi, Sensitivity of climate change mitigation estimates to assumptions about technical change. *Energy Econ.* **20**, 473–493 (1998).
22. E. Kriegler *et al*., The role of technology for achieving climate policy objectives: Overview of the EMF 27 study on global technology and climate policy strategies. *Clim. Change* **123**, 353–367 (2014).
23. V. Barros *et al*., "Impacts, adaptation, and vulnerability-Part B: Regional aspects-contribution of working group II to the fifth assessment report of the intergovernmental panel on climate change" (Cambridge University Press, Cambridge, United Kingdom and New York, NY, 2014).
24. L. Neij, Cost development of future technologies for power generation—A study based on experience curves and complementary bottom-up assessments. *Energy Policy* **36**, 2200–2211 (2008).
25. G. F. Nemet, Interim monitoring of cost dynamics for publicly supported energy technologies. *Energy Policy* **37**, 825–835 (2009).
26. E. Baker, V. Bosetti, L. D. Anadon, M. Henrion, L. A. Reis, Future costs of key low-carbon energy technologies: Harmonization and aggregation of energy technology expert elicitation data. *Energy Policy* **80**, 219–232 (2015).
27. J. D. Farmer, F. Lafond, How predictable is technological progress? *Res. Policy* **45**, 647–665 (2016).
28. B. Nagy, J. D. Farmer, Q. M. Bui, J. E. Trancik, Statistical basis for predicting technological progress. *PLoS One* **8**, e52669 (2013).
29. F. Lafond *et al*., How well do experience curves predict technological progress? A method for making distributional forecasts. *Technol. Forecast. Soc. Change* **128**, 104–117 (2018).
30. V. Masson-Delmotte *et al*., "IPCC: Summary for policymakers in: Global warming of 1.5°C. An IPCC special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty" (World Meteorological Organization, Geneva, Switzerland, 2018), p. 32.
31. G. Dosi, R. R. Nelson, "Technical change and industrial dynamics as evolutionary processes" in *Handbook of the Economics of Innovation*, B. H. Hall, N. Rosenberg, Eds. (Elsevier, 2010), vol. 1, pp. 51–127.
32. R. Wiser *et al*., Expert elicitation survey on future wind energy costs. *Nat. Energy* **1**, 16135 (2016).
33. G. F. Nemet, L. D. Anadon, E. Verdolini, Quantifying the effects of expert selection and elicitation design on experts' confidence in their judgments about future energy technologies. *Risk Anal.* **37**, 315–330 (2017).
34. M. G. Morgan, M. Henrion, M. Small, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis* (Cambridge University Press, 1992).
35. T. P. Wright, Factors affecting the cost of airplanes. *J. Aeronaut. Sci.* **3**, 122–128 (1936).
36. P. Thompson, The relationship between unit cost and cumulative quantity and the evidence for organizational learning-by-doing. *J. Econ. Perspect.* **26**, 203–224 (2012).
37. M. Grubb *et al*., Induced innovation in energy technologies and systems: A review of evidence and potential implications for CO2 mitigation. *Environ. Res. Lett.* **16**, 043007 (2021).
38. US Office of Technology Assessment, *Information Technology R&D: Critical Trends and Issues*. (US Office of Technology Assessment, OTA-CIT-268, 1985).
39. G. F. Nemet, Beyond the learning curve: Factors influencing cost reductions in photovoltaics. *Energy Policy* **34**, 3218–3232 (2006).
40. T. Wiesenthal *et al*., *Technology Learning Curves for Energy Policy Support*. (EUR - Scientific and Technical Research Reports Tech. Rep JRC73231, 2012).
41. Y. Qiu, L. D. Anadon, The price of wind power in China during its expansion: Technology adoption, learning-by-doing, economies of scale, and manufacturing localization. *Energy Econ.* **34**, 772–785 (2012).
42. D. Sahal, A theory of progress functions. *AIIE Trans.* **11**, 23–29 (1979).
43. C. Goddard, Debunking the learning curve. *IEEE Trans. Compon. Hybrids Manuf. Technol.* **5**, 328–335 (1982).
44. G. Sinclair, S. Klepper, W. Cohen, What's experience got to do with it? Sources of cost reduction in a large specialty chemicals producer. *Manage. Sci.* **46**, 28–45 (2000).
45. W. D. Nordhaus, The perils of the learning model for modeling endogenous technological change. *Energy J.* **35**, 1–13(2014).
46. T. Mayer, D. Kreyenberg, J. Wind, F. Braun, Feasibility study of 2020 target costs for PEM fuel cells and lithium-ion batteries: A two-factor experience curve approach. *Int. J. Hydrogen Energy* **37**, 14463–14474 (2012).
47. E. S. Rubin, I. M. Azevedo, P. Jaramillo, S. Yeh, A review of learning rates for electricity supply technologies. *Energy Policy* **86**, 198–218 (2015).
48. International Energy Agency, *Experience Curves for Energy Technology Policy* (OECD Publishing, 2000).
49. J. B. Mazzola, K. F. McCardle, The stochastic learning curve: Optimal production in the presence of learning-curve uncertainty. *Oper. Res.* **45**, 440–450 (1997).
50. O. Schmidt, A. Hawkes, A. Gambhir, I. Staffell, The future cost of electrical energy storage based on experience rates. *Nat. Energy* **2**, 17110 (2017).
51. M. Weiss, M. Junginger, M. K. Patel, K. Blok, A review of experience curve analyses for energy demand technologies. *Technol. Forecast. Soc. Change* **77**, 411–428 (2010).
52. A. Malhotra, T. S. Schmidt, Accelerating low-carbon innovation. *Joule* **4**, 2259–2267 (2020).
53. Y. Yu, H. Li, Y. Che, Q. Zheng, The price evolution of wind turbines in China: A study based on the modified multi-factor learning curve. *Renew. Energy* **103**, 522–536 (2017).
54. E. Williams, E. Hittinger, R. Carvalho, R. Williams, Wind power costs expected to decrease due to technological progress. *Energy Policy* **106**, 427–435 (2017).
55. J. Lilliestam, M. Labordena, A. Patt, S. Pfenninger, Empirically observed learning rates for concentrating solar power and their responses to regime change. *Nat. Energy* **2**, 17094 (2017).
56. M. Weiss *et al*., On the electrification of road transport-Learning rates and price forecasts for hybrid-electric and battery-electric vehicles. *Energy Policy* **48**, 374–393 (2012).
57. A. Van Benthem, K. Gillingham, J. Sweeney, Learning-by-doing and the optimal solar policy in California. *Energy J.* **29**, 131–152 (2008).
58. D. Newbery, Evaluating the case for supporting renewable electricity. *Energy Policy* **120**, 684–696 (2018).
59. G. Shrimali, E. Baker, Optimal feed-in tariff schedules. *IEEE Trans. Eng. Manage.* **59**, 310–322 (2012).
60. M. Grubb, J. Köhler, D. Anderson, Induced technical change in energy and environmental modeling: Analytic approaches and policy implications. *Annu. Rev. Energy Environ.* **27**, 271–308 (2002).
61. B. C. Van der Zwaan, R. Gerlagh, L. Schrattenholzer, Endogenous technological change in climate change modelling. *Energy Econ.* **24**, 1–19 (2002).
62. E. L. Hodson *et al*., US energy sector impacts of technology innovation, fuel price, and electric sector CO2 policy: Results from the EMF 32 model intercomparison study. *Energy Econ.* **73**, 352–370 (2018).
63. A. McDonald, L. Schrattenholzer, Learning rates for energy technologies. *Energy Policy* **29**, 255–261 (2001).
64. P. Söderholm, T. Sundqvist, Empirical challenges in the use of learning curves for assessing the economic prospects of renewable energy technologies. *Renew. Energy* **32**, 2559–2578 (2007).
65. K. Gillingham, R. G. Newell, W. A. Pizer, Modeling endogenous technological change for climate policy analysis. *Energy Econ.* **30**, 2734–2753 (2008).
66. Farmer J. D., Trancik J. *Dynamics of Technological Development in the Energy Sector* London Accord Final Publication (2007) pp. 1–24.
67. S. Alberth, Forecasting technology costs via the experience curve—Myth or magic? *Technol. Forecast. Soc. Change* **75**, 952–983 (2008).
68. G. Chan, L. D. Anadon, M. Chan, A. Lee, Expert elicitation of cost, performance, and RD&D budgets for coal power with CCS. *Energy Procedia* **4**, 2685–2692 (2011).
69. M. G. Morgan, M. Henrion, M. Small, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis* (Cambridge University Press, 1990).
70. R. A. Howard, J. E. Matheson, D. W. North, The decision to seed hurricanes. *Science* **176**, 1191–1202 (1972).
71. L. D. Anadón, G. Nemet, E. Verdolini, The future costs of nuclear power using multiple expert elicitations: Effects of RD&D and elicitation design. *Environ. Res. Lett.* **8**, 034020 (2013).
72. L. T. Lam, L. Branstetter, I. L. Azevedo, A sunny future: Expert elicitation of China's solar photovoltaic technologies. *Environ. Res. Lett.* **13**, 034038 (2018).
73. O. Schmidt *et al*., Future cost and performance of water electrolysis: An expert elicitation study. *Int. J. Hydrogen Energy* **42**, 30470–30492 (2017).
74. E. Baker, H. Chon, J. Keisler, Battery technology for electric and hybrid vehicles: Expert views about prospects for advancement. *Technol. Forecast. Soc. Change* **77**, 1139–1146 (2010).
75. A. Sakti *et al*., Consistency and robustness of forecasting for emerging technologies: The case of Li-ion batteries for electric vehicles. *Energy Policy* **106**, 415–426 (2017).
76. M. G. Morgan, Use (and abuse) of expert elicitation in support of decision making for public policy. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7176–7184 (2014).
77. A. Grubler, The costs of the French nuclear scale-up: A case of negative learning by doing. *Energy Policy* **38**, 5174–5188 (2010).
78. C. Wilson *et al*., Granular technologies to accelerate decarbonization. *Science* **368**, 36–39 (2020).
79. K. Surana, C. Doblinger, L. D. Anadon, N. Hultman, Effects of technology complexity on the emergence and evolution of wind industry manufacturing locations along global value chains. *Nat. Energy* **5**, 811–821 (2020).
80. M. Henrion, B. Fischhoff, Assessing uncertainty in physical constants. *Am. J. Phys.* **54**, 791–798 (1986).
81. L. D. Anadon, M. Bunn, V. Narayanamurti, *Transforming US Energy Innovation* (Cambridge University Press, 2014).
82. Fraunhofer Institute for Solar Energy S, Photovoltaics report (Fraunhofer ISE, Freiburg, Germany, 2018).
83. M. Bolinger, J. Seel, K. H. LaCommare, *An Empirical Analysis of Project Cost, Performance, and Pricing Trends in the United States* (Lawrence Berkeley National Laboratory, 2017).
84. A. E. Curtright, M. G. Morgan, D. W. Keith, Expert assessments of future photovoltaic technologies. *Environ. Sci. Technol.* **42**, 9031–9038 (2008).