PHILOSOPHICAL TRANSACTIONS A

royalsocietypublishing.org/journal/rsta

Review

Cite this article: AghaKouchak A *et al*. 2022 Status and prospects for drought forecasting: opportunities in artificial intelligence and hybrid physical–statistical forecasting. *Phil. Trans. R. Soc. A* **380**: 20210288.
https://doi.org/10.1098/rsta.2021.0288

Received: 24 April 2022
Accepted: 6 September 2022

One contribution of 11 to a Royal Society Science+ meeting issue 'Drought risk in the Anthropocene'.

Subject Areas:
hydrology

Keywords:
drought, prediction, climate, hydrology

Author for correspondence:
A. AghaKouchak
e-mail: amir.a@uci.edu

Status and prospects for drought forecasting: opportunities in artificial intelligence and hybrid physical–statistical forecasting

A. AghaKouchak[1,2], B. Pan[3], O. Mazdiyasni[1], M. Sadegh[4], S. Jiwa[2], W. Zhang[1], C. A. Love[1], S. Madadgar[5], S. M. Papalexiou[6], S. J. Davis[1,2], K. Hsu[1] and S. Sorooshian[1,2]

[1]Department of Civil & Environmental Engineering, and
[2]Department of Earth System Science, University of California, Irvine, CA, USA
[3]Lawrence Livermore National Lab, Livermore, CA, USA
[4]Department of Civil Engineering, Boise State University, Boise, ID, USA
[5]Katrisk LLC, Berkeley, CA, USA
[6]Department of Civil Engineering, University of Calgary, Alberta, Canada

AA, 0000-0003-4689-8357; WZ, 0000-0003-0224-3551; SMP, 0000-0001-5633-0154

Despite major improvements in weather and climate modelling and substantial increases in remotely sensed observations, drought prediction remains a major challenge. After a review of the existing methods, we discuss major research gaps and opportunities to improve drought prediction. We argue that current approaches are top-down, assuming that the process(es) and/or driver(s) are known—i.e. starting with a model and then imposing it on the observed events (reality). With the help of an experiment, we show that there are opportunities to develop bottom-up drought prediction models—i.e. starting from the reality (here, observed events) and searching for model(s) and driver(s) that work. Recent advances in artificial intelligence and machine learning provide significant opportunities for developing bottom-up drought

THE ROYAL SOCIETY PUBLISHING

forecasting models. Regardless of the type of drought forecasting model (e.g. machine learning, dynamical simulations, analogue based), we need to shift our attention to robustness of theories and outputs rather than event-based verification. A shift in our focus towards quantifying the stability of uncertainty in drought prediction models, rather than the goodness of fit or reproducing the past, could be the first step towards this goal. Finally, we highlight the advantages of hybrid dynamical and statistical models for improving current drought prediction models.

This article is part of the Royal Society Science+ meeting issue 'Drought risk in the Anthropocene'.

royalsocietypublishing.org/journal/rsta Phil. Trans. R. Soc. A 380: 20210288

1. Introduction

In May 2012, with no hint of an extreme drought from seasonal prediction models, the United States Department of Agriculture (USDA) predicted a record crop yield. US farmers had planted the largest area of soyabeans and corn in over 75 years [1], possibly due to favourable weather forecasts. Just one month later, the entire US Midwest experienced one of the most extreme droughts on record [2–4]. Approximately 80% of the crop lands experienced drought, resulting in more than $36 billion in damages [1]. The lack of an early warning was a striking aspect of this event, and possibly a determinant of overall damages. An early warning with even a two-month lead time would have likely reduced the drought impacts on the agriculture sector. The 2012 drought is one out of the many extreme drought events from around the world that were not predicted properly due to unreliable seasonal precipitation predictions [5]. Similar unforeseen droughts in East Africa have led to two major famines just in the past 50 years with death tolls in the hundreds of thousands [6,7].

Over the years, several research or operational drought (or hydrological) prediction systems have been developed, including the Climate Prediction Center Seasonal Drought Outlook [8], European Drought Observatory [9], the University of Washington's Surface Water Monitor [10,11], Princeton University's drought forecast system [12–14], US–Mexico Drought Prediction Tool [15], US Drought Monitor [16], Global Integrated Drought Monitoring and Prediction System ([17]) and the Columbia University's International Research Institute seasonal forecasts [18]. Most existing operational drought forecasting models rely on seasonal predictions from major numerical weather prediction centres such as the National Oceanic and Atmospheric Administration's (NOAA's) National Centers for Environmental Prediction (NCEP) and the European Centre for Medium-Range Weather Forecasts (ECMWF). The current drought prediction models (including non-operational research products) can be broadly categorized into three groups: (a) dynamical (process-based) models; (b) statistical models including data-driven artificial intelligence models; and (c) hybrid statistical–dynamical drought prediction systems.

Drought prediction models rely on one or more sources of predictability. Oceanic and stratospheric low-frequency signals, for example, offer valuable predictive information from short range to seasonal [19–21]. The Madden–Julian Oscillation (MJO) and El Niño-Southern Oscillation (ENSO) contribute to the formation of Rossby wave-trains influencing precipitation in mid-latitudes [22–26]. The effect of such teleconnections varies significantly in different parts of the world (e.g. a particular oceanic signal can lead to both wet and drought conditions in different regions). To leverage this information for drought prediction, statistical models typically build simplified relationships between drought and the relevant teleconnections, whereas process-based numerical models focus on improving forecasts by reproducing the known teleconnection processes using physics-based governing equations. Despite major improvements in weather and climate modelling [27] and the substantial increase in remotely sensed satellite and radar data [28], drought prediction remains a major challenge [29–31].

In this article, we offer a review of current models and the evolution of different types of drought prediction approaches. We discuss key gaps and opportunities to improve drought

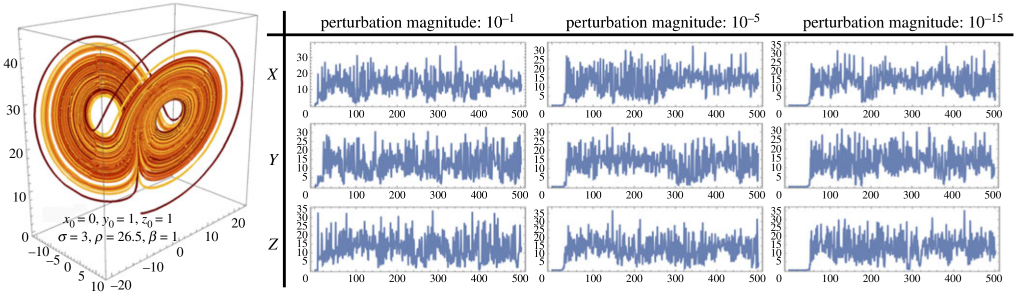

Figure 1. Numerical simulation of the [45] dynamical system across space and time: (left) trajectory simulation; (right) discrepancies in two simulated trajectories projected on the *x*, *y* and *z*-axes. Here, the initial states are perturbed by 10^{-1}, 10^{-5} and 10^{-15} (after [44]). (Online version in colour.)

prediction. We also highlight how the increasing attention to artificial intelligence and machine learning can improve future drought prediction models.

2. Physically based dynamical models

Physically based models (also known as dynamical models) have been widely used for drought prediction [32,33]. From the establishment of primitive equations for atmospheric dynamics (early twentieth century) to the first operational weather forecast model, it has taken nearly half a century to develop the methods needed for drought prediction. However, the current generation of dynamical models used for drought prediction were developed and operationalized mainly in the past two decades. These models deterministically describe the dynamic processes and interactions between the land, atmosphere and ocean using partial differential equations [34], and approximate the unresolved processes (i.e. radiation and cloud microphysics) using a suite of parametrization schemes (e.g. [35–42]). Over the years, biogeochemical and even microbial processes have also been incorporated in dynamic models [43]. Starting from some initial states (values), these equations apply thermodynamic and conservation laws on a control volume to simulate changes over time.

Uncertainty in model predictions comes from a wide range of sources including simplified assumptions, inaccurate initial values and limitations associated with parametrization and model tuning. A major source of uncertainty and limited predictability is the sensitivity to the initial states used for solving the governing equations. Small uncertainties in the initial states typically grow into significant trajectory divergence in the iterative numerical computations [44]. Figure 1, for example, shows the effects of slightly perturbed initial values on the outcomes of the [45] dynamical system—a highly simplified model describing the thermodynamics of a two-dimensional fluid layer warmed from below and cooled from above [45]. The figure shows that relatively small perturbations in the initial values translate into significant discrepancies in simulated trajectories. Despite significant progress in dynamic modelling, inaccurate initial value information has remained as a major source of uncertain predictions. Edward Lorenz (1917–2008), known for his butterfly effect theory, was not the first to point to the uncertainty/error associated with initial values (or broadly accuracy of information) on the theoretical limit to prediction and predictability. In his fundamental work on the so-called ‘Three-Body Problem’, the French mathematical physicist and philosopher, Henri Poincaré (1854–1912), discovered a chaotic yet deterministic system limiting our ability to accurately predict the movement of three planetary bodies—arguably a discovery leading to a new field of research known as chaos theory. With a focus on earth system models, Poincaré’s discovery and Lorenz’s mathematical experiments point to one crucial issue: accurate initial states play a significant role, often more so than other factors, on the accuracy of predictions and impose limits on predictability.

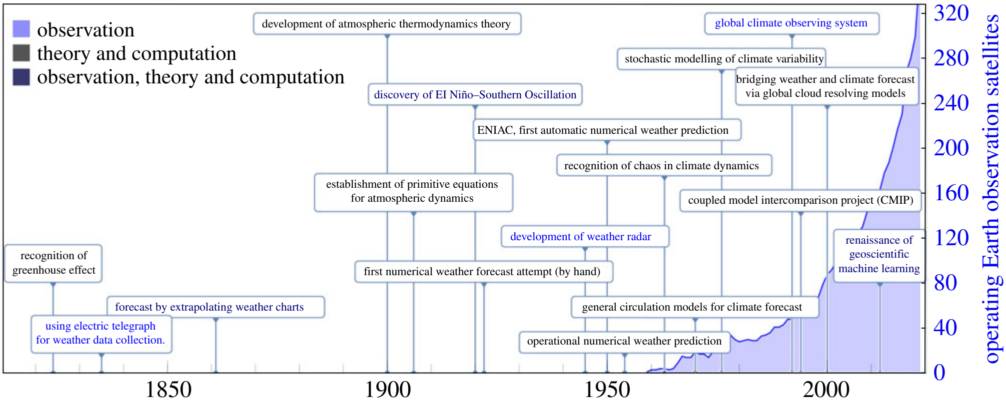

Figure 2. Milestones in observing and modelling the Earth climate system (after [26]). Achievements from observation are labelled in light blue, achievements in theory and computation are labelled in black, achievements from a combination of observation, theory and computation are labelled in navy blue. The curve on the right side denotes the number of operating Earth observation satellites. Satellite data are obtained from World Meteorological Organisation Observing Systems Capability Analysis and Review Tool. (Online version in colour.)

In recent decades, the growing number of observations, including *in situ* measurements and remotely sensed satellite and radar data products, have significantly improved numerical weather models by providing more accurate initial states, and also information that can be used for assimilation [46–50] and forecast verification [46,47]. Along with the improvements in observations, the computational power for solving numerical models has also increased substantially. Figure 2 summarizes major advancements in the development of physically based models and relevant Earth observations. The most striking aspect is the rapid increase in the emergence of satellite observations and ground-based radars increasing the accuracy of initial states and reference data for validation. Bridging the gap between weather and climate forecasts via global cloud resolving models (e.g. [51,52]) has played a pivotal role in improving the accuracy of long-range forecasting (see the timeline in figure 2). Previous studies have also demonstrated an improvement in the forecasting skill of dynamical models when the variability of large-scale oceanic and atmospheric teleconnections is incorporated (e.g. [53–56]).

Over the years, a wide range of models have been developed around the world. To benefit from the strength of a diverse set of models and overcome the limitations of individual dynamical models (e.g. resolution, parametrization), there has been a growing tendency to use multiple models and combine their forecasts [57–59]. A multi-agency effort led to the development of the North American Multi-Model Ensemble (NMME; [60]), which has provided climate forecasts ranging from intra-seasonal to inter-seasonal scales. The NMME ensemble average generally offers better predictions than individual contributing models do for different regions [60–62]; however, the reliability of the precipitation forecasts in general is still rather low [60,61]. Figure 3*a*, for example, shows NMME predictions indicating a wet period in the western USA during December 2014–February 2015 from a forecast initiated in October 2014 (two-month lead time). The region, however, experienced one of the most extreme droughts on record during the same period (see US Drought Monitor maps in figure 3*b*–*d*). That said, NMME, ECMWF and other prediction systems sometimes predict droughts accurately, but rarely beyond a six-month (often three) lead time.

Advancements in physically based climate models in the past decades have led to more accurate and reliable predictions including seasonal drought forecasts [63–65]. However, the predictability of physically based models for drought prediction are still far from ideal and are highly variable in space and particularly at long lead times [54,66,69].

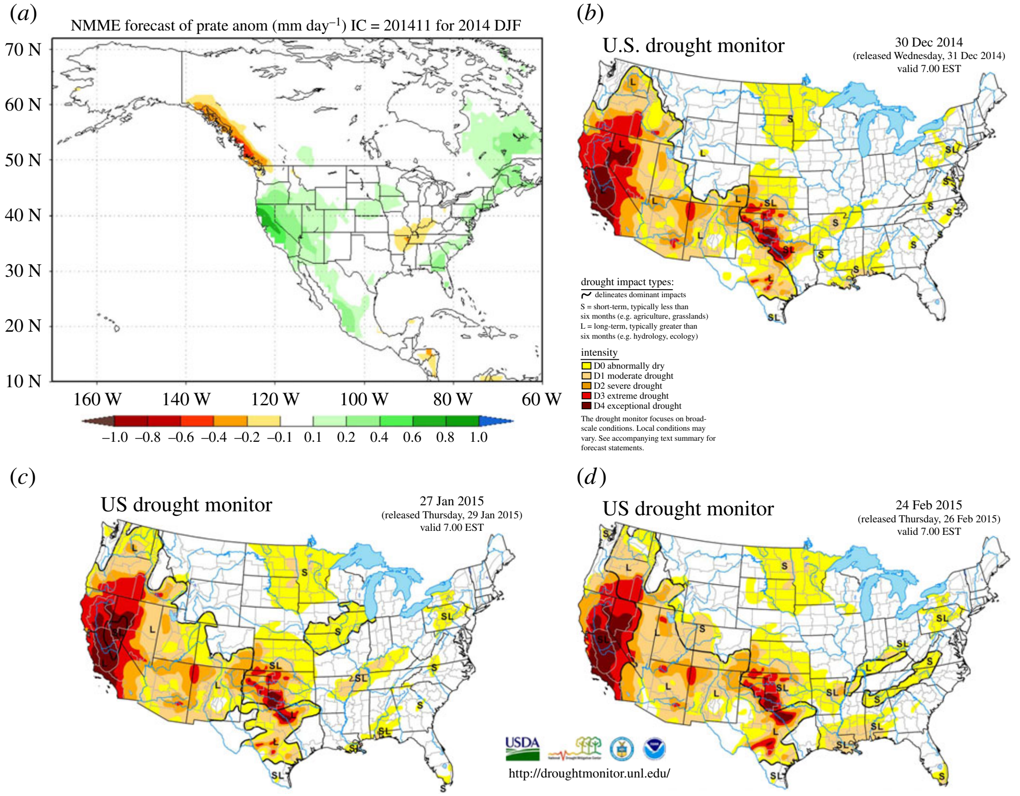

Figure 3. (*a*) Drought predictions based on NNME indicating a wet period in the western USA during December 2014–February 2015 from a forecast initiated in October 2014 (two-month lead time). (*b–d*) Observed extreme and exceptional drought based on the United States Drought Monitor maps. (Online version in colour.)

3. Statistical models

Along with dynamical models, a wide range of statistical drought prediction models have been developed over the years. Instead of solving the underlying governing equations, the premise behind statistical drought forecasting is that (a) there is temporal persistence in drought indicators [70] such as the Standardized Precipitation Index (SPI; [71]) and Palmer Drought Severity Index (PDSI) [72,73], and in land surface characteristics like soil moisture and snow cover [6]; (b) there is a relationship between different drought types. For example, hydrological (i.e. runoff/groundwater deficit) and agricultural (i.e. soil moisture) droughts are preceded by meteorological drought (i.e. deficit in precipitation) and hence the latter can forecast the state of the former (e.g. [74]); and (c) the observed relationship between drought and ocean or atmosphere conditions (aka teleconnections) [75–77] can be used for seasonal drought/precipitation forecasting (e.g. [78–82]).

To address these premises, statistical drought models use a plethora of predictors, representing the underlying processes [83] or relationship between drought-related variables or drought features (e.g. [84–89]). While earlier studies were mostly reliant on the persistence properties of drought indicators [32,33,90], recent studies merge information from initial land conditions and climate/weather data (i.e. persistence) with large-scale climate indices (i.e. teleconnection) [91,92].

Oceanic–atmospheric teleconnections and circulation patterns that affect the large-scale precipitation producing phenomena are forced by sea surface temperature (SST) anomaly [93,94], land atmosphere feedback [95] and natural and anthropogenic changes in radiative forcing [96]. The SST anomaly is presented in terms of large-scale indices such as ENSO, Pacific Decadal

Oscillation (PDO) and Atlantic Multidecadal Oscillation (AMO) [77,97,98]). Also referred to as teleconnections, these ocean precursors are often used for seasonal scale drought forecasting at large spatial scales, such as large river basins [99].

In many drought forecasting models, predictability of drought relies on the slowly varying boundary/climatic conditions (e.g. SST anomaly and land surface characteristics) over time [29]. Atmospheric, oceanic and hydrological predictors of drought are region and season specific [100], and their predictive skill varies by spatial and temporal scales, and lead time [32,101]. Different regions show widely varying correlations with large-scale oceanic and atmospheric indices [96,102,103]. Correlation, composite analyses and principal component methods are commonly used to identify the proper drought predictors at the spatial and temporal scales, season and lead time of interest [19,104]. We will discuss the limitations of this so-called top-down perspective in drought forecasting.

It is tempting to include all variables that have predictive skills in forecasting drought, but the modeller should be cognizant of redundant information in various data sources and model overfitting problems [105]. The latter case may lead to poor performance, especially in cases that involve states that are out of bound of training data [105]. Several methods have been widely used in the literature to reduce dimensionality of the drought forecasting problem, including removing redundant information and finding orthogonal predictors, such as the principal component analysis and canonical correlation analysis [106]. The identified predictors are then used to build a statistical model that can take one or more of the general categories described below.

(a) Stochastic time series models

Time series models rely on the dependence between the predictor–predictand relationship and/or persistence property of drought indices (e.g. soil moisture) [107–111]. This persistence property is mainly due to the accumulation (e.g. six-month SPI) or autoregressive property (e.g. memory or autocorrelation of PDSI) of the drought indices [29] and the fact that drought is relatively slow moving/evolving in nature. For this reason, longer time scales (e.g. 12-month SPI relative to six-month SPI) and shorter lead times (one month relative to three months) are relatively more predictable [101]. In general, autoregressive models ([112,113]) are appealing as they exploit the association of the future state of a process with its current and past states (univariate case), or the association with current and past states of other processes (multivariate case). Autoregressive integrated moving average (ARIMA) models are widely used for drought prediction [32,33]. More advanced models, such as the seasonal autoregressive integrated moving average ([32,90]), have been developed to address the shortcomings of the ARIMA family [114]. Most existing models in this category assume a linear relationship between predictors (e.g. previous SPI) and predictand (e.g. future SPI) and may fall short when this relationship is nonlinear [29]; however, recent advances [115] allow simulations preserving copula dependence structures and non-Gaussian distributions.

(b) Regression models

Regression analysis refers to a large group of predictive models that include linear, multiple linear regression models including the generalized linear models (e.g. [116]), more complex nonlinear regression and machine learning (ML)/artificial intelligence (AI) models [114,117,119]. A large body of literature uses multiple linear regression models (e.g. [4,17,29], and the references therein), and hence we separate this type from ML/AI models (next section). Multiple linear regression models allow for incorporating several predictors, including variables that represent persistence (i.e. antecedent drought indicators) as well as those of teleconnections (i.e. SST anomaly), and hence are more flexible to merge information from various sources for improved drought forecasting, but are more prone to overfitting [29]. These models allow incorporating nonlinearity by transforming the predictors. The logistic, log-linear and other regression function regimes were found useful in forecasting SPI and SPEI as well as other drought impacts [119].

These methods have also been used in pattern-based down-scaling of data useful in forecasting. One downside of regression models is the assumption of linearity or a particular form of nonlinear behaviour, which does not always hold at long time scales.

(c) Bayesian models

Naive Bayes models are classification techniques based on Bayes' theorem and have been used in drought forecasting [120,121]. The naive assumption is that models assume independence among predictors. Naive Bayes' models have been shown to be sensitive to the predictor variables used. The classification power of these models has been used to focus on crop-specific response to drought [122]. Bayesian network models have been used as an accurate statistical scheme for probabilistic modelling aiming to incorporate cause–effect relationships between variables for water resource management purposes [123]. These simple models have been shown to perform well in classification on a low number of training data.

While many papers have shown the strength of statistical models for drought prediction, their overall performance is not significantly different from dynamic models. Some studies have reported the improved performance of statistical models relative to dynamic ones (e.g. [124]), other studies have shown that dynamic models at least outperform persistence-based statistical models (e.g. [125]). However, both methods continue to fall short when it comes to long-range forecasting (e.g. six-month or even three-month lead times).

4. Artificial intelligence and machine learning models

AI refers to a general category of methods and models, including ML models, that leverage computers and machines to imitate the human mind's problem-solving and decision-making capabilities. AI/ML models are statistical by nature and most of the issues discussed in the previous section also apply to them. Continual increases in computing power and the availability of data has allowed researchers to expand on and improve drought forecasting models of this type, the theories for which draw from the evolving fields and subfields of statistics, hydrology, physics and computation. Drought forecasting using machine learning and artificial intelligence seeks to estimate and predict characteristics, indices and intensity of drought phenomenon over different spatial and temporal scales using computer algorithms that approximate models based largely on data. The current state of drought forecasting using ML/AI can be analysed by considering model class characteristics and overall usefulness which evolve and hybridize (including with physically based models) as the field continues to grow. In this section, we discuss the recent advances in AI/ML models separately in more detail.

AI models commonly used for drought forecasting or similar applications with the potential to be used for drought prediction include fuzzy logic based models ([126–129]), genetic algorithm ([130,131]), genetic programming ([132]), clustering methods such as K-means and nearest neighbour [133–135], as well as a variety of ML models such as artificial neural networks (ANNs; [26,118,136,142–]), support vector regression ([72]), support vector machine (SVM; [143,144]), decision tree ([145,146]) and random forest ([147–151]). More recently, boosting algorithms, such as XGBoost [152] and deep generative models [153–155] including variational autoencoders (VAEs; [156]), and generative adversarial networks (GANs; [157]) have shown great promise for drought forecasting performance improvement. Further, models that focus on predicting a distribution rather than a value, e.g. contextual generative adversarial nets ([158]), are particularly attractive for drought prediction applications, specifically when fine-tuned to the drought prediction domain (e.g. by constraining the loss function). AI models vary greatly with respect to complexity and a more complex model (e.g. GANs) may not necessarily outperform a simple ANN model, especially when spatial or temporal patterns/changes are not very important.

ANN models have been widely used in climate data analysis and hydrology for detecting nonlinear interactions between input and output variables [26,159,164]. The structure of these

ANNs consists of layers of nodes, going from input to output, connected by nonlinear functions, and a variety of different neural network architectures exist to better learn different types and structures of data [159]. ANN models have also been used to investigate the influence of global oceanic/atmospheric circulation patterns (e.g. NAO, ENSO) on drought [117,136,138–165]. The so-called recursive neural networks map from several previous states to a single output node to forecast one state ahead, whereas direct approaches estimate several future states from previous states. Recursive models perform better at smaller time scales, whereas direct models are expected to perform better over longer time periods though this issue cannot be simply generalized [118,166]. ANN methods provide great flexibility in computation, making them attractive for modelling complex phenomenon like drought; however, they can be prone to overfitting, and the computations involved can easily become a black box if proper attention is not paid to model construction [167]. ANN models are frequently combined with other approaches, including other probabilistic methods and even deterministic models [168,169].

A more attractive and powerful category of AI models, with significant potential applications to drought prediction, is the so-called deep generative models. Broadly, generative models combine the benefits of neural networks and probabilistic methods for a wide range of applications [153]. Two key architectures in this class include the variational autoencoders (VAEs; [156]) and GANs ([157]). These models have been shown to yield accurate, high-resolution predictions for precipitation, as well as good computational performance [154]; however, work remains in the application of these models [155].

Regardless of the model type, a major challenge in the field of environmental sciences including drought prediction is limited data size for training and validation purposes. There are opportunities to address this issue, for example through transfer learning (e.g. [170]) or training models on climate model simulations [171] that we have not addressed as they are beyond the scope of this drought-focused article.

The aforementioned list is by no means exhaustive, and models are not mutually exclusive as often drought forecasting will employ combinations of model sub-classes to different model functions depending on their relative strengths. Our goal is not to discuss one method versus another. Instead, in the following section we offer some thoughts on how AI/ML models can improve drought prediction.

5. Prospects for improving drought prediction

(a) Machine learning and uncertainty

In AI/ML applications to drought prediction (and arguably most other climate and hydrology applications) model developers justify their method by showing they can reproduce some historical observations through minimizing an error function or maximizing a dependence metric. Here, we argue for the need to shift our focus from just reproducing the past data, to the stability of the uncertainty/error along with reproducing historical events. Data-driven AI/ML models allow processing large datasets from climate simulations, remote sensing and observations, offering unique opportunities to benchmark and examine drought-related processes and prediction. The trend in published work in this area indicates foresights for acceleration of model development and diagnosis in this area. Most current attempts involve using off-the-shelf machine learning tools to replace or refine a process or broadly physics-based models. We claim that continuation of this trend does not necessarily lead to any significant progress or breakthrough in drought prediction, unless we shift our focus from matching simulations with some observations, to reliability and stability of the uncertainty in the evaluation process. This means that we need to make a distinction of aleatoric and epistemic uncertainty before formulating a machine learning model—an important fundamental step that has been systematically ignored in our community.

An AI/ML model is typically a parametric function, whose parameters and parametric form relies on our understanding of the problem. Given some training data (e.g. historical drought events), we optimize the AI/ML parameters of the learning machine to maximize/minimize an objective function (e.g. predicting historical droughts with a certain lead time and smallest error possible). A common formula for understanding this process is as follows [172]:

$$\mathrm{L}_{\text{earning}} = \mathrm{R}_{\text{epresentation}} + \mathrm{E}_{\text{valuation}} + \mathrm{O}_{\text{ptimization}}.$$

Here, representation refers to the parametric form of the AI/ML model, as well as the feature representation of the data; evaluation refers to the objective function that measures the performance of the AI/ML model; and optimization refers to the process of searching parameters for the optimally performing ones. For discussion on how the different learning algorithms discussed above fit into this formula, see [172].

Most recent studies focus on the aspect of representation (R), in the sense that learning algorithms of different parametric forms are proposed and tested for specific tasks (e.g. improving lead time in drought prediction). Deep neural networks, which learn hierarchical feature representations from raw data, have achieved considerable success, compared to conventional learning algorithms that work on predefined feature representations. However, the aspect of evaluation (E) is often overlooked in developing data-driven models, as most studies fail to distinguish various uncertainty sources in defining the learning objective function. To illustrate this point, consider how we use dynamical seasonal forecasts versus statistical (including AI/ML) seasonal forecasts. In dynamical seasonal forecasting, we apply numerical models to explicitly simulate plausible climate state trajectories, starting from an ensemble of possible current climate states. The uncertainties arising from model initialization, model formulation and ensemble sampling can be quantified, though the process might require substantial investments of time and computations resources. For example, one can easily separate the uncertainty associated with initial conditions using the kind of experiments shown in figure 1 (note that here we focus on model uncertainty and we are not referring to the overall uncertainty including lack of knowledge and information).

In AI/ML seasonal forecasting, we apply machine learning tools to simulate the statistical dependency between the predictand and its drivers, based on the available data, mostly from observations. The uncertainties from the training data, internal climate variability, and model representation are often mingled into a single objective function subject to optimization. Unlike processed-based models, separating these uncertainties are not that straightforward, if not impossible at present time. Let us make a distinction between these sources of uncertainty. The uncertainty from the training data is typically beyond the control of the modeller; however, it can be evaluated by testing different fractions of the data randomly for training. While this random sampling offers a general idea about the sample size uncertainty, it does not offer much insight into the uncertainty associated with the quality of the data used for training. For this reason, we can argue that this source of uncertainty is beyond the control of the modeller. Uncertainty associated with model representation, at least in theory, can be tested by trying a wide range of model structures, though it is rarely investigated/considered in practice. Uncertainty from internal climate variability is due to the chaotic nature of climate–atmosphere–ocean dynamic interactions, and varies as a function of the initial climate state and processes involved. In an ideal world, these sources of uncertainty should not be simply minimized, but should be quantified using a probabilistic representation, and reduced as much as possible. Without clear distinction and treatment of these uncertainties, we cannot make significant advances in drought prediction, nor can we gain new foresights about the barriers imposed by these uncertainties. We believe that the first step forward should be a shift in our focus toward quantifying the stability of uncertainty, rather than the goodness of fit (figure 4).

To illustrate this issue, let us assume an AI/ML drought predictions model (blue line in figure 4) in a historical period is being evaluated against observations (red line in figure 4), and the dotted lines show the overall uncertainty bounds. Here, the model is applied to four different locations or four different periods (figure 4*a*–*d*) and from the error perspective all models are equally good. In most applications in the literature, the focus is on the goodness of fit (e.g. the root

royalsocietypublishing.org/journal/rsta *Phil. Trans. R. Soc. A* **380**: 20210288

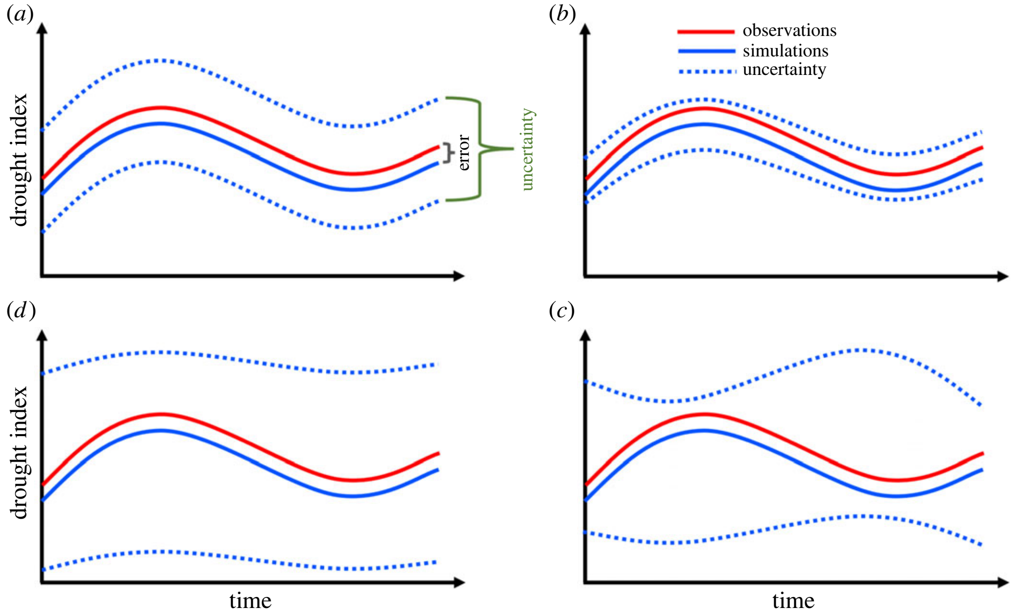

Figure 4. Conceptual illustration of instability of uncertainty bounds in model simulations of four different events with similar overall error. (Online version in colour.)

mean squared error or some indicator of dependence between the red and blue lines). However, the uncertainty bounds are not consistent from one location to the other (or from one event to another). This can be considered as unstable or inconsistent uncertainty when the model is used in different locations or different periods in the same region. While we agree that error is important, we argue that we need methods specifically designed to evaluate the stability of the uncertainty in our AI/ML models or broadly statistical models. This also applies to physically based dynamical models, but it is more important for black box AI/ML models in which separating error sources is not straightforward. Of course, an unusually wide range of uncertainty is not helpful even if the uncertainty bounds are stable and consistent across different locations. We need to make progress in both the spread and its stability. While there are numerous uncertainty quantification methods and ensemble assessment metrics [173–177], we need to develop drought specific uncertainty metrics and objective functions that focus on both stability of uncertainty and its range.

We note that the hydrological and climate sciences community has long realized the need for distinguishing different sources of uncertainty through defining the terms *aleatoric* uncertainty referring to the intrinsic random component of the climate system (e.g. the internal climate variability) and *epistemic* uncertainty referring to uncertainties arising from our imperfect state of knowledge (e.g. systematic biases in data and models) [178–180]. While this distinction of aleatoric and epistemic uncertainty is well understood, we have not yet applied these concepts well in our AI/ML model design and/or objective functions. Neither do we believe that mathematically separating the two types of uncertainty is necessarily required as in practice what matters is the overall uncertainty of the model. What is more important is quantifying and evaluating the stability of uncertainty in different regions or periods. We predict that more efforts in deriving better objective functions that go beyond mere error covering the broader uncertainty characteristics, especially stability of uncertainty, can lead to a major advance in our not AI/ML models, but also dynamical model simulations.

(b) Hybrid dynamical–statistical models

Earlier in this article, we discussed a wide range of statistical and dynamical models. In a seminal work, Hasselmann [181] showed how stochastic models can emerge from deterministic

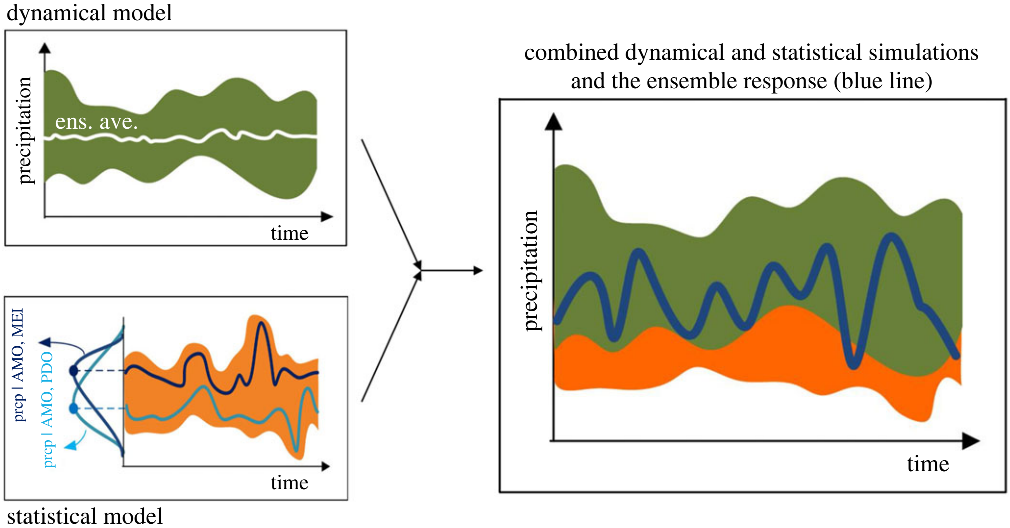

Figure 5. Schematic view of a hybrid statistical–dynamical model in which the statistical and dynamic models are developed and parametrized independently. The outputs are then merged using an ensemble merging/response algorithm. MEI, Multivariate Enso Index (Online version in colour.)

formulations. Since then, many studies have shown that combining statistical and dynamical models can improve seasonal precipitation prediction (e.g. [81,182]). In most applications, however, statistical aspect is mainly a post-processing component applied to the dynamic model outputs for weighting the ensemble members, uncertainty quantification or data assimilation (e.g. [183–187]). We believe there is a need for fully hybrid dynamical and statistical (including AI/ML) seasonal drought prediction frameworks to take advantage of the strengths of both types of models. In a fully hybrid approach, instead of using the two types in series, dynamical and statistical models should be applied in parallel and even independent of each other (e.g. [188–190]) and then combined using an ensemble merging or response method (e.g. [191]). Madadgar *et al*. [188] introduced a fully hybrid dynamical–statistical model using NMME simulations and a Bayesian analogue-year approach based on oceanic and climatic circulation patterns. The two ensembles are then merged using the so-called Expert Advice algorithm [191]—figure 5 for a schematic description and an example application to California.

In addition to improving drought prediction, by combining the strength of two types of models, insight can be gained about the relative contributions of the different types of drought prediction models across space and time. For example, a previous study showed that the ensemble mean of the NMME drought predictions received higher weights in northern California, whereas a statistical model attained higher weighting factors in southern California, increasing the overall hybrid predictability by 5–60% relative to each individual model [188]. While numerous studies have emerged on integrating statistical models through post processing and/or data assimilation, little progress has been made in fully hybrid methods. We believe more efforts should focus in this direction, especially for improving regional drought prediction. Drought forecasting methods often focus on continental-scale or national-level predictions, paying less attention to regional forecasting and assessment. However, given substantial variability across spatial scales in terms of both impacts and drought drivers, for improving water management and decision-making we should focus on more regionally relevant models. At regional scales, using a fully hybrid model, different sources of drought predictability may be harnessed through different forecasting methods (e.g. statistical dependence with climatic and oceanic indices versus physics-based modelling). We note that the earlier point about the need to focus on stability of uncertainty, instead of just performance of the ensemble mean relative to observations (discussed in the previous section), also applies to hybrid statistical–dynamical models.

Also, we should consider whether uncertainty, conceptually and from a theoretical viewpoint, should be considered 'stable' or 'unstable' to better inform objective functions using this concept. We could assume uncertainty as a process itself and if 'unstable'over time or in response to a process then it could be described and studied by a non-stationary process. The question now is if the characteristics of such a process can be identified and quantified. Uncertainty as a non-stationary space–time process results by the space–time variability of its two main components, that is aleatoric and epistemic uncertainty. By definition, aleatoric uncertainty cannot be reduced as it regards the stochastic component of a process given that the process is described by the best possible model. Yet this does not imply that the best possible model performs equally well in different locations or time periods. The same process, e.g. drought, for physical reasons might be more chaotic and unpredictable in a specific location than in another. On the other hand, epistemic uncertainty can be reduced if additional or better information feeds or modifies the model. This does not imply that our potential to improve a model or the accuracy of information is the same at different location or periods in time. It seems, thus, the idea of conceptualizing and studying uncertainty as a non-stationary process in space and time might offer new insights.

(c) Empirical teleconnections

We postulate that the recent focus on predicting droughts based on known teleconnections (e.g. ENSO) has potentially slowed down the progress in improving our forecasting capabilities. Is the actual predictive signal within the known teleconnections or unknown phenomena waiting to be discovered? Observed changes in sea surface temperatures, defining some of the widely used teleconnections, could be the result of undiscovered oceanic currents causing a vertical temperature gradient in certain locations, leading to sea surface temperature changes. Always starting with known teleconnections (or at least what we think we know), leads to a Platonic mindset (after [192]) privileging known teleconnections over less elegant or difficult to explain alternatives. When this Platonic mindset meets the chaotic reality of the ocean–atmosphere interactions and drought prediction, it becomes quite clear that there is a huge gap between what we actually know relative to what we think we know (i.e. our drought predicative power is very limited despite a large number of articles claiming strong teleconnections between different locations). This can become even more clear if we assume that uncertainty is a non-stationary space–time process, and, thus, changes in space and time.

To illustrate this issue, we have designed a simple experiment. Note that we are not presenting this experiment as a new predictive model that should be used from now on. Instead, we use the following experiment to show an alternative approach we term, empirical teleconnections. The experiment conducts an exhaustive, bottom-up, brute-force search algorithm for discovering predictive relationships between specific ocean–land regions (see the details in appendix A). Instead of assuming a certain relationship exists, we search for every combination of oceanic surface temperatures from the entire world (millions of them) to find the combination offering the best predictive drought information at 3–12-month lead times. For the sake of simplicity, assume 'drought' corresponds to precipitation below the 33rd percentile, 'wet' refers to precipitation above the 66th percentile and 'normal' would be everything in between. For the state of California, for example, after exploring all the possible ocean–land combinations, several locations (dots in the inset map in figure 6) appear to give the highest predictive information. The locations are empirically derived and do not overlap collectively with the established teleconnections commonly used for drought prediction. Figure 6 displays predicted November–April precipitation in California using sea surface temperatures from the selected locations against observations after a leave-one-out cross-validation (see appendix A). The figure shows that the empirically derived teleconnections (EmpTe), significantly improve drought prediction relative to the existing teleconnections.

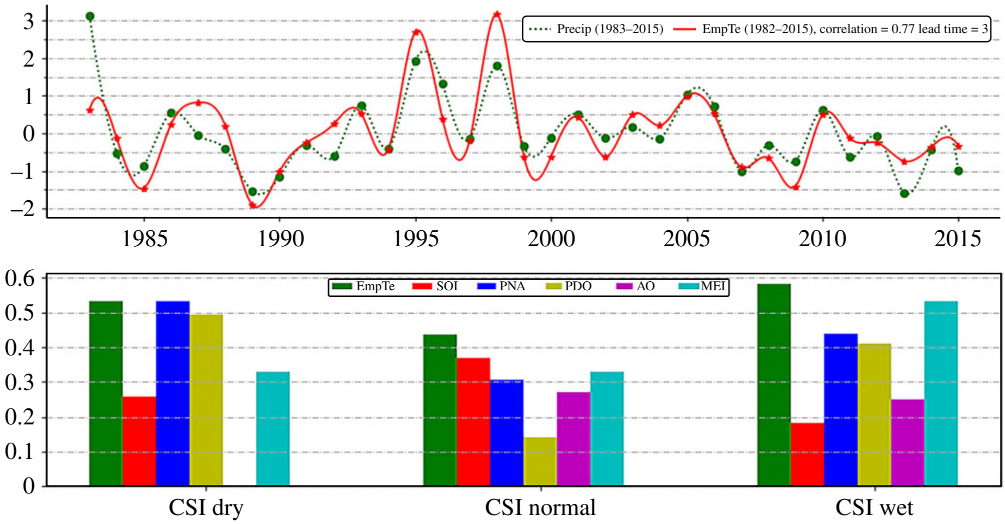

Figure 6. California precipitation, the best empirical teleconnection (EmpTe) and the existing teleconnections. (*a*) California wet season precipitation closely follows EmpTe. (*b*) The bar graph of EmpTe hit and false alarm rates compared to five well-known teleconnections. The locations that provide predictive information are highlighted with dots in the global map in the top right panel. (Online version in colour.)

The first question that comes to mind is whether the seemingly good results are due to overfitting. We will address this issue in the following section (see Closing the loop: towards bottom-up verification). The model captures California's significant wet and dry years, such as the 1997/1998 precipitation event and the 2011–2014 drought (figure 6). Relative to the other known teleconnection, the EmpTe offers higher hits (correctly identifying wet and dry periods) and lower false predictions (see appendix A for a more detailed description of false and hit). This issue highlights the potential of exhaustive search methods for identifying what we call empirical teleconnections. This is where AI/ML models can potentially help up with the next breakthrough. Deep learning models that do not have predefined structures (i.e. can handle a wide range of inputs without making assumptions about the structure of the network) can be particularly attractive for this kind of investigation.

Note that in this example, we only explored all possible combinations of sea surface temperature pixels and California precipitation. We cannot explain the physical reasons for these EmpTe improving predictive information. It could just be a statistical flaw. However, if the predictive results remain consistent (more on this issue later) after rigorous tests, it will indicate that there are potentially unknown processes waiting to be discovered. Ideally, other drought relevant variables should be included, mounting the combinations to billions and billions. In doing so, the search for empirical teleconnections could potentially lead to discovering even new physically explainable teleconnections or unknown phenomena. We emphasize that when there are many potential predictors and input variables (here, millions of combinations of sea surface temperature), spurious predictive patterns/ability may emerge just by chance. For this reason, having rigorous methods for validation and verification is critical. The bottom-up approach discussed in the following section is one out of many ways to avoid potential spurious predictive patterns.

(d) Closing the loop: towards bottom-up verification

Let us revisit the question we asked earlier: Are the empirical teleconnections shown in figure 6 real or possibly a statistical artefact or spurious predictive pattern? Although we conducted

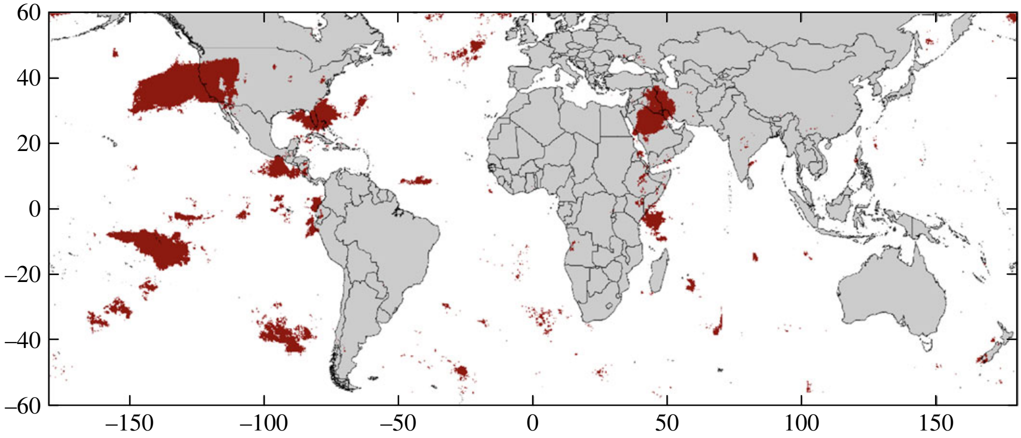

Figure 7. California Empirical Teleconnections (EmpeTe) significant correlation map. This map shows the grids that EmpeTe has a statistically significant (significance level = 0.01) correlation with precipitation across California, as well as parts of the ENSO region and the Middle East. (Online version in colour.)

a commonly used leave-one-out cross-validation, when there are millions and millions of combinations of variables the potential for overfitting is high and, hence, the results could still be physically meaningless. Correlation between two or more variables, even a very strong one, does not necessarily prove causality or predictive information (e.g. [193]). We believe we need to go beyond the typical validation and verification methods. One option is to close the loop by some form of backward analysis. We usually try to find one or more predictors and then we build a top-down model to predict drought. If we start with historical droughts and consider all potential predictors, do we reach the same set of predictors that we typically use in our top-down models? Considering the experiment in figure 6, what would our results look like if we conducted the experiment backwards?

In figure 7, we determine the precipitation grid boxes that have a statistically significant correlation, for example with the empirically derived (EmpTe) SST locations that provide predictive information for California (i.e. the EmpTe locations shown in figure 6). If the empirical teleconnections are statistical artefacts the backward analysis should lead to some spatially random pattern. If the area that shows a significant correlation with the predictive points displays some consistent spatial pattern, we can infer that the relationship may (but still not necessarily) be related to physical processes across coherent geographical regions that were previously unknown. Figure 7 displays the regions where wet season precipitation has a statistically significant (significance level = 0.01) correlation with the locations that provide predictive information. We show that our empirical teleconnection for California has a statistically significant correlation with precipitation across California and the southwestern USA, demonstrating a strikingly coherent spatial pattern. This indicates that the observed relationship is likely not random and involves a data-driven teleconnection that cannot be explained by previously defined teleconnections, moisture transport patterns, etc. It is interesting to note that there is also a significant relationship with our empirically derived teleconnections and the ENSO region in the Pacific Ocean. This figure portrays a relationship between EmpTe locations and California wet-season precipitation that is not a mere coincidence, but rather displays a strong correspondence to the entire region.

Unfortunately, we cannot explain the reason or confidently claim that there is a physical reason behind this promising backward–forward analysis. That is not the point we would like to highlight here. Instead, we emphasize that the AI/ML models (from simple brute force search algorithms to complex deep learning methods) can help us, first, identify such potentially predictive information from a bottom-up approach without any *a priori* assumption on the drivers. The current focus in the AI/ML community to improve drought prediction (i.e. improving the lead time or match between predictions and observations) from a top-down perspective is important but it is unlikely to result in a major leap. Instead, shifting our AI/ML models to build forward–backward predictive assessments, could lead to important discoveries. The outputs of such models can be the starting point to focus on physical understanding and search for process-based reasons for the empirical teleconnections. While this kind of forward–backward analysis cannot be done with the current dynamical models, learning algorithms and pattern recognition approaches are ideal for such analysis. Further, the outputs of an AI/ML-powered EmpTe model can be used as the statistical component of a hybrid statistical–dynamical drought prediction model. Note that a bottom-up or forward–backward assessment can only advance drought prediction when the right model is used, and input data provide relevant information. If the model is not representative or input variables provide insufficient information and/or include substantial overlap, the additional steps for forward–backward assessment does not lead to any improvement.

Having representative models and data, we postulate that by shifting our top-down AI/ML perspective to a more bottom-up, or even better, forward–backward analysis we can not only improve our predictive models, but also potentially discover unknown phenomena. In this example, there appears to be at least a statistical relationship between California and the Middle East precipitation. The literature links the two regions to ENSO, but the mechanics of ENSO influencing precipitation in the Middle East is not well understood. Bottom-up learning algorithms can be used to extract interesting patterns of this kind for further future research. While the example here, and the focus of our paper, is on drought forecasting, one can argue that this applies to other types of predictive models.

6. Final remarks

The purpose of this article is not to discount the significant progress we have made in the area of drought forecasting, especially over the past two decades. By focusing on the current limitations, we aim to highlight important research gaps and key opportunities to improve the current drought forecasting models. For this reason, much of this article focuses on what we can do to potentially transform the next generation of drought forecasting models.

Understandably, decision-makers prefer inaccurate forecasts rather than no forecasts at all [192]. This is not unusual and is not limited to the area of drought forecasting. However, providing inaccurate forecasts is an ever-present concern since forecasts can embolden communities to take on increased risk or not. Recall the US Midwest Drought of 2012 we discuss at the beginning of this article. One could argue that if the predictions were not favourable, farmers would not have planted a record area of cropland. We also believe that having inaccurate or sometimes accurate forecasts is better than no forecast. However, we emphasize that what is missing in our current drought forecasting models is not simply the magnitude of the uncertainty, but a lack of awareness of it. More efforts should focus on quantifying the uncertainty beyond just an ensemble of model simulations.

A shift in our focus toward quantifying the stability of uncertainty in drought prediction models, rather than the goodness of fit or reproducing the past, could be a first step toward testing the robustness of our theories. Instead of solely focusing on improving predictions (e.g. increasing the lead time by a month or two for specific events), we have to put our efforts into improving the robustness of our theories. Some relevant questions to address this issue include: Does the theory behind our predictive model offer similarly reliable predictions in the same location but over different time periods? How does the robustness or uncertainty change if the model is applied to different locations (or at least climatically similar regions)? Quantifying the robustness of uncertainty requires both basic and applied research tailored to drought forecasting, including indicators with realistic measures of uncertainty.

Current theories, including known teleconnections, commonly used for drought prediction are top-down. These theories assume that the process(es) and/or driver(s) for drought forecasting are known. In other words, current approaches start with a model and impose it on the observed events (reality). We believe that using a different perspective may be advantageous to further improving drought forecasting models. Instead of a top-down approach, a bottom-up perspective can be used that starts from reality (here, observed events) and searches for model(s) and driver(s) that work for predicting drought. The idea of EmpTe is one example of how to look at drought prediction from a bottom-up viewpoint. A seemingly reasonable EmpTe will remain empirical until the physics behind it is discovered. An EmpTe that is merely a statistical flaw will be eliminated when tested for robustness and uncertainty across space and time. More efforts in further developing bottom-up approaches, particularly using recent advances in artificial intelligence and machine learning, could potentially even lead to the discovery of new physically explainable teleconnections or drivers.

Finally, we need move towards what we refer to as 'impact-based drought prediction' by linking our drought forecasting tools with drought impact assessment models. Storyline-based concepts [194,195] especially when centred on actual or even expected drought impacts can be the first step toward developing impact-based drought prediction models. We acknowledge that our current drought prediction models are far from ideal and have limited long-range predictability. However, even a short-rage impact-based model (e.g. expected crop failure, water shortage, forest fires activity) can have significant impacts on emergency response and planning.

Data accessibility. All datasets are publicly available. Data sources are acknowledged.

Authors' contributions. A.A.: conceptualization, funding acquisition, project administration, writing—original draft; B.P.: visualization, writing—review and editing; O.M.: visualization, writing—review and editing; M.S.: writing—review and editing; S.J.: writing—review and editing; W..: visualization, writing—review and editing; C.L.: writing—review and editing; S.M.: writing—review and editing; S.M.P.: writing—review and editing; S.J.D.: writing—review and editing; K.H.: writing—review and editing; S.S.: writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This study is partially supported by the NSF INFEWS Grant nos. EAR 1639318 and OISE 2114701, National Oceanic and Atmospheric Administration (NOAA), Modeling, Analysis, Prediction, and Projections program (MAPP) Award NA19OAR4310294 and the National Science Foundation (NSF) Award OAC-1931335.

Appendix A

For the experiment discussed in this article, we obtained the sea surface temperature (SST) dataset from the NOAA Optimum Interpolation Sea Surface Temperature (OISST) V2 (http://www.esrl.noaa.gov/psd/). The OISST dataset has a temporal resolution of one month, and a spatial resolution of $1^\circ \times 1^\circ$ [196]. We used the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks – Climate Data Record (PERSIANN-CDR; [197]). The data are available at a $0.25^\circ \times 0.25^\circ$ spatial resolution from: http://chrsdata.eng.uci.edu.

We determined the empirical teleconnection by finding the two to n locations where the three-month average of SSTs had the highest correlation with California's wet-season, using a brute-force (exhaustive) search algorithm [198]. We considered the wet season as the six highest consecutive months of cumulative precipitation. In this study, n represents the number of grid cells used to create our empirical teleconnection (here up to 10 grid cells). We placed a temporal constraint where the SSTs must have a lead time of between 3 and 12 months, as we are interested in drought forecasting. Once we determined the SST locations, the Raw Empirical Teleconnection ($\mathrm{EmpTe_R}$):

$$\mathrm{EmpTe}_e = (\pm\mathrm{SST1}) + (\pm\mathrm{SST2}) + (\pm\mathrm{SST3}) + (\pm\mathrm{SST4}) + \cdots + (\pm\mathrm{SST}n),$$

where SST1, SST2, SST3, SST4 and SSTn are the three-month averaged SST values, and n is the number of grid cells we use to derive the empirical teleconnection. We multiply the SST value by +1 if the correlation with wet-season precipitation is positive, and by −1 if the correlation is negative. The number of points used in the empirical teleconnection is determined based on which combination of SSTs has the highest correlation. Then we standardize EmpTe_R as follows:

$$\mathrm{EmpTe} = \frac{\mathrm{EmpTe}_R - \mathrm{mean}(\mathrm{EmpTe}_R)}{\mathrm{stdev}(\mathrm{EmpTe}_R)}.$$

We similarly perform a temporal search for the well-known teleconnections (i.e. Multivariate Enso Index (MEI), Southern Oscillation Index (SOI), PDO, Pacific/North American teleconnection pattern (PNA) and Arctic Oscillation (AO)) to find the years with the highest correlation with precipitation, and standardize the mean monthly values. We define the conditional probability of wet hit as

$$\Pr(\mathrm{Precip}_{\mathrm{wet}} > P_T \mid \mathrm{var}_m > \mathrm{var}_T) = \frac{\Pr(\mathrm{Precip}_{\mathrm{wet}} > P_T \cap \mathrm{var}_{\mathrm{wet}-m} > \mathrm{var}_T)}{\Pr(\mathrm{var}_{\mathrm{wet}-m} > \mathrm{var}_T)} m = 3 - 12,$$

and dry hit as

$$\Pr(\mathrm{Precip}_{\mathrm{wet}} < P_T \mid \mathrm{var}_m < \mathrm{var}_T) = \frac{\Pr(\mathrm{Precip}_{\mathrm{wet}} < P_T \cap \mathrm{var}_{\mathrm{wet}-m} < \mathrm{var}_T)}{\Pr(\mathrm{var}_{\mathrm{wet}-m} < var_T)} m = 3 - 12,$$

where $\mathrm{Precip}_{\mathrm{wet}}$ is the sum of precipitation during the wet season (e.g. NDJFMA), P_{T} is the threshold for (wet or dry) precipitation, var is the standardized mean monthly values of our variable, $\mathrm{var}_{\mathrm{T}}$ is the threshold for (wet or dry) var, and m the best 3 months to calculate the empirical teleconnection with a lead time of 3–12 months from the first month of the wet season. The term m refers to the lead time , where the algorithm chooses the lead time between 3 and 12 months with the highest correlation in each gridbox.

We compared the prediction skill of our empirical teleconnection with that of other teleconnections. Model skill is based on the mutual information between the prediction and observation, and measures the dependence between two variables, defined as

$$\mathrm{MI}(\boldsymbol{x}_{t+\tau}, \boldsymbol{o}_t) = \int\int p(\boldsymbol{x}_{t+\tau}, \boldsymbol{o}_t) \log\left[\frac{p(\boldsymbol{x}_{t+\tau}, \boldsymbol{o}_t)}{p(\boldsymbol{o}_t)p(\boldsymbol{x}_{t+\tau})}\right] \mathrm{d}\boldsymbol{x}_{t+\tau} \mathrm{d}\boldsymbol{o}_t,$$

where $p(\boldsymbol{x}_{t+\tau}, \boldsymbol{o}_t)$ is the joint distribution between the prediction ($\mathbf{x}$) and the observation ($\mathbf{o}$), assuming $p(\boldsymbol{x}, \boldsymbol{o}) = p(\boldsymbol{x} \mid \boldsymbol{o})p(\boldsymbol{o})$. In the case that the observations are perfect ($\boldsymbol{o}_t = \boldsymbol{x}_t$), then $I(\boldsymbol{x}_{t+\tau}, \boldsymbol{o}_t) = I(\boldsymbol{x}_{t+\tau}, \boldsymbol{x}_t)$ [199].

Since mutual information is unbounded, we can convert the measure into a 'score' between 0 and 1 using the following transformation [200]:

$$\text{skill score} = 1 - e^{-2\mathrm{MI}}.$$

We present the results of the empirical teleconnection using a leave-one-out cross-validation (LOO-CV) approach. To ensure that the model predictions are independent of the observations for a given year, we hold out the observation data for the prediction year from the calibration data. This LOO-CV process is repeated for all the years throughout the time series.

References

1. USDA. 2012 U.S. Drought 2012: Farm and Food Impacts. See https://web.archive.org/web/20130116030109/http://www.ers.usda.gov/topics/in-the-news/us-drought-2012-farm-and-food-impacts.aspx.
2. Hoerling M *et al.* 2013 Anatomy of an extreme event. *J. Clim.* **26**, 2811–2832. (doi:10.1175/JCLI-D-12-00270.1)
3. Hoerling M, Eischeid J, Kumar A, Leung R, Mariotti A, Mo K, Schubert S, Seager R. 2014 Causes and predictability of the 2012 Great Plains drought. *Bull. Am. Meteorol. Soc.* **95**, 269–282. (doi:10.1175/BAMS-D-13-00055.1)
4. AghaKouchak A. 2014 A baseline probabilistic drought forecasting framework using standardized soil moisture index: application to the 2012 United States drought. *Hydrol. Earth Syst. Sci.* **18**, 2485–2492. (doi:10.5194/hess-18-2485-2014)

5. Kiem AS *et al.* 2016 Natural hazards in Australia: droughts. *Clim. Change* **139**, 37–54. (doi:10.1007/s10584-016-1798-7)
6. AghaKouchak A. 2015 A multivariate approach for persistence-based drought prediction: application to the 2010–2011 East Africa drought. *J. Hydrol.* **526**, 127–135. (doi:10.1016/j.jhydrol.2014.09.063)
7. Vicente-Serrano SM *et al.* 2012 Challenges for drought mitigation in Africa: the potential use of geospatial data and drought information systems. *Appl. Geogr.* **34**, 471–486. (doi:10.1016/j.apgeog.2012.02.001)
8. Steinemann AC. 2006 Using climate forecasts for drought management. *J. Appl. Meteorol. Climatol.* **45**, 1353–1361. (doi:10.1175/JAM2401.1)
9. Acácio V *et al.* 2013 *Review of current drought monitoring systems and identification of (further) monitoring requirements. DROUGHT-R&SPI Technical Report (6)*. Wageningen, The Netherlands: Wageningen Environmental Research.
10. Wood AW, Lettenmaier DP. 2006 A test bed for new seasonal hydrologic forecasting approaches in the western United States. *Bull. Am. Meteorol. Soc.* **87**, 1699–1712. (doi:10.1175/BAMS-87-12-1699)
11. Wood AW. 2008 The University of Washington Surface Water Monitor: an experimental platform for national hydrologic assessment and prediction. In *22nd Conf. on Hydrology,New Orleans, LA, 20–24 January*. Boston, MA : American Meteorological Society.
12. Luo L, Wood EF. 2007 Monitoring and predicting the 2007 US drought. *Geophys. Res. Lett.* **34**, L22702. (doi:10.1029/2007GL031673)
13. Li H, Luo L, Wood EF. 2008 Seasonal hydrologic predictions of low-flow conditions over eastern USA during the 2007 drought. *Atmos. Sci. Lett.* **9**, 61–66. (doi:10.1002/asl.182)
14. Sheffield J *et al.* 2014 A drought monitoring and forecasting system for sub-Sahara African water resources and food security. *Bull. Am. Meteorol. Soc.* **95**, 861–882. (doi:10.1175/BAMS-D-12-00124.1)
15. Lyon B, Bell MA, Tippett MK, Kumar A, Hoerling MP, Quan X-W, Wang H. 2012 Baseline probabilities for the seasonal prediction of meteorological drought. *J. Appl. Meteorol. Climatol.* **51**, 1222–1237. (doi:10.1175/JAMC-D-11-0132.1)
16. Svoboda M *et al.* 2002 The drought monitor. *Bull. Am. Meteorol. Soc.* **83**, 1181–1190. (doi:10.1175/1520-0477-83.8.1181)
17. Hao Z, AghaKouchak A, Nakhjiri N, Farahmand A. 2014 Global integrated drought monitoring and prediction system. *Sci. Data* **1**, 1–10. (doi:10.1038/sdata.2014.1)
18. Moura AD, Sarachik ES. 1997 Seasonal-to-interannual climate prediction and applications: new institutions, new possibilities. *Bull. World Meteorol. Organ.* **46**, 342–347.
19. Mamalakis A, Yu J-Y, Randerson JT, AghaKouchak A, Foufoula-Georgiou E. 2018 A new interhemispheric teleconnection increases predictability of winter precipitation in southwestern US. *Nat. Commun.* **9**, 2332. (doi:10.1038/s41467-018-04722-7)
20. Mundhenk BD, Barnes EA, Maloney ED, Baggett CF. 2018 Skillful empirical subseasonal prediction of landfalling atmospheric river activity using the Madden– Julian oscillation and quasi-biennial oscillation. *NPJ Clim. Atmos. Sci.* **1**, 7. (doi:10.1038/s41612-017-0008-2)
21. Whitaker JS, Weickmann KM. 2001 Subseasonal variations of tropical convection and week-2 prediction of wintertime western North American rainfall. *J. Clim.* **14**, 3279–3288. (doi:10.1175/1520-0442(2001)014<3279:SVOTCA>2.0.CO;2)
22. Hoskins B. 2013 The potential for skill across the range of the seamless weather-climate pre- diction problem: a stimulus for our science. *Quart. J. Roy. Meteor. Soc.* **139**, 573–584. (doi:10.1002/qj.1991)
23. Hoskins BJ, Karoly DJ. 1981 The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atmos. Sci.* **38**, 1179–1196. (doi:10.1175/1520-0469(1981)038<1179:TSLROA>2.0.CO;2)
24. Matthews AJ. 2004 Atmospheric response to observed intraseasonal tropical sea surface temperature anomalies. *Geophys. Res. Lett.* **31**, L14107. (doi:10.1029/2004GL020474)
25. Sardeshmukh PD, Hoskins BJ. 1988 The generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.* **45**, 1228–1251. (doi:10.1175/1520-0469(1988)045<1228:TGOGRF>2.0.CO;2)
26. Pan B, Hsu K, AghaKouchak A, Sorooshian S. 2019 Improving precipitation estimation using convolutional neural network. *Water Resour. Res.* **55**, 2301–2321. (doi:10.1029/2018WR024090)

27. Prein AF *et al.* 2015 A review on regional convection-permitting climate modeling: demonstrations, prospects, and challenges. *Rev. Geophys.* **53**, 323–361. (doi:10.1002/2014RG000475)
28. Huang Y, Chen Z, Tao Y, Huang X, Gu X. 2018 Agricultural remote sensing big data: management and applications. *J. Integr. Agric.* **17**, 1915–1931. (doi:10.1016/S2095-3119(17)61859-8)
29. Hao Z, Singh VP, Xia Y. 2018 Seasonal drought prediction: advances, challenges, and future prospects. *Rev. Geophys.* **56**, 108–141. (doi:10.1002/2016RG000549)
30. WCRP 2010 A WCRP white paper on drought predictability and prediction in a changing climate: assessing current predictive knowledge and capabilities, user requirements and research priorities Technical Report (Geneva: World Climate Research Programme)
31. Ghil M. 2001 Hilbert problems for the geosciences in the 21st century. *Nonlinear Processes Geophys.* **8**, 211–211. (doi:10.5194/npg-8-211-2001)
32. Mishra AK, Desai VR. 2005 Drought forecasting using stochastic models. *Stochast. Environ. Res. Risk Assessment* **19**, 326–339. (doi:10.1007/s00477-005-0238-4)
33. Han P, Wang PX, Zhang SY. 2010 Drought forecasting based on the remote sensing data using ARIMA models. *Math. Computer Model.* **51**, 1398–1403. (doi:10.1016/j.mcm.2009.10.031)
34. Abbe C. 1901 The physical basis of long-range weather forecasts. *Mon. Weather Rev.* **29**, 551–561. (doi:10.1175/1520-0493(1901)29[551c:TPBOLW]2.0.CO;2)
35. Merryfield WJ *et al.* 2013 The Canadian seasonal to interannual prediction system. Part I: models and initialization. *Mon. Weather Rev.* **141**, 2910–2945. (doi:10.1175/MWR-D-12-00216.1)
36. Saha S *et al.* 2014 The NCEP climate forecast system version 2. *J. Clim.* **27**, 2185–2208. (doi:10.1175/JCLI-D-12-00823.1)
37. Zhang S, Harrison M, Rosati A, Wittenberg A. 2007 System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Mon. Weather Rev.* **135**, 3541–3564. (doi:10.1175/MWR3466.1)
38. Alves O, Wang G, Zhong A, Smith N, Tseitkin F, Warren G, Schiller A, Godfrey S, Meyers G. 2003 POAMA: Bureau of Meteorology operational coupled model seasonal forecast system. In *Proc. of National Drought Forum*, *Brisbane*, *Queensland*, *Australia*, pp. 49–56. Brisbane, Australia: DPI Publications.
39. Berner J *et al.* 2017 Stochastic parameterization: toward a new view of weather and climate models. *Bull. Am. Meteorol. Soc.* **98**, 565–588. (doi:10.1175/BAMS-D-15-00268.1)
40. Baggett CF, Barnes EA, Maloney ED, Mundhenk BD. 2017 Advancing atmospheric river forecasts into subseasonal-to-seasonal time scales. *Geophys. Res. Lett.* **44**, 7528–7536. (doi:10.1002/2017GL074434)
41. Best M *et al.* 2011 The Joint UK Land Environment Simulator (JULES), model description–Part 1: energy and water fluxes. *Geosci. Model Dev.* **4**, 677–699. (doi:10.5194/gmd-4-677-2011)
42. Bjerknes J. 1969 Atmospheric teleconnections from the equatorial Pacific. *Mon. Wea. Rev.* **97**, 163–172. (doi:10.1175/1520-0493(1969)097<0163:ATFTEP>2.3.CO;2)
43. Schneider T, Lan S, Stuart A, Teixeira J. 2017 Earth system modeling 2.0: a blueprint for models that learn from observations and targeted high-resolution simulations. *Geophys. Res. Lett.* **44**, 12–396. (doi:10.1002/2017GL076101)
44. Pan B. 2019 *Advancing precipitation prediction using a composite of models and data*. Irvine, CA: University of California.
45. Lorenz EN. 1963 Deterministic nonperiodic flow. *J. Atmosph. Sci.* **20**, 130–141. (doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2)
46. Shukla S, McNally A, Husak G, Funk C. 2014 A seasonal agricultural drought forecast system for food-insecure regions of East Africa. *Hydrol. Earth Syst. Sci.* **18**, 3907–3921. (doi:10.5194/hess-18-3907-2014)
47. Hao Z, Yuan X, Xia Y, Hao F, Singh VP. 2017 An overview of drought monitoring and prediction systems at regional and global scales. *Bull. Am. Meteorol. Soc.* **98**, 1879–1896. (doi:10.1175/BAMS-D-15-00149.1)
48. Ahmadalipour A, Moradkhani H, Yan H, Zarekarizi M. 2017 Remote sensing of drought: vegetation, soil moisture, and data assimilation. In *Remote sensing of hydrological extremes* (ed. V Lakshmi), pp. 121–149. Cham, Switzerland: Springer.

49. Xu L, Abbaszadeh P, Moradkhani H, Chen N, Zhang X. 2020 Continental drought monitoring using satellite soil moisture, data assimilation and an integrated drought index. *Remote Sens. Environ.* **250**, 112028. (doi:10.1016/j.rse.2020.112028)
50. Kumar SV *et al.* 2014 Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation. *J. Hydrometeorol.* **15**, 2446–2469. (doi:10.1175/JHM-D-13-0132.1)
51. Khairoutdinov M, Randall D, DeMott C. 2005 Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes. *J. Atmos. Sci.* **62**, 2136–2154. (doi:10.1175/JAS3453.1)
52. Khairoutdinov MF, Randall DA. 2001 A cloud resolving model as a cloud parameterization in the NCAR Community Climate System Model: Preliminary results. *Geophys. Res. Lett.* **28**, 3617–3620. (doi:10.1029/2001GL013552)
53. Kumar P, Rupa Kumar K, Rajeevan M, Sahai A. 2007 On the recent strengthening of the relationship between ENSO and northeast monsoon rainfall over South Asia. *Clim. Dyn.* **28**, 649–660. (doi:10.1007/s00382-006-0210-0)
54. Wang C, Kucharski F, Barimalala R, Bracco A. 2009 Teleconnections of the tropical Atlantic to the tropical Indian and Pacific Oceans: a review of recent findings. *Meteorol. Z.* **18**, 445–454. (doi:10.1127/0941-2948/2009/0394)
55. Cohen J, Jones J. 2011 A new index for more accurate winter predictions. *Geophys. Res. Lett.* **38**, 3617–3620. (doi:10.1029/2011GL049626)
56. Zhao T, Chen H, Tian Y, Chen X. 2022 Quantifying overlapping and differing information of global precipitation for GCM forecasts and El Niño–Southern Oscillation. *Hydrol. Earth System Sci.* **26**, 4233–4249. (doi:10.5194/hess-26-4233-2022)
57. Hewitt CD. 2004 *Ensembles-based predictions of climate changes and their impacts*. New York, NY: Wiley Online Library.
58. Palmer TN *et al.* 2004 Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull. Am. Meteorol. Soc.* **85**, 853–872. (doi:10.1175/BAMS-85-6-853)
59. Bougeault P *et al.* 2010 The THORPEX interactive grand global ensemble. *Bull. Am. Meteorol. Soc.* **91**, 1059–1072. (doi:10.1175/2010BAMS2853.1)
60. Kirtman BP *et al.* 2014 The North American multimodel ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bull. Am. Meteorol. Soc.* **95**, 585–601. (doi:10.1175/BAMS-D-12-00050.1)
61. Becker E, den Dool Hv, Zhang Q. 2014 Predictability and forecast skill in NMME. *J. Clim.* **27**, 5891–5906. (doi:10.1175/JCLI-D-13-00597.1)
62. Ma S, Zhou T, Dai A, Han Z. 2015 Observed changes in the distributions of daily precipitation frequency and amount over China from 1960 to 2013. *J. Clim.* **28**, 6960–6978. (doi:10.1175/JCLI-D-15-0011.1)
63. Alley RB, Emanuel KA, Zhang F. 2019 Advances in weather prediction. *Science* **363**, 342–344. (doi:10.1126/science.aav7274)
64. Bauer P, Thorpe A, Brunet G. 2015 The quiet revolution of numerical weather prediction. *Nature* **525**, 47–55. (doi:10.1038/nature14956)
65. Wood EF, Schubert SD, Wood AW, Peters-Lidard CD, Mo KC, Mariotti A, Pulwarty RS. 2015 Prospects for advancing drought understanding, monitoring, and prediction. *J. Hydrometeorol.* **16**, 1636–1657. (doi:10.1175/JHM-D-14-0164.1)
66. Kim H-M, Webster PJ, Curry JA. 2012 Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts. *Geophys. Res. Lett.* **39**, L10701. (doi:10.1029/2012GL051644)
67. Infanti JM, Kirtman BP. 2016 North American rainfall and temperature prediction response to the diversity of ENSO. *Clim. Dyn.* **46**, 3007–3023. (doi:10.1007/s00382-015-2749-0)
68. Yang S-C, Keppenne C, Rienecker M, Kalnay E. 2009 Application of coupled bred vectors to seasonal-to-interannual forecasting and ocean data assimilation. *J. Clim.* **22**, 2850–2870. (doi:10.1175/2008JCLI2427.1)
69. Jia L *et al.* 2015 Improved seasonal prediction of temperature and precipitation over land in a high-resolution GFDL climate model. *J. Clim.* **28**, 2044–2062. (doi:10.1175/JCLI-D-14-00112.1)
70. Mishra AK, Singh VP. 2010 A review of drought concepts. *J. Hydrol.* **391**, 202–216. (doi:10.1016/j.jhydrol.2010.07.012)

71. McKee TB, Doesken NJ, Kleist J. 1993 The relationship of drought frequency and duration to time scales. In *Proc. of the 8th Conf. on Applied Climatology, Anaheim, CA, 17–22 January*, pp. 179–183.
72. Belayneh A, Adamowski J, Khalil B, Ozga-Zielinski B. 2014 Long-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet neural network and wavelet support vector regression models. *J. Hydrol.* **508**, 418–429. (doi:10.1016/j.jhydrol.2013.10.052)
73. Karl TR. 1983 Some spatial characteristics of drought duration in the United States. *J. Appl. Meteorol. Climatol.* **22**, 1356–1366. (doi:10.1175/1520-0450(1983)022<1356:SSCODD>2.0.CO;2)
74. Hao Z, Hao F, Singh VP, Sun AY, Xia Y. 2016a Probabilistic prediction of hydrologic drought using a conditional probability approach based on the meta-Gaussian model. *J. Hydrol.* **542**, 772–780. (doi:10.1016/j.jhydrol.2016.09.048)
75. McCABE GJ, Dettinger MD. 1999 Decadal variations in the strength of ENSO teleconnections with precipitation in the western United States. *Int. J. Climatol. J. R. Meteorol. Soc.* **19**, 1399–1410. (doi:10.1002/(SICI)1097-0088(19991115)19:13<1399::AID-JOC457>3.0.CO;2-A)
76. Bradley R, Diaz H, Eischeid J, Jones P, Kelly P, Goodess C. 1987 Precipitation fluctuations over Northern Hemisphere land areas since the mid-19th century. *Science* **237**, 171–175. (doi:10.1126/science.237.4811.171)
77. Ropelewski CF, Halpert MS. 1986 North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO). *Mon. Weather Rev.* **114**, 2352–2362. (doi:10.1175/1520-0493(1986)114<2352:NAPATP>2.0.CO;2)
78. Wu B, Zhou T, Li T. 2009 Seasonally evolving dominant interannual variability modes of East Asian climate. *J. Clim.* **22**, 2992–3005. (doi:10.1175/2008JCLI2710.1)
79. Li X, Zhang K, Gu P, Feng H, Yin Y, Chen W, Cheng B. 2021 Changes in precipitation extremes in the Yangtze River Basin during 1960–2019 and the association with global warming, ENSO, and local effects. *Sci. Total Environ.* **760**, 144244. (doi:10.1016/j.scitotenv.2020.144244)
80. Peng Y, Shen C, Cheng H, Xu Y. 2014 Modeling of severe persistent droughts over eastern China during the last millennium. *Clim. Past* **10**, 1079–1091. (doi:10.5194/cp-10-1079-2014)
81. Schepen A, Wang Q, Robertson D. 2012 Evidence for using lagged climate indices to forecast Australian seasonal rainfall. *J. Clim.* **25**, 1230–1246. (doi:10.1175/JCLI-D-11-00156.1)
82. Hartmann H, Becker S, King L. 2008 Quasi-periodicities in Chinese precipitation time series. *Theor. Appl. Climatol.* **92**, 155–163. (doi:10.1007/s00704-007-0317-1)
83. Fung KF, Huang YF, Koo CH, Soh YW. 2020 Drought forecasting: a review of modelling approaches 2007–2017. *J. Water Clim. Change* **11**, 771–799. (doi:10.2166/wcc.2019.236)
84. Serinaldi F, Bonaccorso B, Cancelliere A, Grimaldi S. 2009 Probabilistic characterization of drought properties through copulas. *Phys. Chem. Earth Parts a/B/C* **34**, 596–605. (doi:10.1016/j.pce.2008.09.004)
85. Dehghani M, Saghafian B, Zargar M. 2019 Probabilistic hydrological drought index forecasting based on meteorological drought index using Archimedean copulas. *Hydrol. Res.* **50**, 1230–1250. (doi:10.2166/nh.2019.051)
86. Pontes Filho JD, Portela MM, Marinho de Carvalho Studart T, Souza Filho FDA. 2019 A continuous drought probability monitoring system, CDPMS, based on copulas. *Water* **11**, 1925. (doi:10.3390/w11091925)
87. Wu H, Su X, Singh VP, Zhang T, Qi J, Huang S. 2022 Comparison between canonical vine copulas and a meta-Gaussian model for forecasting agricultural drought over China. *Hydrol. Earth Syst. Sci.* **26**, 3847–3861. (doi:10.5194/hess-26-3847-2022)
88. Alijanian M, Rakhshandehroo GR, Dehghani M, Mishra A. In press. Probabilistic drought forecasting using copula and satellite rainfall based PERSIANN-CDR and MSWEP datasets. *Int. J. Climatol.* (doi:10.1002/joc.7600)
89. Wu H, Su X, Singh VP, Feng K, Niu J. 2021 Agricultural drought prediction based on conditional distributions of vine copulas. *Water Resour. Res.* **57**, e2021WR029562. (doi:10.1029/2021WR029562)
90. Modarres R. 2007 Streamflow drought time series forecasting. *Stochast. Environ. Res. Risk Assess.* **21**, 223–233. (doi:10.1007/s00477-006-0058-1)

91. Rust W, Holman I, Bloomfield J, Cuthbert M, Corstanje R. 2019 Understanding the potential of climate teleconnections to project future groundwater drought. *Hydrol. Earth Syst. Sci.* **23**, 3233–3245. (doi:10.5194/hess-23-3233-2019)
92. Lima CH, AghaKouchak A. 2017 Droughts in Amazonia: spatiotemporal variability, teleconnections, and seasonal predictions. *Water Resour. Res.* **53**, 10 824–10 840. (doi:10.1002/2016WR020086)
93. Enfield DB, Alfaro EJ. 1999 The dependence of Caribbean rainfall on the interaction of the tropical Atlantic and Pacific Oceans. *J. Clim.* **12**, 2093–2103. (doi:10.1175/1520-0442(1999)012<2093:TDOCRO>2.0.CO;2)
94. Luo JJ, Masson S, Behera S, Shingu S, Yamagata T. 2005 Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble forecasts. *J. Clim.* **18**, 4474–4497. (doi:10.1175/JCLI3526.1)
95. Giannini A, Saravanan R, Chang P. 2003 Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales. *Science* **302**, 1027–1030. (doi:10.1126/science.1089357)
96. Schubert SD *et al.* 2016 Global meteorological drought: a synthesis of current understanding with a focus on SST drivers of precipitation deficits. *J. Clim.* **29**, 3989–4019. (doi:10.1175/JCLI-D-15-0452.1)
97. Mantua NJ, Hare SR. 2002 The Pacific decadal oscillation. *J. Oceanogr.* **58**, 35–44. (doi:10.1023/A:1015820616384)
98. Enfield DB, Mestas-Nuñez AM, Trimble PJ. 2001 The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US. *Geophys. Res. Lett.* **28**, 2077–2080. (doi:10.1029/2000GL012745)
99. Nikraftar Z, Parizi E, Hosseini SM, Ataie-Ashtiani B. 2021 Lake Urmia restoration success story: a natural trend or a planned remedy? *J. Great Lakes Res.* **47**, 955–969. (doi:10.1016/j.jglr.2021.03.012)
100. Kallis G. 2008 Droughts. *Annu. Rev. Environ. Resour.* **33**, 85–118. (doi:10.1146/annurev.environ.33.081307.123117)
101. Anshuka A, van Ogtrop FF, Vervoort RW. 2019 Drought forecasting through statistical models using standardised precipitation index: a systematic review and meta-regression analysis. *Nat. Hazards* **97**, 955–977. (doi:10.1007/s11069-019-03665-6)
102. Hoerling M, Kumar A. 2003 The perfect ocean for drought. *Science* **299**, 691–694. (doi:10.1126/science.1079053)
103. Seager R, Hoerling M. 2014 Atmosphere and ocean origins of North American droughts. *J. Clim.* **27**, 4581–4606. (doi:10.1175/JCLI-D-13-00329.1)
104. Chiew FH, Piechota TC, Dracup JA, McMahon TA. 1998 El Nino/Southern Oscillation and Australian rainfall, streamflow and drought: links and potential for forecasting. *J. Hydrol.* **204**, 138–149. (doi:10.1016/S0022-1694(97)00121-2)
105. Wilks DS. 2011 *Statistical methods in the atmospheric sciences* (vol. 100). New York, NY: Academic Press.
106. Van Der Maaten L, Postma E, Van den Herik J. 2009 Dimensionality reduction: a comparative. *J. Mach. Learn. Res.* **10**, 13.
107. Chen J, Yang Y. 2012 SPI-based regional drought prediction using weighted Markov Chain model. *Res. J. Appl. Sci. Eng. Technol.* **4**, 4293–4298.
108. Nnaji GA, Clark CJ, Chan-Hilton AB, Huang W. 2016 Drought prediction in Apalachicola–Chattahoochee–flint river basin using a semi-Markov model. *Nat. Hazards* **82**, 267–297. (doi:10.1007/s11069-016-2201-8)
109. Sharma TC, Panu US. 2012 Prediction of hydrological drought durations based on Markov chains: case of the Canadian prairies. *Hydrol. Sci. J.* **57**, 705–722. (doi:10.1080/02626667.2012.672741)
110. Cho WKT, Liu YY. 2018 Sampling from complicated and unknown distributions: Monte Carlo and Markov Chain Monte Carlo methods for redistricting. *Physica A* **506**, 170–178. (doi:10.1016/j.physa.2018.03.096)
111. Eddy SR. 2004 What is a hidden Markov model? *Nat. Biotechnol.* **22**, 1315–1316. (doi:10.1038/nbt1004-1315)
112. Thomas JA, Fiering MB. 1962 Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation. In *Design of water-resource systems* (eds A Maass, M Maynard, RD Hufschmidt, HA Thomas Jr, SA Marglin, G. Maskew Fair), pp. 459–1549; 493. Cambridge, MA: Harvard University Press.

113. Box GEP, Jenkins GM. 1970 *Time series analysis: forecasting and control*. San Francisco, CA: Holden Day.
114. Box GE, Jenkins GM, Reinsel GC, Ljung GM. 2015 *Time series analysis: forecasting and control*. New York, NY: John Wiley & Sons.
115. Papalexiou SM. 2022 Rainfall generation revisited: introducing CoSMoS-2s and advancing copula-based intermittent time series modeling. *Water Resour. Res.* **58**, e2021WR031641. (doi:10.1029/2021WR031641)
116. Tosunoğlu F, Onof C. 2017 Joint modelling of drought characteristics derived from historical and synthetic rainfalls: application of generalized linear models and copulas. *J. Hydrol. Reg. Stud.* **14**, 167–181. (doi:10.1016/j.ejrh.2017.11.001)
117. Morid S, Smakhtin V, Bagherzadeh K. 2007 Drought forecasting using artificial neural networks and time series of drought indices. *Int. J. Climatol J. R. Meteorol. Soc.* **27**, 2103–2111. (doi:10.1002/joc.1498)
118. Mishra AK, Desai VR. 2005 Drought forecasting using stochastic models. *Stochastic environmental research and risk assessment* **19**, 326–339.
119. Stagge JH, Kohn I, Tallaksen LM, Stahl K. 2015 Modeling drought impact occurrence based on meteorological drought indices in Europe. *J. Hydrol.* **530**, 37–50. (doi:10.1016/j.jhydrol.2015.09.039)
120. Aiyelokun O, Ogunsanwo G, Ojelabi A, Agbede O. 2021 Gaussian Naïve Bayes classification algorithm for drought and flood risk reduction. In *Intelligent data analytics for decision-support systems in hazard mitigation*, pp. 49–62. Berlin, Germany: Springer.
121. Shin JY, Ajmal M, Yoo J, Kim T-W. 2016 A Bayesian network-based probabilistic framework for drought forecasting and outlook. *Adv. Meteorol.* **2016**, 9472605. (doi:10.1155/2016/9472605)
122. Priya R, Ramesh D, Khosla E. 2018 Crop prediction on the region belts of India: a naïve Bayes MapReduce precision agricultural model. In *2018 Int. Conf. on Advances in Computing, Communications and Informatics (ICACCI), New Orleans, LA, 20–24 January 2008*, pp. 99–104. Boston, MA: American Meteorological Society. (doi:10.1109/ICACCI.2018.8554948)
123. Phan TD, Smart JCR, Capon SJ, Hadwen WL, Sahin O. 2016 Applications of Bayesian belief networks in water resource management: a systematic review. *Environ. Model. Software* **85**, 98–111. (doi:10.1016/j.envsoft.2016.08.006)
124. Xu L, Chen N, Zhang X, Chen Z. 2018 An evaluation of statistical, NMME and hybrid models for drought prediction in China. *J. Hydrol.* **566**, 235–249. (doi:10.1016/j.jhydrol.2018.09.020)
125. Thober S, Kumar R, Sheffield J, Mai J, Schäfer D, Samaniego L. 2015 Seasonal soil moisture drought prediction over Europe using the North American Multi-Model Ensemble (NMME). *J. Hydrometeorol.* **16**, 2329–2344. (doi:10.1175/JHM-D-15-0053.1)
126. Zadeh LA. 1965 Fuzzy sets. *Inform. Control* **8**, 338–353. (doi:10.1016/S0019-9958(65)90241-X)
127. Agboola AH, Gabriel AJ, Aliyu EO, Alese BK. 2013 Development of a fuzzy logic based rainfall prediction model. *Int. J. Eng. Technol.* **3**, 427–435.
128. Chang F-J, Chen Y-C. 2001 A counterpropagation fuzzy-neural network modeling approach to real time streamflow prediction. *J. Hydrol.* **245**, 153–164. (doi:10.1016/S0022-1694(01)00350-X)
129. Özger M, Mishra AK, Singh VP. 2012 Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in Texas. *J. Hydrometeorol.* **13**, 284–297. (doi:10.1175/JHM-D-10-05007.1)
130. Kaur A, Sood SK. 2020 Cloud-fog based framework for drought prediction and forecasting using artificial neural network and genetic algorithm. *J. Exp. Theor. Artif. Intell.* **32**, 273–289. (doi:10.1080/0952813X.2019.1647563)
131. Merabtene T, Kawamura A, Jinno K, Olsson J. 2002 Risk assessment for optimal drought management of an integrated water resources system using a genetic algorithm. *Hydrol. Processes* **16**, 2189–2208. (doi:10.1002/hyp.1150)
132. Omidvar E, Tahroodi ZN. 2019 Evaluation and prediction of meteorological drought conditions using time-series and genetic programming models. *J. Earth Syst. Sci.* **128**, 1–16. (doi:10.1007/s12040-019-1103-z)
133. Abaurrea J, Cebrián AC. 2002 Drought analysis based on a cluster Poisson model: distribution of the most severe drought. *Clim. Res.* **22**, 227–235. (doi:10.3354/cr022227)

134. Ali Z, Hussain I, Faisal M, Shoukry AM, Gani S, Ahmad I. 2019 A framework to identify homogeneous drought characterization regions. *Theor. Appl. Climatol.* **137**, 3161–3172. (doi:10.1007/s00704-019-02797-w)
135. Gelbard R, Goldman O, Spiegler I. 2007 Investigating diversity of clustering methods: an empirical comparison. *Data Knowl. Eng.* **63**, 155–166. (doi:10.1016/j.datak.2007.01.002)
136. Ham Y-G, Kim J-H, Luo J-J. 2019 Deep learning for multi-year ENSO forecasts. *Nature* **573**, 568–572. (doi:10.1038/s41586-019-1559-7)
137. Tangang FT, Tang B, Monahan AH, Hsieh WW. 1998 Forecasting ENSO events: a neural network–extended EOF approach. *J. Clim.* **11**, 29–41. (doi:10.1175/1520-0442(1998)011<0029:FEEANN>2.0.CO;2)
138. Leloup JA, Lachkar Z, Boulanger J-P, Thiria S. 2007 Detecting decadal changes in ENSO using neural networks. *Clim. Dyn.* **28**, 147–162. (doi:10.1007/s00382-006-0173-1)
139. Tao Y, Hsu K, Ihler A, Gao X, Sorooshian S. 2018 A two stage deep neural network framework for precipitation estimation from bispectral satellite information. *J. Hydrometeorol.* **19**, 393–408. (doi:10.1175/JHM-D-17-0077.1)
140. Akbari Asanjan A, Yang T, Hsu K, Sorooshian S, Lin J, Peng Q. 2018 Short-term precipitation forecast based on the PERSIANN System and LSTM recurrent neural networks. *J. Geophys. Res. Atmos.* **123**, 12 543–12 563. (doi:10.1029/2018JD028375)
141. Sadeghi M, Nguyen P, Hsu K, Sorooshian S. 2020 Improving near real-time precipitation estimation using a U-net convolutional neural network and geographical information. *Environ. Model. Software* **134**, 204856. (doi:10.1016/j.envsoft.2020.104856)
142. Gorooh VA, Kalia S, Nguyen P, Hsu K, Sorooshian S, Ganguly S, Nemani RR. 2020 Deep neural network cloud-type classification (DeepCTC) model and its application in evaluation PERSIANN-CCS. *Remote Sensing* **12**, 316. (doi:10.3390/rs12020316)
143. Ganguli P, Janga Reddy M. 2014 Ensemble prediction of regional droughts using climate inputs and the SVM–copula approach. *Hydrol. Processes* **28**, 4989–5009. (doi:10.1002/hyp.9966)
144. Cherkassky V. 1997 The nature of statistical learning theory. *IEEE Trans. Neural Netw.* **8**, 1564–1564. (doi:10.1109/TNN.1997.641482)
145. Nourani V, Molajou A. 2017 Application of a hybrid association rules/decision tree model for drought monitoring. *Glob. Planet. Change* **159**, 37–45. (doi:10.1016/j.gloplacha.2017.10.008)
146. Rahnamay Naeini M, Yang T, Tavakoly A, Analui B, AghaKouchak A, Hsu K, Sorooshian S. 2020 A Model Tree Generator (MTG) framework for simulating hydrologic systems: application to reservoir routing. *Water* **12**, 2373. (doi:10.3390/w12092373)
147. Chen J, Li M, Wang W. 2012 Statistical uncertainty estimation using random forests and its application to drought forecast. *Math. Problems Eng.* **2012**, 915053. (doi:10.1155/2012/915053)
148. Breiman L. 2001 Random forests. *Mach. Learning* **45**, 5–32. (doi:10.1023/A:1010933404324)
149. Park H, Kim K, Lee Dk. 2019 Prediction of severe drought area based on random forest: using satellite image and topography data. *Water* **11**, 705. (doi:10.3390/w11040705)
150. Park S, Seo E, Kang D, Im J, Lee M-I. 2018 Prediction of drought on pentad scale using remote sensing data and MJO index through random forest over East Asia. *Remote Sensing* **10**, 1811. (doi:10.3390/rs10111811)
151. Tang C, Garreau D, von Luxburg U. 2018 When do random forests fail? In *32nd Conf. on Neural Information Processing Systems (NeurIPS 2018), Montreal, Canada*. (doi:10.1007/978-3-030-04167-0)
152. Zhang R, Chen ZY, Xu LJ, Ou CQ. 2019 Meteorological drought forecasting based on a statistical model with machine learning techniques in Shaanxi province, China. *Sci. Total Environ.* **665**, 338–346. (doi:10.1016/j.scitotenv.2019.01.431)
153. Vuyyuru VA, Appa Rao G, Srinivasa Murthy YV. 2021 A novel weather prediction model using a hybrid mechanism based on MLP and VAE with fire-fly optimization algorithm. *Evol. Intell.* **14**, 1173–1185. (doi:10.1007/s12065-021-00589-8)
154. Ravuri S, Lenc K, Willson M, Kangin D, Lam R, Mirowski P, Fitzsimons M *et al.* 2021 Skilful precipitation nowcasting using deep generative models of radar. *Nature* **597**, 672–677. (doi:10.1038/s41586-021-03854-z)
155. Xu J, Li H, Zhou S. 2015 An overview of deep generative models. *IETE Tech. Rev.* **32**, 131–139. (doi:10.1080/02564602.2014.987328)

156. Kingma DP, Welling M. 2019 An introduction to variational autoencoders. *Found. Trends® Mach. Learn.* **12**, 307–392. (doi:10.1561/2200000056)
157. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. 2014 Generative adversarial nets. *Communications of the ACM* **63**, 139–144. (doi:10.1145/3422622)
158. Liu S, Sun Y, Zhu D, Bao R, Wang W, Shu X, Yan S. 2017 Face aging with contextual generative adversarial nets. In *Proc. of the 25th ACM Int. Conf. on Multimedia*, *New Orleans, LA, 20–24 January 2008*, pp. 82–90. Boston, MA: American Meteorological Society.
159. Bishop CM. 1994 Neural networks and their applications. *Rev. Sci. Instrum.* **65**, 1803–1832. (doi:10.1063/1.1144830)
160. Dawson CW, Wilby RL. 2001 Hydrological modelling using artificial neural networks. *Progr. Phys. Geogr. Earth Environ.* **25**, 80–108. (doi:10.1177/030913330102500104)
161. Chong KFE. 2020 'A closer look at the approximation capabilities of neural networks.'
162. Montavon G, Samek W, Müller K-R. 2018 Methods for interpreting and understanding deep neural networks. *Digital Signal Processing* **73**, 1–15. (doi:10.1016/j.dsp.2017.10.011)
163. Kumar S, KP Sudheer AR, Jain SK, Agarwal PK. 2005 Rainfall-runoff modelling using artificial neural networks: comparison of network types. *Hydrol. Process. Int. J.* **19**, 1277–1291. (doi:10.1002/hyp.5581)
164. Hsu K, Gupta HV, Sorooshian S. 1995 Artificial neural network modeling of the rainfall-runoff process. *Water Resour. Res.* **31**, 2517–2530. (doi:10.1029/95WR01955)
165. Wu A, Hsieh WW. 2004 The nonlinear association between ENSO and the Euro-Atlantic Winter sea level pressure. *Clim. Dyn.* **23**, 859–868. (doi:10.1007/s00382-004-0470-5)
166. Poornima S, Pushpalatha M. 2019 Drought prediction based on SPI and SPEI with varying timescales using LSTM recurrent neural network. *Soft Computing* **23**, 8399–8412. (doi:10.1007/s00500-019-04120-1)
167. Saarinen S, Bramley R, Cybenko G. 1993 Ill-conditioning in neural network training problems. *SIAM J. Sci. Comput.* **14**, 693–714. (doi:10.1137/0914044)
168. Khan MMH, Muhammad NS, El-Shafie A. 2020 Wavelet based hybrid ANN-ARIMA models for meteorological drought forecasting. *J. Hydrol.* **590**, 125380. (doi:10.1016/j.jhydrol.2020.125380)
169. Mishra AK, Desai VR, Singh VP. 2007 Drought forecasting using a hybrid stochastic and neural network model. *J. Hydrol. Eng.* **12**, 626–638. (doi:10.1061/(ASCE)1084-0699(2007)12:6(626))
170. Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, Xiong H, He Q. 2020 A comprehensive survey on transfer learning. *Proc. IEEE* **109**, 43–76. (doi:10.1109/JPROC.2020.3004555)
171. Gibson PB, Chapman WE, Altinok A, Delle Monache L, DeFlorio MJ, Waliser DE. 2021 Training machine learning models on climate model output yields skillful interpretable seasonal precipitation forecasts. *Commun. Earth Environ.* **2**, 1–13. (doi:10.1038/s43247-021-00225-4)
172. Domingos P. 2012 A few useful things to know about machine learning. *Commun. ACM* **55**, 78–87. (doi:10.1145/2347736.2347755)
173. Beven K, Binley A. 1992 The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Processes* **6**, 279–298. (doi:10.1002/hyp.3360060305)
174. Gupta H, Beven K, Wagener T. 2005 *Encyclopedia of hydrological sciences, model calibration and uncertainty estimation*. New York, NY: John Wiley & Sons.
175. Kuczera G, Parent E. 1998 Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm. *J. Hydrol.* **211**, 69–85. (doi:10.1016/S0022-1694(98)00198-X)
176. Malinin A, Gales M. 2018 Predictive uncertainty estimation via prior networks. In *Advances in neural information processing systems* (eds S Bengio, H Wallach, H Larochelle, K Grauman, N Cesa-Bianchi, R Garnett). Red Hook, NY: Curran Associates, Inc.
177. Montanari A, Brath A. 2004 A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. *Water Resour. Res.* **40**, W01106. (doi:10.1029/2003WR002540)
178. Apel H, Thieken AH, Merz B, Blöschl G. 2004 Flood risk assessment and associated uncertainty. *Nat. Hazards Earth Syst. Sci.* **4**, 295–308. (doi:10.5194/nhess-4-295-2004)
179. Helton JC, Johnson JD, Oberkampf WL, Sallaberry CJ. 2010 Representation of analysis results involving aleatory and epistemic uncertainty. *Int. J. Gen. Syst.* **39**, 605–646. (doi:10.1080/03081079.2010.486664)

180. Merz B, Thieken AH. 2005 Separating natural and epistemic uncertainty in flood frequency analysis. *J. Hydrol.* **309**, 114–132. (doi:10.1016/j.jhydrol.2004.11.015)
181. Hasselmann K. 1976 Stochastic climate models part I. Theory. *Tellus* **28**, 473–485. (doi:10.3402/tellusa.v28i6.11316)
182. Coelho, Pezzulli S, Balmaseda M, Doblas-Reyes FJ, Stephenson DB. 2004 Forecast calibration and combination: a simple Bayesian approach for ENSO. *J. Clim.* **17**, 1504–1516.
183. Raftery AE, Gneiting T, Balabdaoui F, Polakowski M. 2005 Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Weather Rev.* **133**, 1155–1174. (doi:10.1175/MWR2906.1)
184. Luo L, Wood EF, Pan M. 2007 Bayesian merging of multiple climate model forecasts for seasonal hydrological predictions. *J. Geophys. Res. Atmosph.* **112**, D10. (doi:10.1029/2006JD007655)
185. Schepen A, Wang QJ. 2013 Toward accurate and reliable forecasts of Australian seasonal rainfall by calibrating and merging multiple coupled GCMs. *Mon. Weather Rev.* **141**, 4554–4563. (doi:10.1175/MWR-D-12-00253.1)
186. Buser CM, Künsch HR, Schär C. 2010 Bayesian multi-model projections of climate: generalization and application to ENSEMBLES results. *Clim. Res.* **44**, 227–241. (doi:10.3354/cr00895)
187. Najafi MR, Moradkhani H. 2015 Multi-model ensemble analysis of runoff extremes for climate change impact assessments. *J. Hydrol.* **525**, 352–361. (doi:10.1016/j.jhydrol.2015.03.045)
188. Madadgar S, AghaKouchak A, Shukla S, Wood AW, Cheng L, Hsu KL, Svoboda M. 2016 A hybrid statistical-dynamical framework for meteorological drought prediction: application to the southwestern United States. *Water Resour. Res.* **52**, 5095–5110. (doi:10.1002/2015WR018547)
189. Schepen A, Wang QJ, Robertson DE. 2014 Seasonal forecasts of Australian rainfall through calibration and bridging of coupled GCM outputs. *Mon. Weather Rev.* **142**, 1758–1770. (doi:10.1175/MWR-D-13-00248.1)
190. Slater LJ, Villarini G. 2018 Enhancing the predictability of seasonal streamflow with a statistical-dynamical approach. *Geophys. Res. Lett.* **45**, 6504–6513. (doi:10.1029/2018GL077945)
191. Cheng L, AghaKouchak A. 2015 A methodology for deriving ensemble response from multimodel simulations. *J. Hydrol.* **522**, 49–57. (doi:10.1016/j.jhydrol.2014.12.025)
192. Taleb NN. 2007 *The black swan: the impact of the highly improbable*, vol. 2. New York, NY: Random house.
193. Runge J, Nowack P, Kretschmer M, Flaxman S, Sejdinovic D. 2019 Detecting and quantifying causal associations in large nonlinear time series datasets. *Sci. Adv.* **5**, eaau4996. (doi:10.1126/sciadv.aau4996)
194. Zappa G, Shepherd TG. 2017 Storylines of atmospheric circulation change for European regional climate impact assessment. *J. Clim.* **30**, 6561–6577. (doi:10.1175/JCLI-D-16-0807.1)
195. Chan WC, Shepherd TG, Facer-Childs K, Darch G, Arnell NW. 2022 Storylines of UK drought based on the 2010–2012 event. *Hydrol. Earth Syst. Sci.* **26**, 1755–1777. (doi:10.5194/hess-26-1755-2022)
196. Reynolds, Rayner RA, Smith TM, Stokes DC, Wang W. 2002 An improved in situ and satellite SST analysis for climate. *J. Clim.* **15**, 1609–1625.
197. Ashouri, Hsu K-L, SorooshianS, Braithwaite DK, Knapp KR, Cecil LD, Nelson BR, Prat OP. 2015 PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bull American Meteorological Society* **96**, 69–83. (doi:10.1175/BAMS-D-13-00068.1)
198. KorfRE. 1985 Depth-first iterative-deepening: an optimal admissible tree search. *Artif. Intell.* **27**, 97–109. (doi:10.1016/0004-3702(85)90084-0)
199. DelSoleT. 2004 Predictability and information theory. Part I: measures of predictability. *J. Atmos. Sci.* **61**, 2425–2440.
200. DelSoleT, Shukla. 2010 Model fidelity versus skill in seasonal forecasting. *J. Clim.* **23**, 4794–4806. (doi:10.1175/2010JCLI3164.1)