Climatic Change (2021) 164:33
https://doi.org/10.1007/s10584-021-02985-6

REVIEW ARTICLE

Techniques for constructing climate scenarios for stress test applications

Christine M. Albano[1] · Maureen I. McCarthy[1,2] · Michael D. Dettinger[3] · Stephanie A. McAfee[4]

Received: 9 January 2020 / Accepted: 6 January 2021/Published online: 08 February 2021

Abstract

In this review, we provide guidance on the construction of climate scenarios for stress tests—scenarios that represent disruptive climatic events and can be used to assess the impacts of climate and weather risks at the level of detail that is necessary to identify specific adaptation actions or strategies. While there is a wealth of guidance on scenario-based climate adaptation planning, this guidance typically assumes the selection and use of decadal to century-long time segments of downscaled climate model projections, rather than the creation of a customized scenario depicting a specific extreme event. We address this gap by synthesizing a variety of data sources and analytical techniques for constructing climate scenarios for stress tests that are customized to address specific end-users' needs. We then illustrate the development and application of climate scenarios with a case study that explores water sustainability under changing climate in the Truckee and Carson River basins of California and Nevada. Finally, we assess the potential advantages and disadvantages of the different data sources and analytical techniques described to provide guidance on which are best suited for an intended application based on the system of study, the stakeholders involved, and the resources available. Ultimately, this work is intended to provide the building blocks with which scientist-stakeholder teams can produce their own stress test scenarios to explore place-based weather and climate risks.

Keywords Climate extremes · Stress test · Adaptation planning · Scenario planning · Resilience

✉ Christine M. Albano
christine.albano@dri.edu

1 Desert Research Institute, Reno, NV, USA

2 Department of Physics, University of Nevada, Reno, Reno, NV, USA

3 California-Nevada Climate Applications Program, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA

4 Department of Geography, Extension, and Nevada State Climate Office, University of Nevada, Reno, Reno, NV, USA

1 Introduction

Across the globe, extreme weather and climate-related natural hazards such as floods, droughts, and heat waves have cost thousands of lives and hundreds of billions of dollars and threatened human and ecosystem health. These impacts are expected to increase as global temperatures rise, climates change, and the human footprint continues to expand (Intergovernmental Panel on Climate Change (IPCC) 2012). For this reason, there is an increasing recognition that assessments of climate change impacts should include evaluations of risks from extreme weather and climate (Weaver et al. 2017).

Climate scenarios are commonly used to assess climate change impacts in efforts to identify adaptation strategies that are likely to be robust to a range of climate futures (Rowland et al. 2014), and their use in scenario planning comes in a wide variety of applications and forms (Star et al. 2016). One application of climate scenarios involves the use of stress tests (Stern et al. 2013), also referred to as "wind tunneling" (Van Der Merwe 2008) or vulnerability analyses (Brown and Wilby 2012), to evaluate the risks and impacts of disruptive climatic conditions. In contrast to climate impact assessments that assess a diverse range of potential futures that are expected to be broadly relevant for a particular place or set of resources (e.g., scenario-light or scenario-heavy approaches; Runyon et al. 2020), stress testing involves a more narrow and targeted focus on a discrete event or set of extreme conditions under which a system of interest breaks or fails (National Research Council 2013). This focus on specific climate risks allows the impacts to be explored with enough detail to engage communities in identifying strategies that explicitly address these risks (Weaver et al. 2017) including cascading impacts that can otherwise be difficult to recognize (Albano et al. 2016).

As is the case for most scenario planning applications, scenarios for stress tests may be narrative or quantitative in nature, depending on the resources available to conduct the assessment and the complexity of the system being studied. Narrative stress test scenarios, consisting of a qualitative description of how an extreme event unfolds, are commonly used in emergency response planning exercises (Department of Homeland Security 2013). In general, narrative scenarios provide a valuable basis for diverse stakeholder groups to assess risks at minimal cost. Quantitative stress test scenarios are based on data and model-based simulations and thus require more computational resources to develop. Quantitative stress tests are particularly beneficial for assessing climate impacts to highly managed, complex, or nonlinear systems where responses are difficult to anticipate and when there is a need for spatially explicit response information. In this paper, we focus on methods for developing quantitative climate stress tests.

Quantitative climate stress tests have been used in many planning and research contexts (Groves and Lempert 2007; Mahmoud et al. 2011; Dettinger et al. 2012; Miller et al. 2017; Symstad et al. 2017; Tariq et al. 2017; Ullrich et al. 2018) using a variety of methods and data sources, but generalizable advice on how to develop them is not readily available. Existing guidance documents that describe how to use quantitative scenarios for climate change impact studies (Mckenzie et al. 2012; Wilsey et al. 2013; Rowland et al. 2014) contain information that is highly relevant to the application of climate stress tests; however, these are generally aimed toward the selection and use of already downscaled and bias-corrected global climate model projections that serve as direct inputs to impact models. And while there is substantial advice available for selecting and applying climate model projections for impact assessments that is also useful and applicable to stress tests (e.g., Snover et al. 2013; Vano et al. 2015; Sofaer et al. 2017), there is a lack of explicit guidance for constructing a customized scenario that represents a discrete climatic event or set of extreme conditions for use in stress test applications.

In this study, we seek to address this gap by providing guidance on the types of data and methods that can be used to construct climate stress test scenarios that are customized to the needs of end-users. Our approach is to (1) present a general framework for generating and applying quantitative climate stress test scenarios in the context of stakeholder-driven climate adaptation planning and decision-making; (2) describe a range of climate data sources and analytical techniques that can be used to generate stress tests and present advantages and disadvantages for each of these; (3) provide examples of how these methods have been implemented as part of a water resource sustainability planning effort in the Truckee and Carson River basins of Nevada and California; and (4) provide guidance for choosing which data and methods to use to support stakeholder planning and actions. Our overall objective is to increase the accessibility and use of quantitative climate stress tests by providing guidance on their construction and implementation.

2 A framework for climate adaptation planning based on stress test scenarios

The quantitative climate stress test approaches described here are intended to be used in a bottom-up (Dessai and Hulme 2004) iterative process that is driven by stakeholders' questions, concerns, and needs. The process begins with identifying the climate vulnerabilities stakeholders are most concerned about. From there, a small number of climate stress test scenarios are developed that represent these vulnerabilities. Ideally, multiple scenarios that span a range of uncertainties regarding future climate risks are explored to ensure that management strategies identified are robust across multiple conditions (Lempert et al. 2003), but for applications that seek to focus on particularly complex scenarios or responses, such as emergency response exercises, it may be appropriate to focus on a single scenario. Regardless, iteration with stakeholders throughout this process is critical to ensure that outputs are both realistic and informative (Van Der Merwe 2008) and that stakeholders are invested in, understand, trust, and have a sense of ownership in the results (Dilling and Lemos 2011; Lemos et al. 2012).

An example of such a process is detailed in Fig. 1 and includes (1) an initial problem framing and scenario narrative development step, (2) development and simulation of the stress test(s), and (3) identification of adaptation strategies, with each of these steps involving iterations with stakeholders. The framework recognizes two paths by which climate stress tests can be developed—the storyline approach and the scenario discovery approach. In both cases, preliminary discussions with stakeholders define the key climatic (and other) drivers of interest, the decision-making context, and metrics of performance for the system being studied (e.g., Groves and Lempert 2007; Mahmoud et al. 2009; Rowland et al. 2014). The key difference between these two approaches is that in the storyline's case, stress tests are defined based on stakeholders' perceptions of the climatic events they deem to be most impactful while in the discovery case, stress tests are determined analytically, based on simulated impacts (Groves and Lempert 2007).

The storyline (Garb et al. 2008; Hazeleger et al. 2015; Shepherd et al. 2018) approach starts with identifying one or more climate-event "storylines" that reflect the climate vulnerabilities of the system(s) or resource(s) of interest and constructing a climatic time series that captures

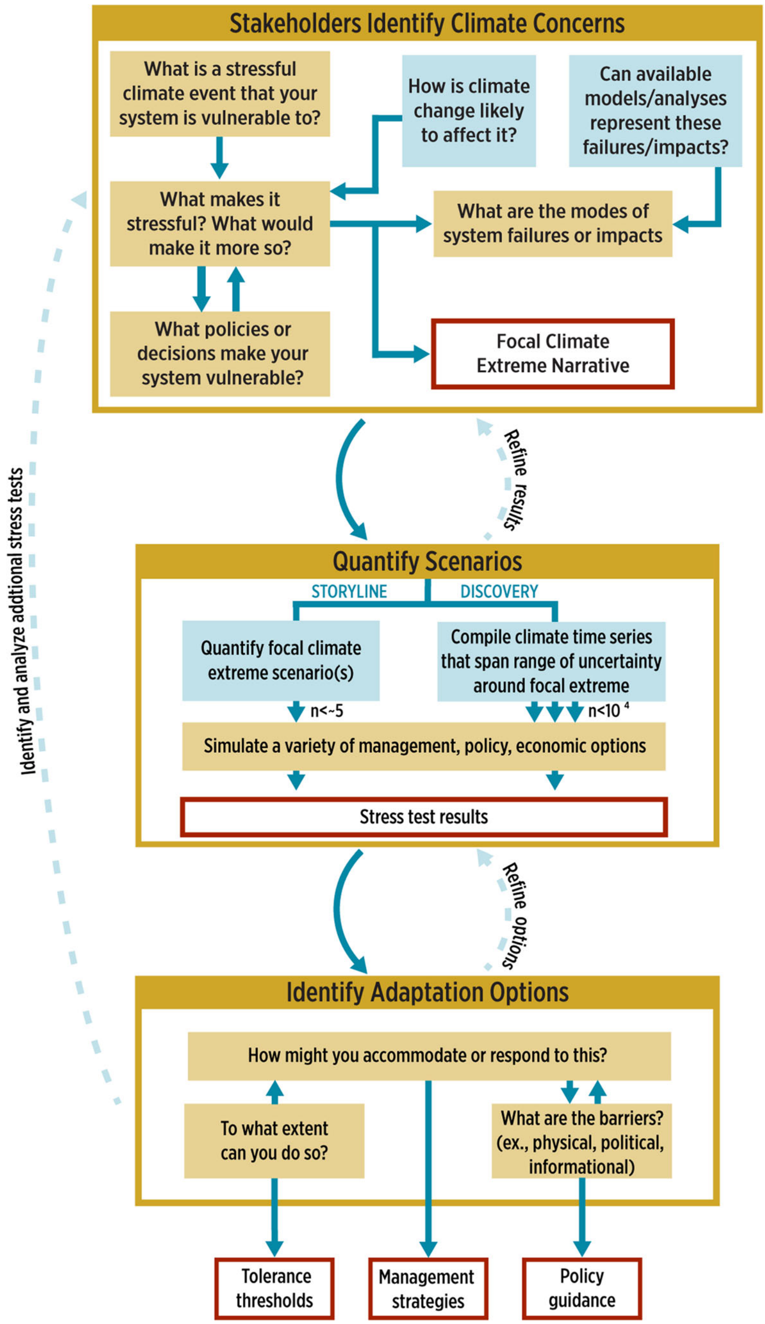

Fig. 1 Iterative process for translating stakeholder concerns about extreme climate and weather events into quantitative stress tests that can be used to explore adaptation strategies. Blue and gold boxes indicate points in the process where the dominant sources of information come from scientists and stakeholders, respectively, though both parties are ideally providing input in all cases. Red outlined boxes indicate outputs from each step

one or more events described in the storyline. This could start with a broad characterization of the types of events stakeholders are concerned about (e.g., extreme winter floods, extended drought), followed by the identification of more specific attributes of the event or change that make it of higher risk or concern. Once a small number of climate storylines are identified, they are translated into quantitative stress tests using techniques described in the next section of this paper. The stress tests are then input into one or more simulation models, and system performance is evaluated across a range of management or policy options based on pre-defined performance indicators (Fig. 1).

In the case of scenario discovery (Groves and Lempert 2007; Lempert 2013), a large ensemble of climate time series is run through one or more impact simulation models to identify the climatic conditions or decision alternatives under which the system does or does not meet objectives, based on a priori–defined performance metrics. Climate sequences where the system becomes stressed can then be identified and examined in greater detail. Regardless of the approach (storylines vs. discovery), the output from the stress test construction and simulation step is a small and digestible set of stress tests that represent a diversity of climate stresses and management options that can serve as a focal point for discussion, iteration, and shared learning with stakeholders.

The third step involves identification of management strategies that can accommodate the impacts of the climate stress (i.e., adaptation strategies). Although it is identified as a “third” step here, identification and simulation of adaptation strategies may come earlier in the process, depending on stakeholders’ initial understanding of risks and initial desire to explore various alternative strategies. Regardless, this step serves as the point at which stress tests are explored in depth with stakeholders to identify viable management strategies or policy guidance that mitigates risks, while also providing insights into potential system behaviors and tolerance thresholds. This point in the process also serves as a driving force behind iteration as uncertainties in climate stressors and/or management strategies are further explored, simulations are refined to achieve greater realism or incorporate new ideas and information, and more novel or stressful scenarios are explored as scientist-stakeholder teams develop greater rapport (Fig. 1).

3 Data sources and analytical techniques for specifying climate stress tests

3.1 Data sources

Data sources for stress test applications include historical climate records, stochastically generated weather time series, paleoclimate proxies, and climate model projections (Fig. 2). Here, we describe how each of these could be used as a basis for stress tests and summarize their advantages and disadvantages (Table 1).

3.1.1 Historical data

Historical climate records from weather stations or gridded climate products are one of the most readily available sources of climatic time series that can be used to construct a climate scenario, often with daily or potentially higher temporal resolution that can capture some types of extreme weather quite well. A disadvantage is that observational records cover a limited

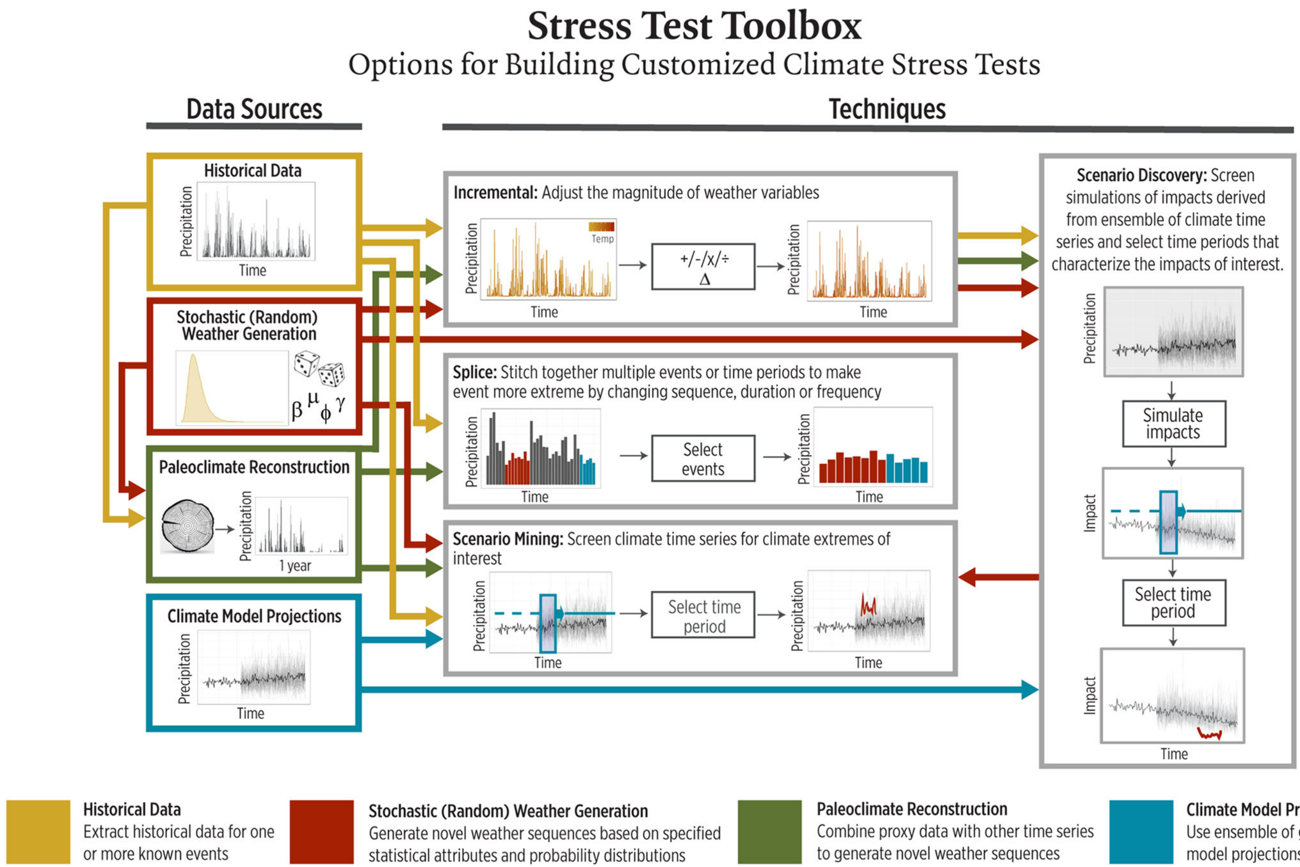

Fig. 2 Overview of four potential data sources and four potential techniques for constructing climate scenarios for stress tests. Arrows indicate different ways techniques and data sources could be combined

Table 1 Summary of advantages and disadvantages of data sources and climate stress test scenario construction techniques discussed in this paper

	Advantages	Disadvantages
	Data types	
Historical	• Familiar events provide good basis for discussion • Data are readily accessible and available • Sub-daily resolution data and observations may be available	• Climate sequences limited to timeframe of historical record • Use of familiar historical event risks 'de-facto' thinking by stakeholders if not sufficiently perturbed
Stochastic weather generator	• Ability to generate infinite number of climate sequences for subsequent sensitivity analysis • Novel sequences offer opportunities for surprise	• Lack process representation and may not reasonably reflect climate dynamics • Adequate simulation of spatial and temporal autocorrelation over long time periods or large spatial extents is difficult • Do not capture variability and extremes well
Paleoclimate reconstructions	• Represent real-world events outside the experiences of many stakeholders • May include wider range of variability of climatic conditions than climate model projections or historical data • Offers opportunities for surprise	• Additional processing needed to generate sub-annual resolution climate sequences • Lack of good spatial representation • Potential to draw incorrect inferences from proxy data
Climate model projections	• Data are readily accessible and available • Best known representations of potential climate futures and processes driving them • Offer opportunities for surprise	• Available projections limit the range of uncertainty that can be explored • Potential for bias/poor representation of extremes due to model structural errors and bias correction
	Method types	
Incremental	• Straightforward to implement • Useful for sensitivity and uncertainty analyses • Useful for pseudo-global warming experiments (see Section 3.2.1)	• Lack of climate process representation and may not reasonably reflect dynamics or thermodynamics of the event
Splice	• Simple to implement • Familiar events provide good basis for discussion • Avoids adjustments of magnitudes, changing only duration or sequencing of event	• Unless combined with another method, one cannot draw inferences about causal factors, system sensitivities, or uncertainties
Scenario mining	• Results in large database of information for additional analysis or iteration with stakeholders • Ability to generate conditional probabilities of impact occurrence	• Requires initial investment in generating stress test relevant metrics • Extreme meteorology does not necessarily translate to extreme impacts • Ability to generate conditional probabilities gives false sense of uncertainty range
Scenario discovery	• Objective identification of scenarios based on their simulated impacts reduces perception of bias and more precisely targets impacts of interest • Results in large database of information for additional analysis or iteration with stakeholders • Enables comprehensive exploration of vulnerability space and analysis of system sensitivities, conditional probabilities, and uncertainties	• Computationally demanding. Requires high-quality data and models and substantial investment in simulation and data analysis. • Ability to generate conditional probabilities gives false sense of uncertainty range

time span and thus may not capture the most extreme events, which tend to occur less frequently, nor do they reflect novel climate time sequences that might arise due to non-stationarity of climate. One of the greatest strengths of using historical data is that stress tests can be loosely based on events that were recently experienced by, or are familiar to, stakeholders. Scenarios based on familiar events are much more likely to resonate with stakeholders (Vasileiadou and Botzen 2014) and are more likely to be viewed as plausible (Alexander 2000). They also provide a tangible starting point for discussions of climate-related risks as stakeholders can use their past experiences (i.e., historical analogs) as a reference point (Ford et al. 2010; Albano et al. 2016). The main challenge with this approach is ensuring that the stress test is sufficiently different from the historical event upon which it is based that it represents a novel experience for stakeholders so as to avoid “de facto solutions” (Alexander 2000).

3.1.2 Stochastic weather generation

Time series developed using stochastic weather generators (SWGs) are a second data resource. These generators draw random values from probability distributions fitted to a reference climate record (past or future) to simulate novel weather sequences that preserve statistical attributes such as means, variances, and frequencies, as well as their spatial and temporal autocorrelation characteristics (Wilks and Wilby 1999). Stochastic weather (and streamflow) generators have been used in a variety of contexts to assess climate impacts (e.g., Wilks and Wilby 1999; Fatichi et al. 2011; Kirsch et al. 2013). The strength of using SWGs is that an infinite number of time series can be generated that allow for statistically robust sensitivity analyses to be conducted (Brown and Wilby 2012). The weaknesses of this approach are that many SWG models do not capture variability, extremes, or spatial patterns well and those that do are complicated in structure (i.e., require the use of multiple parameters or data distributions; Chen and Brissette 2014) and therefore require significant expertise to implement. Moreover, in the absence of specialized interventions, this approach can result in unrealistic sequences that lack the multi-scaled spatiotemporal autocorrelations and long-term structure of observed climatic time series, especially at multi-year to decadal timescales and longer (Chen et al. 2010), so their use is best limited to short-duration (< 1 year) events over localized areas.

3.1.3 Paleoclimate reconstructions

Paleoclimate reconstructions based on proxy data provide a third alternative data source for stress test development. A variety of proxy data, such as ice or sediment cores, pollen, or biological organisms (e.g., tree rings, corals, plankton, diatoms) are available. These can be useful for characterizing a variety of past climatic conditions and are especially helpful in quantifying their variability and persistence, given the long-term records they provide (Sorooshian and Martinson 1995). Tree-ring reconstructions are one of the more convenient data sources because they are widely available (e.g., Cook and Krusic 2004), can represent conditions that occurred in relatively recent times (i.e., the past few thousand years), and are generally available at annual resolution. Tree-ring data can be used to reconstruct temperature and precipitation conditions, as well as some of the ecological or hydrological effects of droughts or pluvial events. For example, tree rings have been used to reconstruct fire histories and forest dynamics (Amoroso et al. 2017), flood stages (Ballesteros-Canovas et al. 2015), and streamflows (e.g., Meko and Graybill 1995; Woodhouse and Lukas 2006), any of which could be useful in the context of stress tests.

For some extreme event applications, sub-annual time series may be desirable. One relatively straightforward approach to achieving this using paleoclimate reconstructions is to apply daily time series from the historical climate record to analog years in the proxy record (Dettinger et al. 2017). This "temporal analog" approach is best applied when conditions in a given year in the paleoclimatic record are within the bounds of historical data records. In cases where reconstructed conditions are outside the bounds of the historical record, incremental adjustments (see Section 3.2.1 on incremental techniques) could be made to closely analogous years to match the proxy record (Dettinger et al. 2017). More complicated approaches involve the stochastic generation (see previous section) of sub-annual weather or streamflow sequences (e.g., Gangopadhyay et al. 2009, 2015; Sauchyn and Ilich 2017).

Overall, the strength of paleoclimate proxy-based tests is that they can represent real-world events that are outside the bounds of what many stakeholders have experienced. They may also capture a wider range of variability than is contained in future climate projections, which are often limited to 100 years. Using paleoclimate proxies of climate extremes may also be more accepted by some stakeholders who are skeptical of climate models or by indigenous groups with traditional knowledge of the impact of climatic events on their communities (Norton-Smith et al. 2016). Conversely, this data source can be more challenging to work with relative to others in cases when proxy information must be integrated with historical or stochastic time series to generate sufficiently high temporal resolution but simple techniques, such as those described in Dettinger et al. (2017), are available.

3.1.4 Climate model projections

Climate model projections, including those from global climate models (GCMs; e.g., from the Fifth Coupled Model Intercomparison Project (CMIP5); Taylor et al. 2012) or regional climate models (RCMs), are a fourth resource for developing climate stress test scenarios. The advantages of using climate models are that their outputs have become readily available in recent years, they provide the best-known representations of potential futures, and they can simulate connections and changes that may surprise researchers and stakeholders in useful ways. Disadvantages of their use are that they can underestimate variability and extremes (Mote et al. 2011), they may share structural errors that result in systematic biases (e.g., Seager et al. 2019), and typically, only a limited number of projections (compared to the full range of possible climate changes) is available, and thus, they provide a narrow and biased view of the uncertainty space (Brown and Wilby 2012), which may be seen as a limitation for applications where exploring a wide range of uncertainties is desired (see Section 3.2.4). In addition, climate model outputs often need to be bias-corrected and downscaled before they can be used for adaptation planning, which has the potential to create unanticipated but influential artifacts (Maraun et al. 2017) that are not directly observable by end-users (Sofaer et al. 2017), and it requires expertise to understand how post-processing steps may have affected the extreme of interest.

To represent an extreme event at a local scale, climate model outputs are typically downscaled to a resolution appropriate to the extreme and the impacts being analyzed while also correcting for bias at that resolution. For many hydrologic applications, daily resolution data are needed, and several statistical downscaling approaches have been developed to achieve this. Statistical downscaling methods can differ substantially in their abilities to reproduce different types of climate and hydrologic extremes (Werner and Cannon 2016), and some (e.g., Pierce et al. 2014) resolve these with finer spatial and temporal detail than others and can be useful for representing daily-resolution extremes. If finer than daily resolution is required, as might be the case for some applications of short-duration events such as storms or fires, an RCM such as the Weather Research and Forecasting (WRF; Skamarock et al. 2008) model may be used to dynamically downscale an event to get the necessary spatial and temporal detail, using a bias-corrected GCM to supply boundary conditions.

3.2 Stress test specification techniques

Here, we describe four analytical techniques that use the data sources described above to construct quantitative stress tests. Although we outline methods individually, here many of these methods are used in combination to achieve study objectives (Fig. 2). We further discuss their strengths and weaknesses (Table 1).

3.2.1 Incremental method

The incremental method (Mearns et al. 2001), which is inclusive of techniques known as change factor, delta change factor (Anandhi et al. 2011), or delta (Sofaer et al. 2017) methods, involves the adjustment of climate variables such as temperature or precipitation magnitudes by a specified amount. This approach can be applied to historical, paleoclimate, or stochastically generated sequences (Fig. 2). It is a straightforward technique to apply, is the simplest to explain to diverse stakeholder groups, and is useful when a goal of the study is to understand behaviors and sensitivities of a system driven by mean changes in one or more variables at a time (Sofaer et al. 2017). There are numerous ways in which adjustments can be made (see Mearns et al. 2001, Anandhi et al. 2011, and Sofaer et al. 2017 for more comprehensive reviews), ranging from simple applications of spatially and temporally uniform changes to space- and time-varying adjustments (e.g., to modify variance or to make season-specific adjustments). Adjustments could also include random components to add elements of realistic uncertainty and fluctuation to the spatial or temporal changes or could be informed by climate model projections (Bureau of Reclamation 2015). For stress test applications, temperature or precipitation adjustments could be applied to an already-extreme event in the climate record to alter its intensity. For example, a drought period could be made warmer or drier, and an extreme storm could be adjusted to be wetter or to change the relative amounts of rain vs. snow. If system sensitivities and response thresholds are of interest, incremental adjustments could be systematically applied to a climate sequence, simulated to generate response surfaces, and combined with climate projection information, following methods described in the decision-scaling literature (Brown et al. 2012; Whateley et al. 2015).

Incremental adjustments to one or more climate variables, whether uniform or selective, do not realistically capture the physics behind spatial and temporal variations of change. In cases where such internal consistency is important (e.g., when synergistic interplays between several different climatic variables have the potential to be as important as any single variable), some applications have instead used incremental methods to adjust the boundary conditions of numerical weather prediction models (Dominguez et al. 2018; Mahoney et al. 2018; Ullrich et al. 2018). In doing so, this “pseudo global warming” method (Kawase et al. 2009) offers a potentially more internally consistent depiction of how the event would unfold relative to a simple adjustment of the climate record, albeit at significantly greater computational cost.

3.2.2 Splice method

The splice method (Fig. 2) involves stitching together selected episodes from within a climate record to alter the sequencing, duration, or frequency of extreme events. Once the desired type and key attributes of an extreme event are initially identified, the climate record is searched for time periods that could potentially contribute to the storyline. These time periods are then spliced at points where the transition between them is as plausible as possible from a meteorological standpoint. For example, Dettinger et al. (2012) developed a 23-day extreme winter storm scenario by selecting and splicing together two large historical storm sequences, with the concatenation point chosen specifically to be the time when geopotential height fields over the Pacific and North America were best matched to ensure reasonable continuity of precipitation patterns over the two events. For other types of extremes, like droughts, attention to larger scale atmospheric patterns becomes less important and continuity can be accomplished by splicing during dry periods with similar temperatures, for example. Although this method may require some expertise to create internally consistent event sequences, this approach is relatively simple to implement, as a scenario can be developed by making straightforward and very plausible adjustments such as increasing the duration of an event or changing its sequencing, without adjusting observed magnitudes of meteorological variables.

3.2.3 Scenario mining method

The scenario mining method applies moving-window calculations of climate metrics that characterize extreme events of interest (e.g., a 10-year drought, a month-long storm, a week-long heatwave) across the duration of one or more climate time series to identify a time segment that best represents the climate-extreme condition(s) of interest (Fig. 2). The selection of the time period could involve plotting the resultant moving window calculations in one or more dimensions of the extreme of interest, akin to approaches to scenario selection that involve mapping potential scenarios along dimensions of high-priority external drivers (Rowland et al. 2014). This time period is then extracted and becomes the focus of detailed assessment. A variety of metrics could be evaluated in this application (see Karl et al. 1999 for examples). For example, extremes could be defined relative to the historical record (i.e., as occasions when a historical-period percentile or recurrence interval is exceeded) or based on specific thresholds.

Extremes in meteorological variability, such as alternating wet and dry periods (i.e., “weather whiplash”; Swain et al. 2018), may also be of interest. These can be identified in several ways, depending on how one chooses to measure variability. If the goal is to maximize or minimize variability for a specified time period, this could be characterized by standard deviations or coefficients of variation. Alternatively, if the goal is to enhance variation around a specific frequency band, variability could be estimated and ranked by first applying low- or high-pass filters (Chatfield 2004) to the raw projections and then calculating variances or coefficients of variation of the filtered series (compared to raw variances or coefficients of variation). Spectral methods, which are used to transform time series into component frequencies (Ghil et al. 2002), could be used to identify time periods with enhanced variability at the desired frequency. In other cases, a specific sequencing of variability might be desired; for example, a series of alternating wet and dry years or a series of multiple wet years followed by a series of multiple dry years. In these cases, the specified sequence could be represented as a schematic time series of highs and lows (e.g., 1 s and 0 s). Correlations between this sequence and a running time series of the same length in the climate projections can then be calculated, with the time periods with the highest correlations indicating those that most closely correspond to the desired sequencing of variability.

The scenario mining method is more computationally demanding than those previously described due to the initial investment required to screen the climate model projections. However, once this step is completed, there is a trove of information that can be tapped for subsequent analysis and iteration with stakeholders. We include a detailed workflow and example of how this approach was used in the case study described later in this paper. Another disadvantage of this method is that meteorological extremes do not necessarily translate to extreme impacts (Van Der Wiel et al. 2020); this can be addressed by the scenario discovery method described in the next section.

3.2.4 Scenario discovery method

The scenario discovery approach (Bryant and Lempert 2010), coined by the RAND Corporation as part of a formalized robust decision-making framework (RDM; Groves and Lempert 2007), can also be used to identify climate stress test scenarios. For example, in RDM, a large (i.e., hundreds to millions) ensemble of climate sequences and management alternative combinations spanning a range of uncertainties or decisions that a stakeholder group wishes to explore are fed through one or more impact simulation models to characterize the range of possible outcomes, identify the circumstances under which the system succeeds or fails, and identify management strategies that are robust across that range of outcomes (Groves and Lempert 2007). Data mining algorithms are then used to identify representative scenarios to be examined in greater depth that illustrate system stress or failures across a range of conditions (Bryant and Lempert 2010). This approach was developed to address the problem of how to objectively identify a small number of scenarios that are both relevant to decision-making and span a large range of uncertainty, given that the storyline approach (see Section 2) may be too subjective for some audiences (Groves and Lempert 2007).

In scenario discovery, input climate sequences could take the form of an ensemble of climate projections, incrementally adjusted historical or paleoclimate time series, or stochastically generated time series (Fig. 2). Climate models provide the best physical process representation of these time series inputs; however, one downside to this approach is that there are a limited number of climate realizations to draw from, which limits the range of uncertainty in climate parameter space that can be considered (Brown and Wilby 2012). Thus, studies have typically either (1) systematically applied incremental changes of temperature and precipitation across a range of expected uncertainty to the historical record and used this to analyze sensitivity (Schwarz et al. 2018) or (2) used stochastic weather generators to generate a very large number of climate sequences (Brown et al. 2012). While the advantages of these latter approaches are that they permit a more comprehensive exploration of the uncertainty space, the scenarios selected for in-depth exploration with stakeholders may not be climatically plausible, given that incremental and stochastic weather generated time series are unlikely to be inconsistent with larger scale dynamics that drive precipitation variability (Breinl et al. 2017). If physical process representation is important, scenario mining techniques (Fig. 2) could be applied to find analogs in historical records or in climate model projections that are known to simulate the climate and weather characteristics of the focal extreme event realistically.

The ability to objectively identify stress test scenarios based on a comprehensive analysis of simulated impacts is a key advantage of the scenario discovery method, as it ensures that the scenario selected stresses the system at the level that is intended. The very large simulated datasets produced by this method have considerable scientific value as causal factors leading to system failure can be analyzed and system sensitivities can be identified with high statistical power. The primary disadvantages of this approach are its substantial data and computational requirements. In addition, the process of developing the scenario may be more complicated to explain or may appear more "black box" to stakeholders than for a scenario that originates from a storyline approach (see Section 2).

4 Water for the Seasons case study

In the previous sections, we described an iterative stakeholder process and various approaches for generating climate stress test scenarios. Here, we describe a study in which these have been applied in practice. The Water for the Seasons (WftS) project (http://waterfortheseasons.com) engaged federal water agencies, state water engineers, local water utilities and districts, farmers, ranchers, native American communities, and ecosystem managers throughout the Truckee and Carson River basins in California and Nevada in the development of stakeholder-informed climate scenarios. These scenarios were used in combination with hydroclimatic models and social science collaborative modeling approaches to simulate water supplies and demand outcomes and assess options to enhance water sustainability in a changing climate (Singletary and Sterle 2017). The Truckee and Carson River systems rely almost entirely on snowpack-derived water supplies, which are very vulnerable to climate change due to increased temperatures that reduce snowpack storage, increase evaporative water demand, and increase demand from agriculture and urban users (Bureau of Reclamation 2015).

The WftS project followed the process identified in Fig. 1. Primary climate-related concerns, policy constraints, adaptation options, and indicators of impacts to the system were initially identified based on semi-structured interviews of 66 individuals from organizations with regulatory or water management responsibilities (Singletary and Sterle 2017). This initial data collection effort was complemented by semi-annual meetings with a Stakeholder Affiliate Group representing 12 key water management organizations (Sterle and Singletary 2017). These engagements facilitated a broad characterization of the types of events stakeholders were concerned about, followed by a narrowing to the more specific and quantifiable attributes (e.g., timing, magnitude, frequency, duration, extent) of events of greatest concern (Singletary et al. 2016).

Four sets of climate stress test scenarios were developed using a storyline approach and data and methods described in previous sections to explore the potential effects of (1) increased temperatures, (2) extended drought conditions, and (3) increased precipitation variability on water supplies and demands. The fourth set of scenarios focused on a scenario discovery–based sensitivity analysis of water levels in the largest and most upstream reservoir in the system, Lake Tahoe, which is one of the most important indicators of water supply availability. Fundamentally, each of these scenarios was constructed from 4- to 6-km-resolution gridded daily temperature (minimum and maximum) and precipitation datasets spanning the study region. These grids, along with meteorological station data, formed the inputs to the suite of hydrological (GSFLOW, Markstrom et al. 2008; MODFLOW, Harbaugh 2005), open-water evaporation (Complementary Relationship Lake Evaporation (CRLE) model; Huntington and

McEvoy 2011), and operation (MODSIM, http://modsim.engr.colostate.edu/; RiverWare, http://riverware.org/) models that were used to simulate impacts.

The stakeholder process also identified several potential adaptation options, including specific strategies related to farm- and municipal-scale water conservation measures, enhancing water storage, altering crops, and changing reservoir management within the bounds of existing policy constraints (Sterle and Singletary 2017). Several of these strategies were simulated under historical and climate changed conditions (Sterle et al. 2020a, 2020b). Streamflows and reservoir levels at specific times and locations (i.e., system performance indicators) were evaluated under the climate and management scenarios to quantify the ability of the system to meet user demand.

4.1 Increased temperature scenarios (incremental method)

A large majority of the stakeholders interviewed identified increased temperatures as a top concern. They recognized that warming temperatures drive decreased snowpack storage, earlier snowmelt runoff resulting in altered streamflow timing, increased crop evapotranspiration and reservoir evaporation, impacts to water quality, and increased demand by agricultural, municipal, and industrial users (Singletary et al. 2016). A temperature increase of 4.3 °C was applied to historical weather station records from the area to form this scenario using the incremental method (Fig. 2). This increase was selected because stakeholders felt it was an extreme change, and it was informed by an ensemble of 15 downscaled GCM projections responding to moderating (RCP 4.5) and accelerating (RCP 8.5) greenhouse gas emissions where 4.3 °C was the ensemble average temperature change projected by the year 2080 and was also near the maximum projected by the year 2050 (Sterle et al. 2020b). Scenarios consisting of historical precipitation with baseline and warmer temperatures (historical, historical + 4.3 °C) were run through linked hydrological and river operation models to isolate the influence of temperature change on water supplies and demands relative to the historical time period. This scenario set was used to quantify reductions in reservoir storage under warmer temperatures that could occur under status quo, fixed-date, reservoir operations and explore alternative management strategies that might help to mitigate these effects (Sterle et al. 2020a). Use of the incremental approach also enabled a temperature sensitivity analysis to be conducted whereby the hydrology was simulated at 0.5 °C increments of temperature increase from historical up to 4.5 °C. This provided estimates of the potential timing and magnitude of change in percentage of winter precipitation that fell as snow (vs. rain) as a function of temperature change that captured a range of temperature increases projected by the ensemble (as recommended in Sofaer et al. 2017) through the year 2050.

4.2 Extended drought scenarios (splice + incremental)

When the WftS project began in 2015, the Truckee-Carson River System was experiencing the fourth year of a drought that many stakeholders viewed as one of the worst in history (Singletary et al. 2016). Throughout 2012–2015, precipitation was below average with exceptionally low precipitation in 2014, record-breaking high temperatures in both 2014 and 2015, and a record low snowpack in 2015. This resulted in an accumulated precipitation deficit of 100–180% of the annual average, extensive forest die-off, crop loss, native fish loss, and dramatic economic losses across the California-Nevada region (Ullrich et al. 2018). Water levels at Lake Tahoe dropped below the natural rim, ceasing water flows from Lake Tahoe into the Truckee River. This compromised the ability of water providers to meet legal required flows in the Truckee River and fulfill the myriad of downstream demands. As a result, urban water users were encouraged to reduce water use by 10%. Agricultural producers were forced to sell livestock and many suffered significant crop losses. The recreation industry and local economy were also severely impacted, as the meager winter snowpack reduced tourism visits (DeLong 2015).

At the request of the Stakeholder Affiliate Group, an "extended drought" scenario was developed to assess how water supplies would be affected if this drought were to continue for an extended period of time under contemporary temperatures and how this same extended drought might be worsened under mid-century (2050) conditions (Dettinger et al. 2017). To develop a plausible baseline (current temperature) extended drought scenario, the research team used the splice method (Fig. 2) to extend the 2012–2015 drought by following it with a second recent drought (1987–1995), resulting in a 13-year drought sequence. The time series was based on daily gridded PRISM (Daly et al. 2008) data. Because the 1987–1995 time period was cooler than the 2012–2015 period, all daily grid-cell temperatures for the 1987–1995 record were uniformly adjusted by the difference between period temperature averages for 1987–1995 and 2012–2015 in that grid cell. The incremental method was used to develop a climate change version from this historical baseline by adding 2.5 °C to all grid cells and time steps to reflect mid-century conditions. The resulting set of 13-year drought scenarios represented a challenging set of conditions that were used to explore drought impacts to modeled streamflows, reservoir levels, and water supply and demand under current and

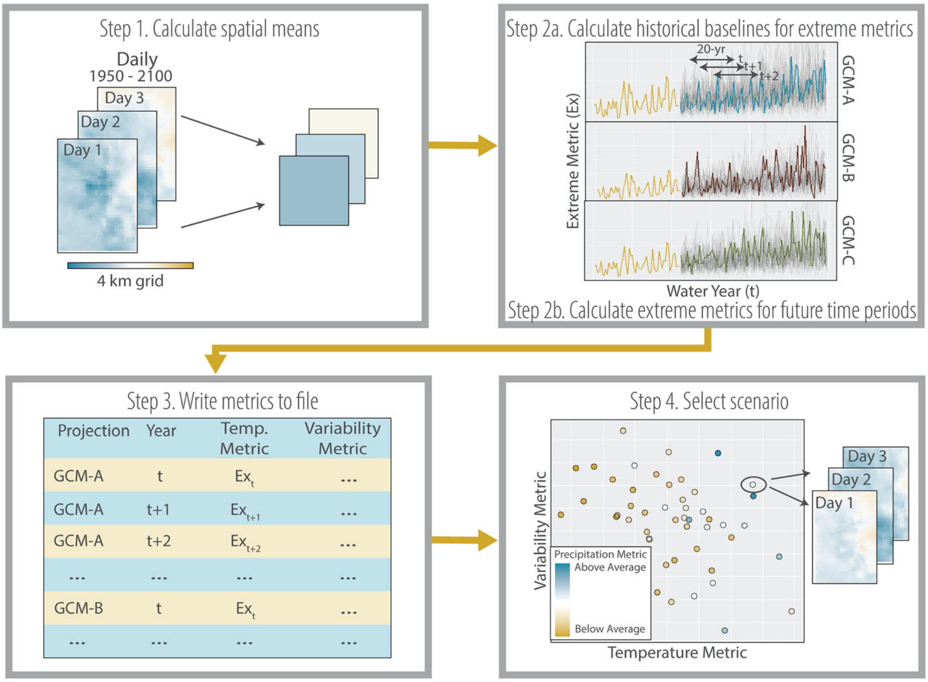

Fig. 3 Workflow for identifying stress tests that was developed for the Water for the Seasons project using the scenario mining technique. For each GCM and time step, a spatial mean is calculated across the study area for each climate variable. Metrics relating to 20-year precipitation variability, average precipitation, and average temperature were calculated on a rolling basis for the historical time period and for each GCM time series. These values were then screened to identify the 20-year period that best matched the criteria for the scenario identified by the stakeholder-researcher team, and the time series corresponding to this was used for the scenario

warmer temperatures, as well as how to minimize these impacts through operational changes using a dynamic decision support system (Boyer 2018).

4.3 Precipitation variability scenarios (scenario mining)

The WftS stakeholder group represented a wide diversity of water use interests, with each having the potential to be challenged by climate change impacts in very different ways. The third set of scenarios was intended to explore some key differences between how water supplies for stakeholders might be differentially impacted in the Truckee and Carson River basins: The Truckee River basin contains significant reservoir storage capacity and provides the majority of municipal water for the Reno-Sparks area, with groundwater providing supplemental water when surface water reserves are insufficient. In contrast, the Carson River basin contains no constructed reservoir storage and most municipal and industrial water supplies are drawn from groundwater and supplemented by surface water. The dominant water use in the basin is for agriculture, and these supplies are drawn directly from surface flows in the Carson River, which are supplemented by water diverted from the Truckee River in lower parts of the basin (Sterle and Singletary 2017). These differences had the potential to result in contrasting responses among the two basins to rapid fluctuations between drought years and wet years (i.e., precipitation whiplash (Swain et al. 2018)) or to multi-year droughts and wet periods. To explore precipitation variability frequency ("whiplash" vs. multi-year) effects, a third set of scenarios was constructed from the GCM ensemble of future climatic episodes with high-frequency (roughly quasi-biennial) precipitation variations vs. lower-frequency (roughly 5-year) precipitation variations. These scenarios illustrate both the general characteristics of the projected future climate (warmer with new precipitation regimes) and specific frequency traits of local interest.

A novel workflow (Fig. 3) for the "scenario mining" method (Fig. 2) was developed as part of this project and used to construct the "precipitation variability" scenarios. To accomplish this, daily temperature and precipitation data from an ensemble of 30 localized constructed analog (LOCA)-downscaled (6-km resolution; Pierce et al. 2014) GCM projections for 2006–2100 were first spatially averaged across the study area. The resulting daily time series from each projection was then screened on a rolling basis to identify 20-year periods that met the following criteria: (1) had similar average temperature (+ 4.3 °C from historical) and precipitation as the "increased temperature" scenarios (see Section 4.1) so that they all could be compared and (2) exhibited dominantly low- (~5-year) or high- (~2–3 year) frequency precipitation variations.

Low-frequency variations (LF) were identified in terms of the ratio of the standard deviations of 5-year moving averages of precipitation ($\mathrm{MA}(P_{5\ \mathrm{yr}})$) to the unfiltered annual totals ($P_{1\ \mathrm{yr}}$) (Eq. 1). Larger values of this ratio indicate relatively greater contributions of 5-year variability to the total variability during the 20-year period and thus a stronger lower frequency variability signal.

$$\mathrm{LF} = \frac{\sigma_{20\ \mathrm{yr}}\left(\mathrm{MA}\left(P_{5\ \mathrm{yr}}\right)\right)}{\sigma_{20\ \mathrm{yr}}\left(P_{1\ \mathrm{yr}}\right)} \tag{1}$$

High-frequency variability (HF) was calculated by subtracting a 2-year moving average precipitation ($\mathrm{MA}(P_{2\ \mathrm{yr}})$) from unfiltered annual totals ($P_{1\ \mathrm{yr}}$) and calculating the ratio of the standard deviation of these values to the standard deviation of the unfiltered annual totals for the 20-year period (Eq. 2). Larger values indicate greater relative contributions of the interannual component and thus higher frequency variability.

$$HF = \frac{\sigma_{20\ \mathrm{yr}}\left(\{P_n\}_{n=1}^{n=20}\right)}{\sigma_{20\ \mathrm{yr}}\left(P_{1\ \mathrm{yr}}\right)}, \text{where } P_n = P_{1\ \mathrm{yr}} - \mathrm{MA}\left(P_{2\ \mathrm{yr}}\right) \tag{2}$$

The 20-year precipitation variability, average temperature, and average precipitation metrics were calculated for all 20-year periods in all climate model projections and in the historical period (Fig. 3, Step 3). These values were then plotted against each other to identify 20-year periods that met all criteria (Fig. 3, Step 4). Variability during future time periods was also compared to a gridded dataset of historical observations upon which the LOCA downscaling is based (Livneh et al. 2013), to serve as a reference point. The 20-year periods that best fit all criteria were then selected, and daily gridded data from the selected ensemble model and time period were extracted to construct the scenario.

Because the hydrologic models used in the WftS project are parameterized with meteorological station data, gridded precipitation projections required additional bias correction in order to reduce the effects of "drizzle," small precipitation amounts that are an artifact of GCM representations and downscaling (Pierce et al. 2014) that were still observed in the data despite the LOCA process including a bias correction step. This "drizzle" resulted in cumulatively large positive biases in the gridded precipitation data as compared to observed. Bias-correction was achieved by modeling the relationship between the square roots of observed and gridded precipitation for the base period 1980–2010 using asynchronous linear regression (Stoner et al. 2013), then applying adjustments to the future projections, based on this modeled relationship.

The variability scenarios highlighted the adverse effects that low-frequency variability would have on most Truckee River water users due to reduced upstream reservoir capture during repeated wet years and lessened reserves during multi-year droughts. They further demonstrated how warming conditions under status quo management differentially affect water users, depending on the basin (lack of storage in the Carson River results in greater impacts from warming) and position in the watershed (downstream water users benefit from uncaptured upstream runoff) (Sterle et al. 2020b). These scenarios were further explored to understand the implications of alternative groundwater management strategies in the Carson River basin, in particular (Sterle et al. 2020b).

4.4 Lake Tahoe water surface elevation scenarios (scenario discovery)

A final set of scenarios was explored to address how climate change may affect future Lake Tahoe water surface elevation. As the largest and uppermost reservoir in the Truckee River system, Lake Tahoe water surface elevation is an important indicator of the water supply situation and is used to trigger operational decisions regarding downstream flows and water allocations to users in both the Truckee and Carson systems under the Truckee River Operating Agreement. Notably, only a small fraction of water in the lake is available for downstream water supply, as the natural rim of the lake is only about 6 ft below the height of the dam. As a result, lake level can only rise so far (about 6 ft) above the natural rim at which point water must be released, but the lake level can fall below that rim in times of drought by tens of feet and more. Thus, stakeholders view lake levels as a particularly strong and impactful indicator of water supply availability in the study area.

To address these questions, 16 downscaled climate projections were used to drive an integrated set of hydrologic, lake evaporation, and water management operations models for the Lake Tahoe basin. Researchers employed the multivariate adaptive constructed analogs set of downscaled GCMs that included estimates of solar radiation and wind speed (Abatzoglou and Brown 2012) so that more precise estimates of lake surface evaporation and evapotranspiration could be developed. The resulting ensemble of projected Lake Tahoe levels was used to estimate probabilities that the Lake will fall below the natural rim and the potential frequencies and magnitudes of this type of system failure under future climate changes.

As described, these calculations modify the description presented above for scenario discovery (Fig. 2), because they were designed to analyze and parameterize the overall ensemble instead of focusing on a particular extreme or challenging time periods within the overall ensemble. Scenario discovery could be applied to the resulting lake-level ensemble, if a more detailed exploration of consequences and causes were desired. One example of this would be to calculate appropriate statistics in all 20-year windows in the lake-level ensemble to identify the cases of lake-level extremes (low and high). Time periods with the fastest declines (or rises) in lake levels could also be identified and retrieved from the ensemble for more study.

5 Discussion

In the preceding sections, we reviewed several ways that climate stress tests can be developed and key advantages and disadvantages (Table 1), as well as concrete examples of how some of these methods were implemented in an ongoing stakeholder-driven water-climate resiliency project. We discuss here some of the key factors that might influence the choice of approach, data, or techniques, drawing, in part, from observations and lessons from our case study. Although we have focused on a climate adaptation planning case study here, we emphasize that the climate scenario construction techniques described herein can be applied in a wide variety of contexts to address weather and climate risks. For example, scenarios are especially useful for emergency response exercises and planning (Committee on Homeland Security and Governmental Affairs 2006; Dettinger et al. 2012; Albano et al. 2016) and many of the same considerations apply in this context.

5.1 Audience

Choices related to approach, data, and methods are ideally based on needs and priorities of the stakeholders involved to ensure that end products are usable (Dilling and Lemos 2011), including their planning horizons, their interests and expertise, and their experiences with climatic events and climate science. For example, if the local planning horizon is 20 years, a stress test drawn from end-of-century climate projections or from the paleoclimatic record may not be well-accepted, even if scientific evidence suggests that it is physically possible for this to occur within that timeframe. Engagement with an audience to understand the limits of what they are willing to explore and how this relates to the scientific evidence regarding climate or weather risks will be required to find the common ground that is necessary for productive response and adaptation planning.

The storyline approach and techniques (all but scenario discovery) tend to draw heavily on the experiences of stakeholders and researchers, including climatic events that have impacted them (or their communities) and connections they see between climate drivers and impacts. The strengths of the storyline approach are that stress tests can be quite straightforward to develop, requiring fewer resources. Moreover, by linking personal experiences to the scenario, stakeholders can easily draw connections between past and future risks (Brewer 2007), understand the impacts more vividly (Marx et al. 2007), and may express greater concern and a stronger desire to adapt (Vasileiadou and Botzen 2014). This approach may resonate in cases when experiences with extreme climatic events have already spurred adaptation actions. In contrast, scenario discovery draws heavily from analytical information and model-based conceptualizations of how climate may impact a study system. These analytical approaches may resonate more with stakeholder groups who tend to think and interpret information analytically.

In the Water for the Seasons case study, the first two scenarios included increased temperature and extended drought scenarios—both drawing from historical data and using simple, storyline approaches to scenario construction. The latter scenario was designed in direct response to an ongoing climatic event that stakeholders wanted to explore and provided an excellent jumping off point for the co-development of climate adaptation options. All these scenarios provided opportunities for developing relationships while exploring familiar but challenging territory at the forefront of stakeholders' minds. Building from that foundation, the group went on to explore less familiar climate changed scenarios, drawing from GCM projections. These later scenarios explored more challenging and unique weather sequences, not experienced historically, providing more opportunity for novel policy considerations and surprise.

5.2 Study goals

While most climate adaptation planning efforts share a common goal of improving management strategies to accommodate climate impacts, study goals may vary depending on (1) the most pressing risks, (2) the scientific information available to the study, and (3) the degree to which the group wishes to conduct modeling and analysis versus leveraging existing knowledge. All of these factors have the potential to influence what data sources or techniques are used and are worth considering in selection and construction of scenarios (Table 1).

The type of climate extreme that the group wishes to explore and the spatial extent of the study area will contribute to the data and methods that can be used. Short-duration extremes on the scale of days to months likely require daily or even sub-daily data. Scenarios in these cases may need to use historical data or dynamically downscaled projections to achieve the desired temporal resolutions. Stochastic weather generator–based sequences may be reasonable to use for a small study area or limited number of locations and short-duration extremes (if properly tuned to provide realistic sub-daily weather transitions), but as the size of the study area or number of locations modeled increases, generating sequences with reasonable spatial autocorrelation structures (Chen et al. 2010) becomes increasingly difficult.

The choice of methods might also vary depending on how much analysis is desired by the group. For example, if the purpose of the exercise or planning effort is geared toward facilitating stakeholder discussions of cascading hazards or adaptation strategies, less analysis may be required than if the main goal is to increase scientific understanding, quantify sensitivities, or estimate risk. For example, for a tabletop emergency response exercise where the scenario is merely a focal point for exploring risks and responses, the scenario may not need to have accurate process representation and internal climatic consistency, whereas this could be quite important for engineering design purposes. For design purposes, estimates of the probability of occurrence of the particular scenario explored might also be of interest, and the raw materials used for scenario mining or discovery can be revisited to provide plausible estimates of that needed ancillary information (Brown et al. 2012; Weaver et al. 2013). In these instances, any estimated probabilities should be interpreted with a clear understanding of the biases and limitations of these model ensembles (Mote et al. 2011), along with their inability to represent true probabilities (Dessai and Hulme 2004).

If one of the objectives of the study is to comprehensively explore sensitivities of a system across a range of uncertainties, the incremental (i.e., for analysis of a defined source of uncertainty) or scenario discovery (for analysis of multiple sources of uncertainty) is a good approach. The advantage of both of the more analytically intensive approaches (mining and discovery) is that once the initial processing of data is complete, there is ample opportunity to explore outputs in response to stakeholder questions and iterations. If there is less interest in quantification and analysis but more interest in detailed exploration of a small number of scenarios as a foundation for information exchange and discussion, the splice or incremental methods may offer a simpler and less computationally demanding alternative.

As described in the previous section, the WftS project explored several of these methods. Since none of the scenarios were focused on short-duration extreme events, the types of data and methods used were less constrained by the need to use sub-daily or fine spatial resolution data, allowing the full range of options to be used. Some of the scenarios (e.g., increased variability) were not attached to any sensitivity or probabilistic analysis, and thus, they could not be compared to each other in a controlled way since they differed in several ways (e.g., timing and phase of precipitation), despite having similar average temperature and precipitation amounts. While these were useful for stimulating discussions with stakeholders about the impacts of increased variability, opportunities to draw robust scientific inferences about the impacts of climate change on variability based on these scenarios were more limited in comparison to the other scenario sets, which allowed for system sensitivity assessment and estimation of (conditional) probabilities.

5.3 Resources available

The resources available for the climate adaptation planning effort, including availability of data, models, technical expertise, and participants' time, are also an important consideration. Although storyline-based approaches can, in some cases, require significant data processing and analytical capacity, they do not have to and are likely to be a better choice under time or resource constraints. Incremental or splice methods based on historical data are technically the simplest while scenario discovery is the most computationally resource intensive. In the scenario mining and scenario discovery methods, up-front investments in data processing provide good opportunities to explore additional stress tests if an initial test does not lead to the desired discussions and discoveries.

The WftS project had significant resources to build upon, given the availability of long-term data; well-developed models of the hydrologic system; several subject matter experts in climate, hydrology, economics, and social sciences; and a stakeholder group that was eager to invest time and energy. This permitted iteration and exploration of a variety of scenarios. Nonetheless, most of the scenarios simulated in the Water for the Seasons project were constructed using relatively simple methods. First, the scenarios were merely an important avenue to explore the greater problems of how the physical and social systems work and fail and what engineering, management, or policy options might be developed to reduce the costs and likelihoods of failure and were thus not a central focus of the study. Second, the Truckee-Carson River system being studied is highly complex and interlinked geographically and hydrologically. In order to properly represent the consequences of scenarios and stakeholder options, the hydrologic simulations and analyses required linking seven different hydrologic and management models. Due to the complexity of modeling this system, the scenario discovery approach was applied only for the Lake Tahoe level analysis. That analysis was focused only on the Upper Truckee watershed, so that it only required three of the seven hydrologic and management models. Despite this only being applied to one part of the system, the scientists on the team deemed this to be a worthwhile investment of resources given the opportunities to continue to explore these simulations far beyond the project life cycle.

6 Conclusions

In this paper, we describe a variety of data sources and techniques available for constructing climate stress tests that are customized to stakeholders' needs. We also provide examples of their implementation and preliminary guidance for determining which options might be best suited for a given application. In summary, we recommend considering the following key factors when selecting a data source or technique for scenario construction: the spatial and temporal extent and resolution needed to characterize the climatic event or time series, how essential it is for the scenario to be physically internally consistent, how different data sources or techniques are likely to be perceived by stakeholders, what resources are available to develop the scenario, and what research questions the team wishes to answer. In addition, we suggest that connecting a scenario to stakeholders' experiences with past climatically stressful events provides a good opportunity to develop trust, gain traction by discussing experiences and concerns that are readily available and tangible, and leverage the practical knowledge they can offer based on their experiences to identify adaptation strategies. From there, the group can collectively work outward to explore impacts and adaptation options under more novel or extreme conditions.

We expect the information provided here to be applicable across a range of contexts, from nearer term emergency response exercises and planning to longer term climate adaptation planning. Generating usable science requires the co-production and shared ownership of information (Dilling and Lemos 2011). Ultimately, this work is intended to support these objectives by providing the necessary guidance to allow scientist-stakeholder teams to develop their own customized stress test scenarios that address their specific weather and climate risks of concern so that specific strategies for mitigating or adapting to these risks can be identified (Weaver et al. 2017).

Acknowledgments This work could not have been accomplished without conversation and insights from Dale Cox (U.S. Geological Survey) and Michael Anderson (California Division of Water Resources) and the contributions from the Stakeholder Advisory Group and research team members of the Water for the Seasons project. We also wish to thank Katherine Hepworth, Erin Delapena Lucas, Shelby Hockaday, Trina Kleist, and Sudhiti Naskar for graphic design support and two anonymous reviewers for feedback that helped to improve this manuscript.

Funding This material is based upon work supported by the Science Applications for Risk Reduction (SAFRR) program in the U.S. Geological Survey's Natural Hazards Mission Area under Grant Nos. G16AC00304 and G18-00405 and through a Water Sustainability and Climate program grant from the National Science Foundation and the U.S. Department of Agriculture (#1360505/1360506).

Data availability Data and material are available upon request.

Compliance with ethical standards

Conflict of interest The authors declare no conflicts of interest.

Code availability Code developed for this study is available upon request.

References

Abatzoglou JT, Brown TJ (2012) A comparison of statistical downscaling methods suited for wildfire applications. Int J Climatol 32:772–780. https://doi.org/10.1002/joc.2312

Albano CM, Dettinger MD, McCarthy MI et al (2016) Application of an extreme winter storm scenario to identify vulnerabilities, mitigation options, and science needs in the Sierra Nevada mountains, USA. Nat Hazards 80:879–900. https://doi.org/10.1007/s11069-015-2003-4

Alexander D (2000) Scenario methodology for teaching principles of emergency management. Disaster Prev Manag An Int J 9:89–97

Amoroso M, Daniels L, Baker P, Camarero J (2017) Dendroecology:tree-ring analyses applied to ecological studies. Springer, Berlin Heidelberg

Anandhi A, Frei A, Pierson DC et al (2011) Examination of change factor methodologies for climate change impact assessment. Water Resour Res 47:1–11. https://doi.org/10.1029/2010WR009104

Ballesteros-Canovas JA, Stoffel M, St George S, Hirschboeck K (2015) A review of flood records from tree rings. Prog Phys Geogr 39. https://doi.org/10.1177/0309133315608758

Boyer WO (2018) A dynamic decision support system for drought resiliency and climate change. University of Nevada, Thesis

Breinl K, Di Baldassarre G, Girons Lopez M et al (2017) Can weather generation capture precipitation patterns across different climates, spatial scales and under data scarcity? Sci Rep 7:5449. https://doi.org/10.1038/s41598-017-05822-y

Brewer GD (2007) Inventing the future: scenarios, imagination, mastery and control. Sustain Sci 2:159–177. https://doi.org/10.1007/s11625-007-0028-7

Brown C, Ghile Y, Laverty M, Li K (2012) Decision scaling: linking bottom-up vulnerability analysis with climate projections in the water sector. Water Resour Res 48:1–12. https://doi.org/10.1029/2011WR011212

Brown C, Wilby RL (2012) An alternate approach to assessing climate risks. EOS Trans Am Geophys Union 93:401–412. https://doi.org/10.1029/2012EO410001

Bryant BP, Lempert RJ (2010) Thinking inside the box: a participatory, computer-assisted approach to scenario discovery. Technol Forecast Soc Change 77:34–49. https://doi.org/10.1016/j.techfore.2009.08.002

Bureau of Reclamation (2015) Truckee basin study. Basin study report. https://www.usbr.gov/watersmart/bsp/docs/finalreport/truckee/tbsbasinstudy.pdf

Chatfield C (2004) The analysis of time series: an introduction, 6th edn. Chapman and Hall/CRC Texts in Statistical Science Ser

Chen J, Brissette FP (2014) Stochastic generation of daily precipitation amounts: review and evaluation of different models. Clim Res 59:189–206. https://doi.org/10.3354/cr01214

Chen J, Brissette FP, Leconte R (2010) A daily stochastic weather generator for preserving low-frequency of climate variability. J Hydrol 388:480–490. https://doi.org/10.1016/j.jhydrol.2010.05.032

Committee on Homeland Security and Governmental Affairs (2006) Hurricane Katrina: a nation still unprepared. Special Report of the Committee on Homeland Security and Governmental Affairs, S. Rept. 109–322

Cook ER, Krusic P (2004) North American summer PDSI reconstructions. IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series # 2004-045. NOAA/NGDC paleoclimatology program. https://www.ncdc.noaa.gov/data-access/paleoclimatology-data

Daly C, Halbleib M, Smith JI et al (2008) Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. Int J Climatol 28:2031–2064

DeLong J (2015) The drought is relentless ; here's what it's affecting. Reno Gaz. Journal. https//www.rgj.com/story/news/2015/06/17/heres-how-the-drought-is-affecting-your-life/28882073/

Department of Homeland Security (2013) Homeland Security Exercise and Evaluation Program (HSEEP). https://preptoolkit.fema.gov/documents/1269813/1269861/HSEEP_Revision_Apr1

Dessai S, Hulme M (2004) Does climate adaptation policy need probabilities? Clim Policy 4:107–128. https://doi.org/10.1080/14693062.2004.9685515

Dettinger MD, Martin Ralph F, Hughes M et al (2012) Design and quantification of an extreme winter storm scenario for emergency preparedness and planning exercises in California. Nat Hazards 60:1085–1111. https://doi.org/10.1007/s11069-011-9894-5

Dettinger MD, Sterle K, Simpson K et al (2017) Climate scenarios for the Truckee-Carson River System. University of Nevada Cooperative Extension Special Publication 17-05

Dilling L, Lemos MC (2011) Creating usable science: opportunities and constraints for climate knowledge use and their implications for science policy. Glob Environ Chang 21:680–689. https://doi.org/10.1016/j.gloenvcha.2010.11.006

Dominguez F, Dall'Erba S, Huang S et al (2018) Tracking an atmospheric river in a warmer climate: from water vapor to economic impacts. Earth Syst Dyn 9:249–266. https://doi.org/10.5194/esd-9-249-2018

Fatichi S, Ivanov VY, Caporali E (2011) Simulation of future climate scenarios with a weather generator. Adv Water Resour 34:448–467. https://doi.org/10.1016/j.advwatres.2010.12.013

Ford JD, Keskitalo ECH, Smith T et al (2010) Case study and analogue methodologies in climate change vulnerability research. Wiley Interdiscip Rev Clim Chang 1:374–392. https://doi.org/10.1002/wcc.48

Gangopadhyay S, Harding BL, Rajagopalan B et al (2009) A nonparametric approach for paleohydrologic reconstruction of annual streamflow ensembles. Water Resour Res 45:1–14. https://doi.org/10.1029/2008WR007201

Gangopadhyay S, Mccabe GJ, Woodhouse CA (2015) Beyond annual streamflow reconstructions for the Upper Colorado River Basin: a paleo-water-balance approach. Water Resour Res 51:9763–9774. https://doi.org/10.1002/2014WR016259

Garb Y, Pulver S, Vandeveer SD (2008) Scenarios in society, society in scenarios : toward a social scientific analysis of storyline-driven environmental modeling. Environ Res Lett 3:045015. https://doi.org/10.1088/1748-9326/3/4/045015

Ghil M, Allen MR, Dettinger MD et al (2002) Advanceed spectral methods for climatic time series. Rev Geophys:40. https://doi.org/10.1029/2000RG000092

Groves DG, Lempert RJ (2007) A new analytic method for finding policy-relevant scenarios. Glob Environ Chang 17:73–85. https://doi.org/10.1016/j.gloenvcha.2006.11.006

Harbaugh A (2005) MODFLOW-2005, The U.S. Geological Survey modular ground-water model—the ground-water flow process. U.S. Geological Survey techniques and methods 6-A16. 253 p

Hazeleger W, Van Den Hurk BJJM, Min E et al (2015) Tales of future weather. Nat Clim Chang 5:107–113. https://doi.org/10.1038/nclimate2450

Huntington J, McEvoy D (2011) Climatological estimates of open water evaporation from selected Truckee and Carson River basin water bodies, California and Nevada. Desert Research Institute. Publication No. 41254

Intergovernmental Panel on Climate Change (IPCC) (2012) Managing the risks of extreme events and disasters to advance climate change adaptation. Special Report of the Intergovernmental Panel on Climate Change. https://www.ipcc.ch/report/managing-the-risks-of-extreme-events-and-disasters-to-advance-climate-change-adaptation/

Karl T, Nicholls N, Ghazi A (1999) CLIVAR/GCOS/WMO workshop on indices and indicators for climate extremes. Clim Chang 42:3–7

Kawase H, Yoshikane T, Hara M et al (2009) Intermodel variability of future changes in the Baiu rainband estimated by the pseudo global warming downscaling method. J Geophys Res 114:1–14. https://doi.org/10.1029/2009JD011803

Kirsch BR, Characklis GW, Zeff HB (2013) Evaluating the impact of alternative hydro-climate scenarios on transfer agreements: practical improvement for generating synthetic streamflows. J Water Resour Plan Manag 139:396–406. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000287

Lemos MC, Kirchhoff CJ, Ramprasad V (2012) Narrowing the climate information usability gap. Nat Clim Chang 2:789

Lempert R (2013) Scenarios that illuminate vulnerabilities and robust responses. Clim Chang 117:627–646. https://doi.org/10.1007/s10584-012-0574-6

Lempert RJ, Popper SW, Bankes SC (2003) Shaping the next one hundred years: new methods for quantitative, long-term policy analysis. RAND, Santa Monica, CA

Livneh B, Rosenberg EA, Lin C et al (2013) A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: update and extensions. J Clim 26:9384–9392. https://doi.org/10.1175/JCLI-D-12-00508.1

Mahmoud M, Liu Y, Hartmann H et al (2009) A formal framework for scenario development in support of environmental decision-making. Environ Model Softw 24:798–808. https://doi.org/10.1016/j.envsoft.2008.11.010

Mahmoud MI, Gupta HV, Rajagopal S (2011) Scenario development for water resources planning and watershed management: methodology and semi-arid region case study. Environ Model Softw 26:873–885. https://doi.org/10.1016/j.envsoft.2011.02.003

Mahoney K, Swales D, Mueller MJ et al (2018) An examination of an inland-penetrating atmospheric river flood event under potential future thermodynamic conditions. J Clim 31:6281–6297. https://doi.org/10.1175/JCLI-D-18-0118.1
Maraun D, Shepherd TG, Widmann M et al (2017) Towards process-informed bias correction of climate change simulations. Nat Clim Chang 7:764–773. https://doi.org/10.1038/nclimate3418
Markstrom SL, Niswonger RG, Regan RS et al (2008) GSFLOW-coupled ground-water and surface-water flow model based on the integration of the precipitation-runoff modeling system (PRMS) and the modular ground-water flow model (MODFLOW-2005). U.S. Geological Survey techniques and methods 6-D1
Marx SM, Weber EU, Orlove BS et al (2007) Communication and mental processes: experiential and analytic processing of uncertain climate information. Glob Environ Chang 17:47–58. https://doi.org/10.1016/j.gloenvcha.2006.10.004
Mckenzie E, Rosenthal A, Bernhardt J et al (2012) Developing scenarios to assess ecosystem service tradeoffs : guidance and case studies for InVEST users. World Wildlife Fund, Washington D.C
Mearns LOO, Hulme M, Carter TR et al (2001) Climate scenario development. In: Houghton J, Ding Y, Griggs D et al (eds) Climate change 2001: the scientific basis, contribution of working group 1 to the third assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp 739–768
Meko D, Graybill DA (1995) Tree-ring reconstruction of upper Gila River discharge. Water Resour Bull 31:605–616. https://doi.org/10.1111/j.1752-1688.1995.tb03388.x
Miller BW, Symstad AJ, Frid L et al (2017) Co-producing simulation models to inform resource management: a case study from southwest South Dakota. Ecosphere 8. https://doi.org/10.1002/ecs2.2020
Mote PW, Brekke LD, Duffy PB, Maurer EP (2011) Guidelines for constructing climate scenarios. EOS Trans Am Geophys Union 92:2010–2012
National Research Council (2013) Climate and social stress: implications for security analysis. The National Academies Press, Washington D.C
Norton-Smith K, Lynn K, Chief K et al (2016) Climate change and indigenous peoples : a synthesis of current impacts and experiences. Gen. Tech. Rep. PNW-GTR-944, Department of Agriculture, Forest Service, Pacific Northwest Research Station. 1–142
Pierce DW, Cayan DR, Thrasher BL (2014) Statistical downscaling using localized constructed analogs (LOCA). J Hydrometeorol 15:2558–2585. https://doi.org/10.1175/JHM-D-14-0082.1
Rowland E, Cross M, Hartmann H (2014) Considering multiple futures : scenario planning to address uncertainty in natural resource conservation. U.S. Fish and Wildlife Service, Washington D.C.
Runyon AN, Carlson AR, Gross J et al (2020) Repeatable approaches to work with scientific uncertainty and advance climate change adaptation in US national parks. Park Steward Forum 36. https://doi.org/10.5070/p536146402
Sauchyn D, Ilich N (2017) Nine hundred years of weekly streamflows: stochastic downscaling of ensemble tree-ring reconstructions. Water Resour Res 53:9266–9283. https://doi.org/10.1002/2017WR021585
Schwarz A, Ray P, Wi S et al (2018) Climate change risk faced by the California Central Valley water resource system. California's fourth climate change assessment. Publication number:CCCA4-EXT-2018-001
Seager R, Cane M, Henderson N et al (2019) Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases. Nat Clim Chang 9:517–522. https://doi.org/10.1038/s41558-019-0505-x
Shepherd TG, Boyd E, Calel RA et al (2018) Storylines: an alternative approach to representing uncertainty in physical aspects of climate change. Clim Chang:555–571. https://doi.org/10.1007/s10584-018-2317-9
Singletary L, Sterle K (2017) Collaborative modeling to assess drought resiliency of snow-fed river dependent communities in the Western United States: a case study in the Truckee-Carson river system. Water (Switzerland) 9:1–19. https://doi.org/10.3390/w9020099
Singletary L, Sterle K, Simpson K (2016) Assessing and enhancing the climate resiliency and adaptive capacity of the Truckee-Carson River System: results of a survey of local organizations. Spec Publ 16-03:1–48
Skamarock WC, Klemp JB, Dudhia J et al (2008) Description of the advanced research WRF version 3. CAR Technical Note, NCAR/TN-475+STR
Snover AK, Mantua NJ, Littell JS et al (2013) Choosing and using climate-change scenarios for ecological-impact assessments and conservation decisions. Conserv Biol 27:1147–1157. https://doi.org/10.1111/cobi.12163
Sofaer HR, Barsugli JJ, Jarnevich CS et al (2017) Designing ecological climate change impact assessments to reflect key climatic drivers. Glob Chang Biol 23:2537–2553. https://doi.org/10.1111/gcb.13653
Sorooshian S, Martinson DG (1995) Proxy indicators of climate: an essay. In: Natural climate variability on decade-to-century time scales. The National Academies Press, Washington D.C., pp 489–494
Star J, Rowland EL, Black ME et al (2016) Supporting adaptation decisions through scenario planning: enabling the effective use of multiple methods. Clim Risk Manag 13:88–94. https://doi.org/10.1016/j.crm.2016.08.001
Sterle K, Jose L, Coors S et al (2020a) Collaboratively modeling reservoir reoperation to adapt to earlier snowmelt runoff. J Water Resour Plan Manag 146. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001136

Sterle K, Singletary L (2017) Adapting to variable water supply in the Truckee-Carson River System : results of focus groups conducted in 2016 with local water managers. Univ Nevada Coop Ext Spec Publ 17-15

Sterle K, Singletary L, Lee G et al (2020b) Water sustainability and climate in the Truckee-Carson River System , western United States 10 key takeaways from the Water for the Seasons collaborative research program. University of Nevada, Reno Extension Special Publication SP-20-02

Stern PC, Ebi KL, Leichenko R et al (2013) Managing risk with climate vulnerability science. Nat Clim Chang 3: 607–609. https://doi.org/10.1038/nclimate1929

Stoner AMK, Hayhoe K, Wuebbles DJ (2013) An asynchronous regional regression model for statistical downscaling of daily climate variables. Int J Climatol 33:2473–2494. https://doi.org/10.1002/joc.3603

Swain DL, Langenbrunner B, Neelin JD, Hall A (2018) Increasing precipitation volatility in twenty-first-century California. Nat Clim Chang 8:427–433. https://doi.org/10.1038/s41558-018-0140-y

Symstad AJ, Fisichelli NA, Miller BW et al (2017) Multiple methods for multiple futures: integrating qualitative scenario planning and quantitative simulation modeling for natural resource decision making. Clim Risk Manag 17:78–91. https://doi.org/10.1016/j.crm.2017.07.002

Tariq A, Lempert RJ, Riverson J et al (2017) A climate stress test of Los Angeles' water quality plans. Clim Chang 144:625–639. https://doi.org/10.1007/s10584-017-2062-5

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 93:485–498

Ullrich PA, Xu Z, Rhoades AM et al (2018) California's drought of the future: a midcentury recreation of the exceptional conditions of 2012–2017. Earth's Futur 6:1568–1587. https://doi.org/10.1029/2018EF001007

Van Der Merwe L (2008) Scenario-based strategy in practice: a framework. Adv Dev Hum Resour 10:216–239. https://doi.org/10.1177/1523422307313321

Van Der Wiel K, Selten FM, Bintanja R et al (2020) Ensemble climate-impact modelling: extreme impacts from moderate meteorological conditions. Environ Res Lett 15. https://doi.org/10.1088/1748-9326/ab7668

Vano JA, Kim JB, Rupp DE, Mote PW (2015) Selecting climate change scenarios using impact-relevant sensitivities. Geophys Res Lett 42:5516–5525. https://doi.org/10.1002/2015GL063208

Vasileiadou E, Botzen WJW (2014) Communicating adaptation with emotions: the role of intense experiences in raising concern about extreme weather. Ecol Soc 19. https://doi.org/10.5751/ES-06474-190236

Weaver CP, Lempert RJ, Brown C et al (2013) Improving the contribution of climate model information to decision making: the value and demands of robust decision frameworks. Wiley Interdiscip Rev Clim Chang 4:39–60. https://doi.org/10.1002/wcc.202

Weaver CP, Moss RH, Ebi KL et al (2017) Reframing climate change assessments around risk: recommendations for the U.S. National Climate Assessment. Environ Res Lett 12:

Werner AT, Cannon AJ (2016) Hydrologic extremes - an intercomparison of multiple gridded statistical downscaling methods. Hydrol Earth Syst Sci 20:1483–1508. https://doi.org/10.5194/hess-20-1483-2016

Whateley S, Walker JD, Brown C (2015) A web-based screening model for climate risk to water supply systems in the northeastern United States. Environ Model Softw 73:64–75. https://doi.org/10.1016/j.envsoft.2015.08.001

Wilks DS, Wilby RL (1999) The weather generation game: a review of stochastic weather models. Prog Phys Geogr 23:329–357. https://doi.org/10.1191/030913399666525256

Wilsey CB, Lawler JJ, Maurer EP et al (2013) Tools for assessing climate impacts on fish and wildlife. J Fish Wildl Manag 4:220–240. https://doi.org/10.3996/062012-JFWM-055

Woodhouse CA, Lukas JJ (2006) Multi-century tree-ring reconstructions of Colorado streamflow for water resource planning. Clim Chang 78:293–315. https://doi.org/10.1007/s10584-006-9055-0