AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE

10.1002/2013WR014561

Key Points:

- Traditional methods treat trend detection separate from decision analysis
- Under-preparation risks due to failure to detect trends are ignored
- Our approach combines hypothesis test and risk-based decisions

Correspondence to:
A. Rosner,
arosner@usgs.gov

Citation:
Rosner, A., R. M. Vogel, and P. H. Kirshen (2014), A risk-based approach to flood management decisions in a nonstationary world, *Water Resour. Res.*, *50*, doi:10.1002/2013WR014561.

Received 15 AUG 2013
Accepted 28 JAN 2014
Accepted article online 3 FEB 2014

A risk-based approach to flood management decisions in a nonstationary world

Ana Rosner[1], Richard M. Vogel[2], and Paul H. Kirshen[3]

[1]U.S. Geological Survey Conte Anadromous Fish Research Center, Turners Falls, Massachusetts, USA, [2]Department of Civil and Environmental Engineering, Tufts University, Medford, Massachusetts, USA, [3]Civil Engineering Department and Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, New Hampshire, USA

Abstract Traditional approaches to flood management in a nonstationary world begin with a null hypothesis test of "no trend" and its likelihood, with little or no attention given to the likelihood that we might ignore a trend if it really existed. Concluding a trend exists when it does not, or rejecting a trend when it exists are known as type I and type II errors, respectively. Decision-makers are poorly served by statistical and/or decision methods that do not carefully consider both over- and under-preparation errors, respectively. Similarly, little attention is given to how to integrate uncertainty in our ability to detect trends into a flood management decision context. We show how trend hypothesis test results can be combined with an adaptation's infrastructure costs and damages avoided to provide a rational decision approach in a nonstationary world. The criterion of *expected regret* is shown to be a useful metric that integrates the statistical, economic, and hydrological aspects of the flood management problem in a nonstationary world.

1. Introduction

Within the context of flood management in a nonstationary setting, the first question normally posed is whether or not a statistically significant nonstationary flood signal is apparent. As a result, there are now hundreds, possibly thousands, of studies which have sought to evaluate whether or not statistically significant trends have been observed in historical flood records. All of the many previous studies we have reviewed which have sought to determine whether a trend exists in flood discharges, storm surges, precipitation extremes, or other hydroclimatic processes, have employed a null hypothesis, H_0, of no trend and most have chosen a significance level of $\alpha = 0.05$ (see reviews by *Easterling et al*. [2000], *Nicholls* [2000], and *Huntington* [2006]). A significance level of 0.05 implies that if there really is no trend (that is the assumption of H_0), we will only (mistakenly) report trends 5% of the time. The societal consequences of making such a mistake are that we will prepare for a flood trend, when it does not exist, which we term over-preparedness or over-investment. Should society not also be interested in the likelihood of under-preparedness? Surely there are situations in which society will regret having been under-prepared for consequences of events which could have been avoided [see *Trenberth*, 2011; *Vogel et al*., 2013, for further discussion]. Our brief review reveals that most previous studies concerning hydroclimatic trends have assumed a null hypothesis of stationarity or no hydroclimatic trends. *Trenberth* [2011] argues that because "global warming is now unequivocal and very likely caused by human activities" that our null hypothesis should now be reversed to the conditional assumption that hydroclimatological regimes are now nonstationary. We agree with *Matalas* [2012] who argues that regardless of whether our future world is stationary or nonstationary, traditional decision oriented and statistical decision methods can still play a critical role in water resources management.

Statistical analysis of a null hypothesis of no trend, termed Null-Hypothesis Significance Testing (NHST), focuses only on type I error, α, of concern under stationary conditions, because all such hypothesis tests were derived under conditions of no trend. Type II error, β, of interest when one is more concerned about missing the effect of the alternative nonstationary hypothesis, H_A, is usually ignored and H_A is too often dismissed. The decision matrix for the general trend detection decision problem is depicted in Figure 1 along with the type I and type II error probabilities α and β. Statisticians define the term "power" as the likelihood of detecting a trend, when it exists which is equal to $1 - \beta$ in Figure 1. Of particular concern to us are the likelihood and consequence of type II errors, which are entirely out of our control, because it is only the probability of a type I error α that can be specified (and thus controlled) in a hypothesis test. Here type II

10.1002/2013WR014561

	No Trend in Floods H_o	Trend in Floods H_A
Do Not Adapt	$1-\alpha$	β Type II Error (under-prepare)
Adapt	α Type I Error (over-invest)	$1-\beta$ Power

Figure 1. Decision matrix and definitions of type I and type II errors.

errors (under-preparation) involve significant societal consequences because they imply no societal response is necessary when one is warranted. For example, the physical implication of a type I or over-preparedness error in adaptation decisions for flood management is wasted money on unneeded infrastructure. The physical repercussions of a type II or under-preparedness error, on the other hand, are major flood damages due to inadequate protection. Decision-makers are poorly served by statistical and/or decision methods that do not carefully consider both sources of error, which is a central point of *Vogel et al*. [2013]. The brief communication by *Vogel et al*. [2013] focused on the likelihood associated with both sources of error; this study extends those ideas by introducing a methodology for considering both sources of error into a rational decision process for making adaptation decisions in what may be either a stationary or nonstationary world.

Numerous fields including psychology, economics, social sciences, meteorology, and medical research, have called into question the value of NHST tests due to its focus on a single, often arbitrary, significance level α [*Ziliak and McCloskey*, 2008; *Cohen*, 1994; *Nicholls*, 2000] among other concerns. *Cohn and Lins* [2005] stated these concerns succinctly when they said: "Because statistical tests are proof by contradiction, any inconsistency between the null hypothesis and the natural system can itself lead to rejection of the null hypothesis." Concerns over the use of NHST are now widespread, though remarkably, none of those studies we have reviewed dwell on the most important criticism of all, that of ignoring the probability of type II errors (see, for example, critiques by *Cohen* [1994], *Nicholls* [2000], and *Ziliak and McCloskey* [2008]). It is our goal to develop a methodology which integrates both the probabilities of type I and type II errors into a rational decision framework for weighing the consequences of those errors and for making decisions under uncertainty and potential nonstationarity.

Criticisms about NHST are of vital concern to the fields of geophysics, climate science, and water resources engineering, where the trend analysis could have an impact on major infrastructure decisions. It is only very recently and rarely that researchers have raised concern over the importance and impacts of type II errors in the climate and hydrologic sciences [*Cohn and Lins*, 2005; *Trenberth*, 2011; *Morin*, 2011; *Ziegler et al*., 2003, 2005]. Though those studies discussed the importance of considering type II errors in the analysis of trends, they did not consider the resulting impacts on infrastructure decisions and societal preparedness, as is our primary focus. A type II error in the context of an infrastructure decision implies under-preparedness, which can be more costly to society than the type I error (over-preparedness) which the NHST focuses on. For example, *ASCE* [2007] concluded that much of the nearly $150 billion in flood damage caused by Hurricane Katrina resulted from underdesign of levees and other components of the New Orleans hurricane protection system. *Sarewitz et al*. [2003] document that an invalid assumption of stationarity can lead to increases in vulnerability due to underestimation of flood risk. Note that type II errors corresponding to under-preparedness are paramount even in a stationary world as was rigorously shown by *Stedinger* [1982] for risks posed by floods.

The common and traditional trend analyses focus only on our understanding of conditions of no trend, because NHST was derived under conditions (null hypothesis) of no trend. Of particular concern to us is that, due to the design of the hypothesis test, we cannot control the type II errors when making decisions that hinge on whether there is a trend. These types of error signify substantial societal consequences, the result of having decided not to adapt to a nonstationary world, when the world IS nonstationary. There are two possible approaches to integration of a trend hypothesis test into a decision framework: (a) derive a new suite of hypothesis tests, so that H_0 is the case of a trend and H_A is no trend, in which case the type I error corresponds to the outcome with the most serious impacts; or (b) keep H_0 as the "no trend" hypothesis, estimate the type I and II errors, and use both error probabilities in evaluating the statistical evidence for the trend. We favor the latter approach.

Given the significant attention given to the problem of flood management in a nonstationary world [see *Milly et al*., 2008], we are surprised that:

1. All hypothesis tests still focus on the type I or "overdesign" error when the type II or "underdesign" error is often of equal or more relevance to society. Of critical concern is the possibility promoted by *Trenberth* [2011] that "As a whole the community is making too many type II errors."

2. Few if any studies have gone to the trouble of estimating the probability of under design, β, or its complement $1-\beta$ known as the power of the test; we could only find a few papers in the water and climate literature which discuss the likelihood of type II errors [*Lettenmaier*, 1976; *Bowling et al*., 2000; *Ziegler et al*., 2003, 2005; *Morin*, 2011].

3. We are unaware of anyone who has developed a hypothesis test which focuses attention on a null hypothesis of "trend," regardless of whether the evidence for that trend has already been established in the scientific literature.

4. We are unaware of any studies which have sought to consider both of these probabilities in a risk-based decision framework, as is the goal of this study.

We begin by reviewing risk-based decision making (RBDM) methods formulated under stationary conditions. We then provide background on trend detection and the estimation of the type I and type II error probabilities. We describe the traditional approach which treats trend detection and RBDM as independent steps in the decision process. Next, we introduce a new method risk-based approach which integrates trend detection with RBDM. We bring in the concept of regret, which is extremely important in the context of decision making under uncertainty, because it reflects the difference between the benefits associated with a particular option and the benefits associated with the best option available if one had perfect foresight. We introduce the calculation of expected regret, which combines the trend detection error probabilities with expected damages and infrastructure costs. Finally, we then combine all the elements of our analysis into a case study application for a coastal flood management adaptation decision in Mystic, Connecticut, U.S.

2. Risk-Based Decision Making

Risk-based decision making (RBDM) is a well-established methodology that determines appropriate levels of infrastructure based on the expected damages avoided versus the cost of the infrastructure required [*Tung*, 2005, *U.S. Army Corps of Engineers*, 2000; *National Research Council*, 2000]. For example, the *U.S. Army Corps of Engineers* [2000, section 3.3] recommends that flood damage reduction studies are to be conducted using an analytical, risk-based approach calculating expected performance including the use of stage-damage functions, a probabilistic display of benefits and costs, and residual damages. RBDM can be used in place of the traditional design storm approach which selects a particular design event (a specific T-year event usually specified by regulation), and then selects the necessary infrastructure to protect against the flood event with that specified average return period T. Instead, the goal of RBDM is to choose a level of infrastructure protection that minimizes the total expected cost, which is a sum of the costs of infrastructure and of the expected residual damages as shown in Figure 2. One can also use RBDM to evaluate specific infrastructure alternatives by calculating net benefits: the expected cost of damages avoided, less the cost of infrastructure. A positive net benefit indicates that the alternative is economically attractive. While protection against the $T=100$ year flood is the most common design target under traditional analysis, a RBDM process may lead to a protection target either smaller or larger than the 100 year flood, depending on the probability and the consequences of the flood as well as the costs of the needed infrastructure.

3. Traditional Decision-Making Process in a Nonstationary World

When one employs a traditional decision-making process, whether based on an arbitrary design event or RBDM, a trend is evaluated for statistical significance separately from the economic project evaluation. Such an approach treats trend detection as a purely statistical issue, separate from the decision problem. First, a hypothesis test is performed and the statistical significance α of the trend is estimated. If α is below a prespecified critical value, usually $\alpha_{critical}=0.05$, the economic analysis is performed to evaluate the economic viability of a proposed flood management plan. If α exceeds the critical value, the trend is dismissed and the economic analysis is not performed. If the outcome of the hypothesis test does not definitively indicate that a trend is evident, the consequences of under-preparation, however substantial, would not even be computed and would not be considered in the analyses. Instead we recommend use of statistical decision

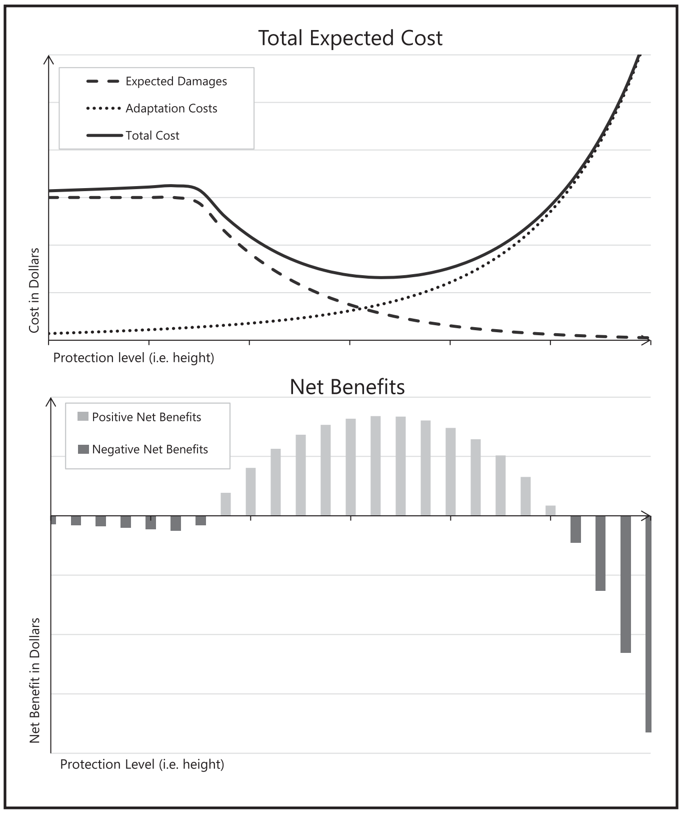

Figure 2. Risk-based decision making: choosing protection level through either minimization of total expected cost or net benefits.

theory which combines both the statistical knowledge gained from the trend detection with classical decision theory.

4. A Risk-Based Approach to Flood Management in a Nonstationary World

We propose a new approach which uses both risk-based decision theory and hypothesis testing, together. Our approach integrates the uncertainty inherent in the trend detection process, with the probabilistic nature of floods and the economic analysis of infrastructure alternatives. The resulting statistical decision process enables the decision-maker to ask the question whether enough information is available to warrant making a particular adaptation decision, and whether the economic impact of a trend is great enough that it is advisable to plan for increased storms, even if the statistical significance of the trend falls short of some customary arbitrary threshold.

The statistical trend detection yields both type I and type II error probabilities based on historical evidence up to the date of the decision, as well as an estimate of the trend's magnitude for use in the economic analysis. Flood damage estimates are needed both with and without adaptation, under both stationary and nonstationary conditions. All of this information is then combined to determine the potential damages avoided by adaptation whether or not we experience a nonstationary world. The damages avoided may not exceed the cost of the adaptation if we plan for a flood trend that never materializes. Similarly, we may incur damage costs that could have been prevented by implementation of an adaptation strategy if a trend that we were not expecting actually materializes. Each of these avoidable costs is termed regret and form the basis of our economic analysis.

The idea here is to integrate all relevant information concerning adaptation options and flood outcomes. The resulting decision is based on two different measures of "expected regret": the expected cost of avoidable damages if we under-prepare; and the expected cost of unneeded infrastructure if we over-invest. Expected regret integrates both the statistical (un)certainty and the potential consequences of a trend, if it exists. The expected regret of over-investment can then be weighed against the expected regret of under-preparation. Generally, one would recommend to invest in adaptation when the expected regret of under-preparation is greater than the expected regret of over-preparation, and to not adapt otherwise. The proposed method provides an integrated understanding of the uncertainty associated with our knowledge about flood trends along with the economic consequences of various adaptation plans in an intuitive, meaningful way that will help stakeholders make well-informed decisions for climate change adaptation.

5. Trend Testing and the Probability of Under- and Over-preparation

Adaptation planning in the context of flood management depends critically on trend detection, hence it is important to understand the limitations and concerns surrounding statistical trend tests. This study employs both the type I and II error probabilities associated with trend tests, while acknowledging the tremendous uncertainty associated with our ability to discern trends from other natural phenomenon such as persistence [see *Cohn and Lins*, 2005] as well as complications due to seasonality, censoring, change points, skewness, and other issues *Helsel and Hirsch* [2002], *Khaliq et al*. [2009], *Kropp and Schellnhuber* [2011], and *Sonali and Kumar* [2013] provide an overview of recent innovations in trend detection with attention given to most of the above-mentioned complications involving detection of trends. In this initial effort to develop a decision-oriented methodology for incorporation of flood trend hypothesis testing into flood management we ignore those complications. A natural extension to this study would be to explore how these various complications influence our ability to make effective flood management decisions.

One of the main arguments against Null-Hypothesis Significance Testing (NHST) is its adherence to a single critical value $\alpha_{\mathrm{critical}}$ of 0.05. Concerns about NHST are of vital concern to climate sciences [see *Nicholls*, 2000; *Trenberth*, 2011] and water resources engineering, where the trend analysis could have an impact on major infrastructure decisions. Use of NHST implicitly places disproportionate emphasis on the type I error probability α, while the power $1-\beta$ is rarely reported, despite the importance and linkage between type II errors, under-preparedness, and its likelihood in flood management applications. Our approach avoids the need to define a critical value for either the type I or type II error probability and gives crucially needed emphasis on type II under-preparation errors.

Lettenmaier [1976] first introduced to the water resources literature analytical expressions for the power of a hypothesis test based on ordinary least squares (OLS) regression in the context of trend detection in water quality management. *Bowling et al*. [2000] performed a similar analysis to determine the minimum detectable difference or the smallest trend one could discern to be statistically significant. Remarkably, of the hundreds and possibly thousands of studies which have examined trends in hydro climatic variables, we could only find a few studies which computed the probability of type II errors. For example, *Ziegler et al*. [2005] used GCMs to predict trends in annual precipitation on the Mississippi basin, and then performed simulations to determine the minimum length of record which would be needed to detect trends of those magnitudes. Focusing on the Mississippi River basin, they calculated the detection time required to predict the magnitude of trends predicted by the GCMs. They found that between 82 and 143 years would be required to detect the trend corresponding to type I and II error probabilities of $\alpha=0.05$ and $\beta=0.10$, respectively. *Ziegler et al*. [2003] performed similar evaluations for three GEWEX basins. *Ziegler et al*. [2003, 2005] employed a simple analytical approximation to the power of a *t* test introduced earlier by *Lettenmaier* [1976]. *Morin* [2011] performed a similar analysis using Monte-Carlo simulations to estimate the minimum magnitude of change in annual precipitation at over 9000 stations globally, that could be detected over a 50 year period. He reports minimum detectable trends given type I and II error probabilities of $\alpha=0.05$ and $\beta=0.50$.

In this study, we employ a simple linear model to characterize trends in flood levels. Here we only consider a trend in the mean of the natural logarithms of the annual maximum flood series as a function of time. More complex trend analyses are possible by incorporating other covariate predictors of the trend such as climatic indices [*Jain and Lall*, 2001; *Kwon et al*., 2008] and/or trends in other moments [*Villarini et al*., 2009]. *Vogel et al*. [2011] found that a linear model relating the logarithm of instantaneous annual maximum streamflow and its year of occurrence provided an excellent approximation for thousands of river gages across the continental U.S. Even for highly nonlinear trends, ordinary least squares (OLS) regression can often provide a good approximation by employing the “ladder of powers” to linearize the relationship. *Mosteller and Tukey* [1977] provide a guide to selecting appropriate (and possibly different) power transformations of *y* and *x* using a plot of *y* versus *x* and their so-called “bulging rule” [*Helsel and Hirsch*, 2002, Figure 9.5]. Given the power transformation x^θ, and y^θ, going up the “ladder of powers” corresponds to setting $\theta>1$ (i.e., x^2, x^3, etc.), and going down the ladder of powers means setting $\theta<1$, (i.e., ln(x), 1/x, $\sqrt{x}$, etc.).

Interestingly, even though exact analytical expressions exist for computing the power of a trend test based on OLS regression, we found it quite difficult to locate textbooks or primer papers which document such analyses. This is especially surprising given the widespread use of linear regression for performing trend analyses. *Lettenmaier* [1976], *Rothenberg* [1988], and *Dupont and Plummer* [1990, 1998] describe an analytical calculation of β for a linear regression.

Consider a simple linear model between the natural logarithm of the annual maximum flood *y* and its year of occurrence *x* so that $y=\ln(q)=\beta_o+\beta_1 x+\varepsilon$, where *q* is the annual maximum flood and β_o and β_1 are

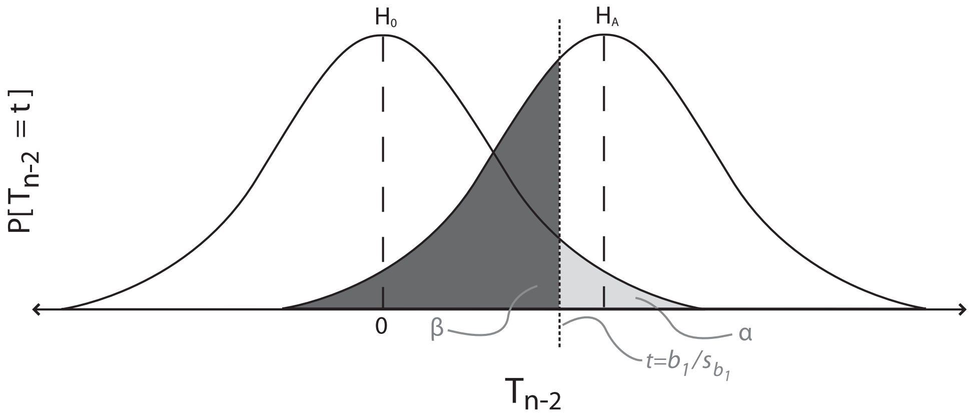

Figure 3. Student's t probability distributions of H_A and H_0, showing type I and type II error associated with estimation of slope (trend) coefficient in linear regression model.

regression coefficients. Such a model is highly nonlinear in real space and appears to capture flood trends over wide geographic regions in the U.S. [see *Vogel et al.*, 2011] with approximately independent, homoscedastic and normally distributed residuals ε, required to perform meaningful statistical inference on the fitted model. Our methodology can be applied regardless of the sign of the linear model slope (or trend) term. The fitted model takes the form $y=\ln(q)=b_o+b_1x$ with parameters b_o and b_1 computed from a time series of length n. It is important to realize that this simple linear regression model is a model of the conditional mean of y so that $E[y|x]=b_o+b_1x$ where $E[y|x]$ denotes the expectation of y conditioned on the value of x. This fact is used later on to provide a nonstationary model of flood magnitudes and frequency.

The trend test amounts to a Students t test on the estimate of the slope term b_1 as illustrated in Figure 3. Given the null hypothesis H_0: $b_1=0$ versus the one-sided alternative hypothesis H_A: $b_1>0$ one can estimate the type I error probability using

$$\alpha=1-F[t] \tag{1}$$

where F denotes the cumulative distribution function of a Student's t random variable with $n-2$ degrees of freedom, T_{n-2}, and $t=b_1/s_{b_1}$,where b_1 is the OLS estimate of the trend slope and s_{b_1} is an estimate of the standard deviation of b_1. Note from Figure 3 that α is simply the shaded region depicted to the right of the value of t under the null hypothesis H_0. Similarly, the probability of the type II error β corresponding to the value of α determined in (1) corresponds to the shaded region to the left of the value of t, under the alternative hypothesis H_A. Under H_A, the true trend slope is assumed to be known to be β_1 in which case the value of $t'=t_{1-\alpha,n-2}-(\beta_1/\sigma_{b_1})$ follows a Student's t distribution with $n-2$ degrees of freedom so that the type II error probability is simply $\beta=F(t')$. Here $t_{1-\alpha,n-2}$ is that value of a Student's t random variable with $n-2$ degrees of freedom and nonexceedance probability $1-\alpha$. To provide a general expression for β, we use the basic theoretical expressions corresponding to the linear model $y=\ln(q)=\beta_o+\beta_1x+\varepsilon$ including the facts that $\beta_1=\rho\sigma_y/\sigma_x$, $\sigma_{b_1}=\sigma_\varepsilon/(\sigma_x\sqrt{n})$, and $\sigma_\varepsilon/\sigma_x=\sqrt{1-\rho^2}$, where σ_y, σ_x, and σ_ε are the standard deviation of, y, x, and ε respectively, and ρ is the correlation coefficient between y and x. Now combining all these facts, the expression for $\beta=F(t')$ can be written as:

$$\beta=F\left(t_{1-\alpha,n-2}-\delta\sqrt{n}\right) \tag{2}$$

where $\delta=\frac{1}{\sqrt{\frac{1}{\rho^2}-1}}$

where $t_{1-\alpha,n-2}$ is that value of a Student's t random variable with $n-2$ degrees of freedom and nonexceedance probability $1-\alpha$. The result in (2) corresponds to the use of a one-sided test, because our assumption here is that a positive trend exists. *Lettenmaier* [1976], *Dupont and Plummer* [1990], and *Bowling et al.* [2000] provide derivations which result in equivalent expressions for β given above, though they consider the two-sided alternative hypothesis that the trend slope could be either positive or negative. Our result in (2)

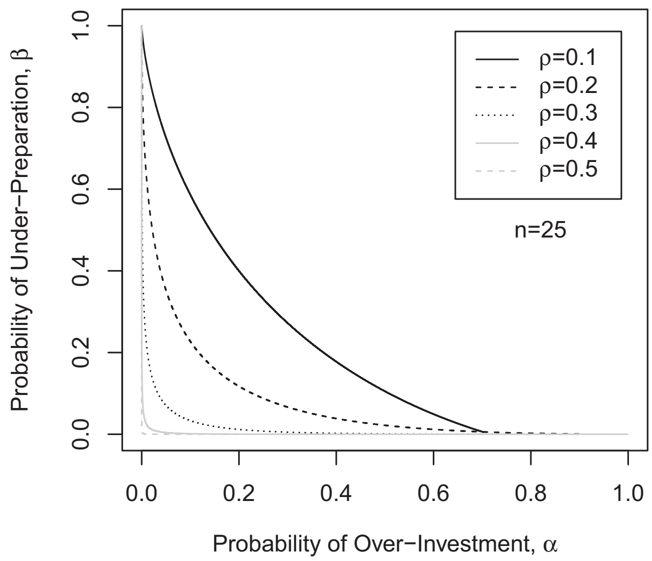

Figure 4. Relationship between probability of under-preparedness and over-investment, β and α, respectively, as a function of the goodness of fit of the trend model ρ, and the length of record *n*, used to fit the trend model.

applies to situations in which one expects either a positive or negative trend, a priori. Equations (1) and (2) only apply to situations in which one has fit a linear trend model using ordinary least squares regression resulting in model residuals ε which are homoscedastic, independent, and normally distributed.

The values of α and β are inversely related to each other as shown in Figure 4 and their relationship only depends on the values of n and ρ. Note that the trend term b_1 is related to ρ via the relation $b_1 = \rho\sigma_y/\sigma_x$, where σ_x and σ_y are the standard deviation of x and y, respectively. Note that no correlation implies no trend ($b_1 \rightarrow 0$ as $\rho \rightarrow 0$) and a perfect correlation $\rho \rightarrow 1$ implies a trend term equal to $b_1 = \sigma_y/\sigma_x$.

Several important conclusions can be drawn from the inverse relationship between α and β shown in Figure 4 that have valuable implications for weighing flood management risks. Recall from Figure 1 that the values of α and β may be interpreted as the probability of over-investing or and under-preparing, respectively. Further, they are inversely related so that, when n and ρ are fixed, to ensure a very low probability of over-investment, one must accept a fairly high probability of under-investment. Only if the values of n or ρ are increased will the values of both α and β decrease. Increasing the value of n is tantamount to waiting for additional years of data, which might mean waiting until it is too late. Increasing the value of ρ if possible, is a much more viable option, as it could result from improvements in our ability to perform trend detection, attribution and prediction. In this initial study, we employ the parametric linear regression method for estimating the type II error probability. We caution the reader that when parametric tests are applied to nonlinear trend detection problems which cannot be linearized using the ladder of powers, their power is lower than for equivalent nonparametric tests [see *Helsel and Hirsch*, 2002].

6. Risk-Based Decisions Using Decision Trees

A decision tree is the graphical equivalent of a stochastic dynamic program and is an ideal tool for implementation of sequential statistical decision problems. It describes the sequence of possible decisions for numerous alternatives along with their probabilistic and economic outcomes. It is a very powerful statistical decision approach because it combines a graphical representation of the overall set of alternatives and decisions, with a framework for making decisions in light of expected outcomes. *Fiering and Matalas* [1990] provide one of the earliest examples of a statistical decision process for evaluating various alternatives in the context of climate change. *Chao and Hobbs* [1997] give a brief history of decision analysis applications to climate change; and apply a mathematical version of a decision tree known as a stochastic dynamic program for evaluating breakwater adaptation under possible climate change impacts on Lake Erie. *Hobbs et al.* [1997] were the first to apply a decision tree approach to water resources management under climate change. They demonstrate use of a Bayesian approach to analyzing the necessary probabilities in the decision tree for evaluating alternative adaptation strategies for climate change for the Great Lakes. Approaches such as these require a number of assumptions to generate values for the probabilities required. For example, *Hobbs et al.* [1997] and *Manning et al.* [2009] describe a Bayesian analysis consisting of aggregating predictions from suites of model predictions, such as Global Circulation Models (GCMs) or Regional Circulation Models (RCMs). We are unaware of any previous studies which have employed a decision tree along with

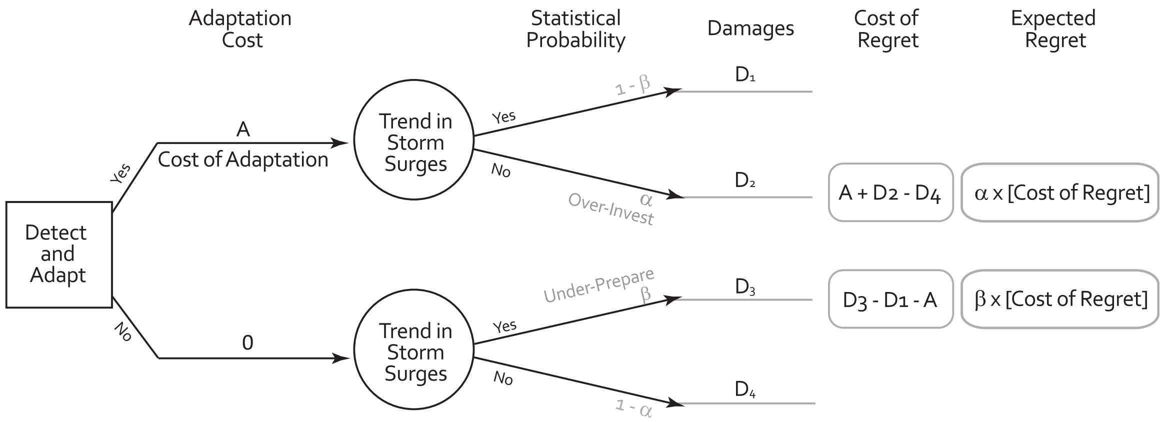

Figure 5. Decision tree for risk-based flood adaptation in a nonstationary world. Estimates of expected damages and probabilities of type I and II errors, α and β, are used to calculate expected regret, integrating statistical uncertainty and potential physical consequences of a trend to inform adaptation decisions.

the probabilistic results of trend hypothesis testing for evaluating alternative adaptation strategies for responding to climate change.

A decision tree for our coastal adaptation problem is shown in Figure 5. As is the convention for decision trees, a square represents decision outcomes while circles represent chance outcomes, which in this case, are floods due to coastal storm surges. The four nodes emanating from the chance nodes are represented by the four physically defined outcomes, corresponding to the four possible trend detection hypothesis test outcomes represented in Figure 1. For example, over-investment results from a decision to adapt followed by no trend in storm surges, which has a conditional probability equal to $\alpha = $ P[no trend in storm surge|adaptation]. Correspondingly, under-preparation results from a decision not to adapt followed by a trend in storm surges, with a conditional probability equal to $\beta = $ P[trend in storm surges|no adaptation].

The adaptation costs and the damages corresponding to each outcome are also shown in Figure 5. The cost of regret is a net cost for each of the outcomes that would result from a type I or type II error. The cost of regret accounts for the costs of adaptation, the damages, and the damages avoided. Expected regret for each of the two undesirable outcomes is computed as the product of the cost of regret and its associated probability. Figure 6 describes the calculation of the expected regret associated with under- and over-preparation. Generally, one would recommend investment in adaptation when the expected regret of under-preparation is greater than the expected regret of over-investment, and to not adapt otherwise.

7. Risk-Based Decision Analysis in a Nonstationary World: Case Study, Mystic, CT

The expected regret decision tree method outlined in Figures 5 and 6 is applied to an adaptation decision for coastal protection in the village of Mystic in Groton, CT. The town of Groton, CT has been actively engaged in the task of climate change evaluation and preparedness. In January–June of 2010, three workshops were held engaging local, state, and federal government officials and various stakeholders; the workshops addressed climate change impacts and adaptation with a special focus on Mystic, CT. [see *Rosner*, 2012]. Only the most relevant aspects of adaptation plan are included here. The reader is referred to *Rosner* [2012] for further details on adaptation alternatives.

We have utilized both the selection of adaptation alternatives and the cost and damage estimates from these workshops to demonstrate the application of the approach outlined in Figures 5 and 6 for this adaptation decision process. Of the nine coastal adaptation alternatives identified in the workshops, *Rosner* [2012] found that if one conducted a traditional RBDM analysis, regardless of whether or not a trend exists, the net benefits associated with several alternatives were always positive, and in other cases were always negative. In those situations, it is straightforward to decide whether or not a particular alternative is attractive, as those decisions do not hinge on uncertain climate change impacts. However, for some adaptation options, such as Mystic alternative B in *Rosner* [2012], the net benefits were negative if sea level trends are ignored, and positive otherwise. Therefore, the decision of whether to recommend that adaptation option is dependent on our degree of certainty concerning the observed trend. This is the type of situation for

10.1002/2013WR014561

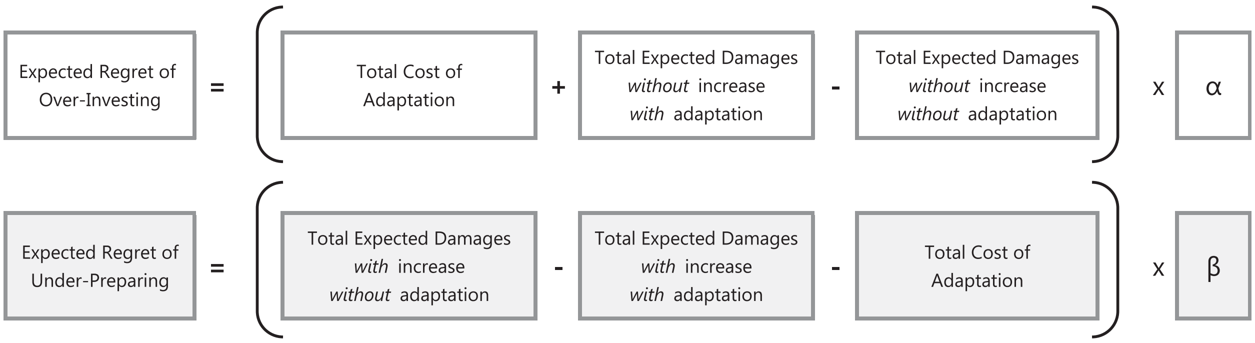

Figure 6. Computation of cost of regret.

which our approach is ideally suited, because our approach using expected regret integrates the uncertainty associated with our knowledge of future trends along with the costs and uncertainty associated with the various adaptation strategies. Thus, the following analysis focuses only on one of the nine adaptation strategies summarized by *Rosner* [2012, Mystic alternative B]. Alternative B in *Rosner* [2012] consists of a hurricane barrier at the entrance to the Mystic River which would provide protection to a level of 3.28 m (above the station datum) at a capital cost of $18 million and annual operations and maintenance cost of $75,000 per year. We have chosen to use a 25 year period ending in year 2035 for evaluating the project.

8. Trend Detection in Sea Level Anomalies

An annual series of sea level anomalies over the period 1938–2010 corresponding to the NOAA sea level gauge at New London, CT, was assembled using the methods outlined by *Kirshen et al.* [2008] and *Rosner* [2010]. Numerous parametric trend models were considered for describing the time trend in sea level anomalies. A log linear model was found to provide a good approximation to the sea level anomaly series because the resulting model residuals were heteroscedastic and well approximated by a normal distribution, thus enabling us to compute the type I and type II error (over- and under-preparation, respectively) probabilities given in equations (1) and (2). Further model diagnosis is provided in *Rosner* [2012], including influence statistics and an evaluation of the serial correlation of the model residuals. The fitted model $y=\ln(q)=b_o+b_1x=-3.840+0.00188x$ is shown in Figure 7, where q are the annual maximum sea level anomalies for year x over the $n=73$ year period.

The type I error probability is computed from equation (1) so that $\alpha=P[T_{n-2}\geq t]=0.115$, where $t=b_1/s_{b_1}=1.212$ is the t ratio for b_1 and $s_{b_1}=0.00155$ is the standard deviation of b_1. Thus, $\alpha=0.115$ is simply the probability that a Student's t variate, with 71 degrees of freedom, exceeds 1.212, and reflects the likelihood of a type I error. Similarly from equation (2) $\beta=P\left(T_{n-2}\leq\left(t_{1-\alpha,n-2}-\delta\sqrt{n}\right)\right)$ with $\delta=1/\sqrt{(1/r^2)-1}=0.144$, where $r=0.142$ is an estimate of the correlation coefficient ρ between y and x. Here $t_{1-\alpha,n-2}=t_{0.885,71}=1.212$ so that $\beta=P(T_{n-2}\leq 0.115)=0.493$ which reflects the likelihood of a type II error.

In this instance, the likelihood of both over- and under-preparation is quite high. Such a high value of $\alpha=1.115$ would normally result in a decision to ignore the possible trend in sea levels. However, the value of $\beta=0.493$ indicates a high likelihood of missing the trend if it really exists. This is a situation in which a complete risk-based trend-detection and decision analysis, as depicted in Figures 5 and 6, is needed to fully account for all possible costs and outcomes.

9. Stationary and Nonstationary Frequency Analysis

Previous studies examining the frequency distribution of sea levels have concluded that a Generalized Extreme Value (GEV) distribution can provide a good approximation to the frequency-magnitude relationship of storm surges [*Warner and Tissot*, 2012; *Huang et al.*, 2008; *Kirshen et al.*, 2008; *van den Brink et al.*, 2003]. The GEV model is fit to the annual maximum series (AMS) of sea level anomalies using the method of

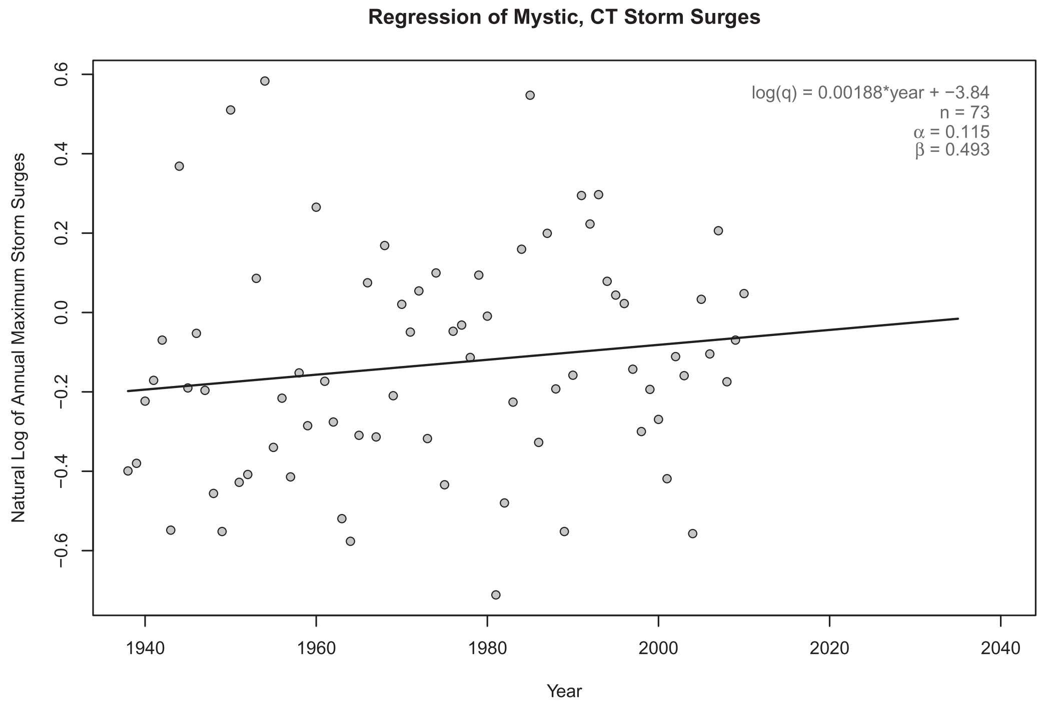

Figure 7. Relationship between sea level anomalies and time for the Mystic, CT case study along with fitted regression line and the likelihood of over-preparation ($\alpha = 0.115$) and under-preparation ($\beta = 0.493$).

moments, resulting in a stationary GEV model (see Figure 11 for a detailed outline of calculations). *Rosner* [2012] documents that the GEV distribution provides an excellent fit to the sea level data for this case study area. The trend model summarized in Figure 7 represents the mean sea level anomaly conditioned on the year of occurrence, thus a nonstationary GEV model is created by replacing the mean of the AMS series by the trend model in Figure 7 so that the resulting nonstationary GEV model becomes a function of the year.

10. Stationary and Nonstationary Damage-Frequency Analysis

The coastal flood damages at varying elevations for the Mystic site have been estimated for the Town of Groton, CT climate change preparedness study (P. H. Kirshen, Modeling Potential Adaptation Actions for Groton, 2010, http://www.icleiusa.org/action-center/planning/Modeling%20Adaptation%20Actions%20-for%20Groton.pdf), and are summarized in Figure 8. A simple exponential function was fit to the damages and storm surge elevation data, and used for interpolation only, to estimate damages for intermediate sized events. The damage elevation curve in Figure 8 is combined with each of the stationary and nonstationary GEV models of storm surge elevation frequency, resulting in the damage frequency relationships for the stationary and nonstationary cases, examples of which are seen in Figure 9. These damage frequency curves are developed over the planning horizon 2011–2035 for each case, with no adaptation or with a hurricane barrier in place. Note that while a single damage frequency curve is all that is needed in a stationary world, under nonstationarity this relationship changes slightly with each progressing year. Thus, there are different damage frequency curves for each year, with the curve for 2035 shown as an example in Figure 9. With each adaptation alternative and associated protection level considered by *Rosner* [2012], the threshold for damage is raised. Note that no damage is caused by events below 2.85 m without adaptation, or below 3.28 with the hurricane barrier in place, and that the exceedance probability (shown on the *x* axis) for these elevation thresholds differs slightly between the stationary and 2035 nonstationary curves.

The area under each damage-frequency curve in Figure 9 represents the Expected Annual Damages (EAD) [see *National Research Council*, 2000, for further details]. In a stationary world, a single estimate of EAD is all that is needed though here a different EAD is computed for each year under nonstationary conditions. The expected damages for the entire planning period is the summation of the expected damages in each year, after first converting each value to a present worth equivalent using the current U.S. Army Corps of

10.1002/2013WR014561

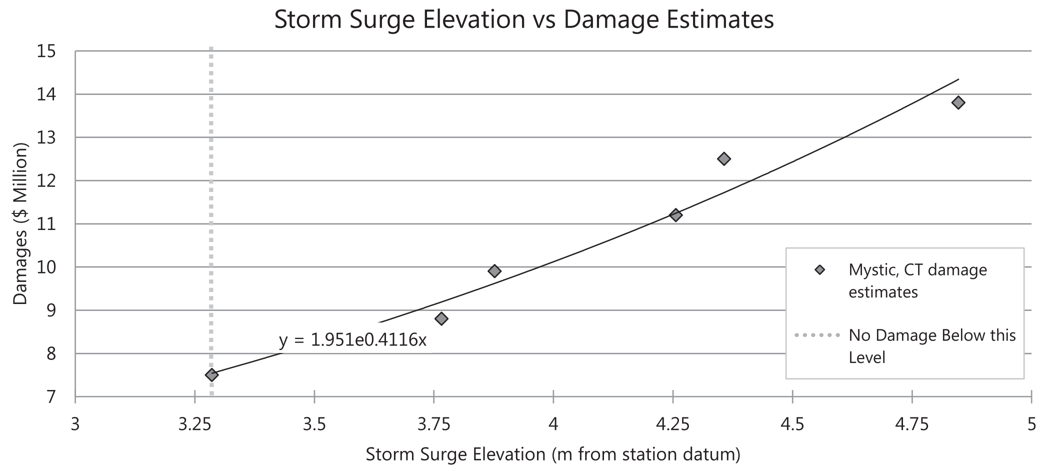

Figure 8. Mystic, CT flood damage estimates as a function of storm surge elevation along with fitted relationship used only for interpolation.

Engineers Project "Evaluation and Formulation Rate" of 3.5% recommended in *U.S. Army Corps of Engineers* [2013]. Annual operations and maintenance costs of the adaptation options are also converted to present worth dollars and added to the initial capital cost.

11. Risk-Based Trend Detection Results

In this section, we contrast the results of our Expected Regret decision-making approach based on risk and trend uncertainty, with the results of a traditional risk-based decision-making analysis which often ignores either the over- or under-preparation options. Figure 10 summarizes the results of our expected regret decision-making approach. It shows that the expected regret of over-preparation is $249,000, while the expected regret of under-preparation is close to three times greater at $685,000. This disparity results because the cost of regret of under-preparation is significantly higher than that of over-investment along

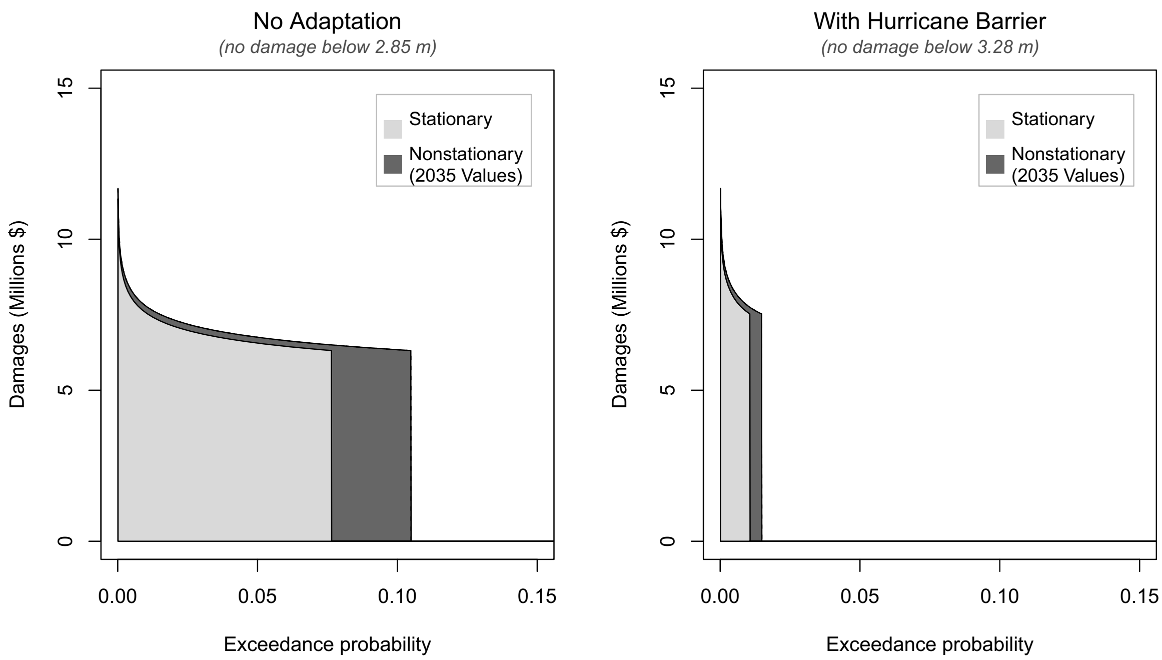

Figure 9. Exceedance probability versus expected damages for mystic, CT case study. Damages are estimated with and without proposed hurricane barrier, for stationary and nonstationary conditions. Area under each damage-frequency curve represents Expected Annual Damages (EAD). A different EAD is computed for each year under nonstationary conditions, and the curve for year 2035 is shown as an example.

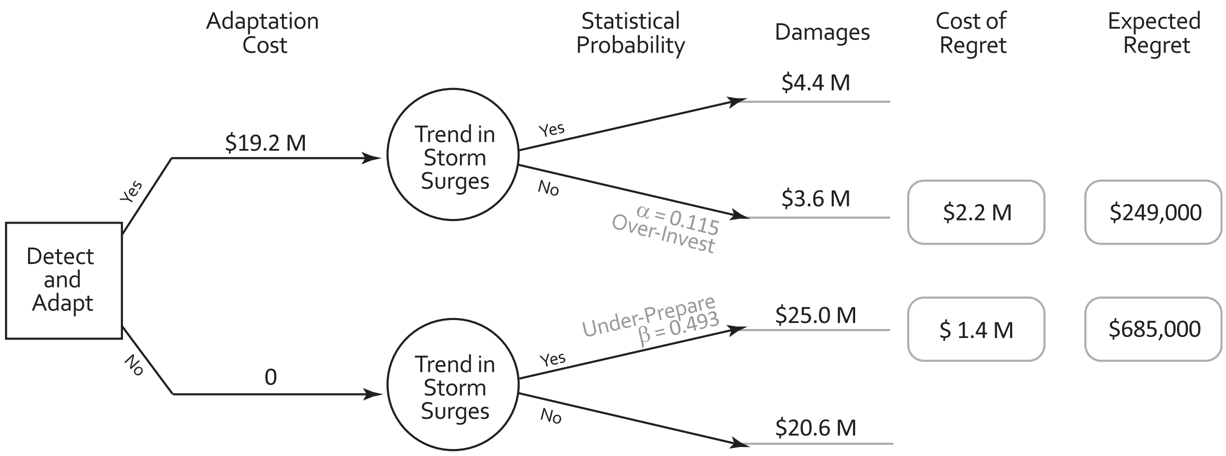

Figure 10. Decision tree for risk-based coastal flood adaptation in a nonstationary world, case study: Mystic, CT.

with the fact that the type II (under-preparation) error probability (0.493) is far greater than that of type I (over-investment) error (0.115). The result of our proposed approach would be a strong recommendation to adapt, despite the uncertainty associated with our current state of knowledge concerning potential increasing trends.

In contrast, using the traditional net benefits approach, one normally either first accepts or rejects the increasing trend in storm surge heights, and then considers the damages avoided less the cost of adaptation. Here we obtain expected net benefits under stationary and nonstationary conditions of −$2,175,000 and +$1,389,000, respectively. Recall that since probability of a type I error (over-investment) of the trend shown in Figure 7 was higher than 0.05, engineers would normally dismiss the possibility of nonstationary conditions under these conditions, and thus conclude not to build the hurricane barrier, since the net benefits are negative under stationary conditions.

12. Conclusions

We have introduced an approach that combines hypothesis testing, risk-based decision making, and a decision tree to address climate change adaptation decisions made under conditions of nonstationarity and uncertainty. Traditional risk-based approaches which attempt to deal with nonstationarity tend to separate the statistical hypothesis testing (trend detection) from the decision aspects of the analysis. They frequently fail to assess the risk of under-preparation resulting from a failure to detect the trend. Our integrated approach based on statistical decision theory addresses both uncertainties associated with the damages from future hazards (storm surges), and uncertainties due to our limited information and knowledge concerning future trends in those natural phenomena. We accomplish this by combining risk-based decision making to assess the economic consequences of possible adaptation strategies and their uncertain outcomes, and hypothesis testing to assess the degree of uncertainty associated with our knowledge about future trends.

Our combined risk-based decision and trend detection method gives needed attention to the risks and damages of under-preparation. The concept of expected regret integrates the probabilities of each under- and over-preparation along with their associated economic repercussions under all possible futures, and provides decision-makers with an objective and physically relevant method of addressing risk and uncertainty in the adaptation process. Instead of simply presenting the costs and benefits of the recommended adaptation plan, with a footnote about the uncertainty of the trend, this method integrates all of relevant information to address the question, “Should we invest now, despite all relevant sources of costs and uncertainty?” Our “systems” approach attempts to integrate all relevant information concerning adaptation decisions in an uncertain and possibly nonstationary world to enable decision-makers to weigh competing adaptation strategies. Central to our methodology is the assumption that all such information is available, including information concerning future flood damage costs which is very challenging to obtain.

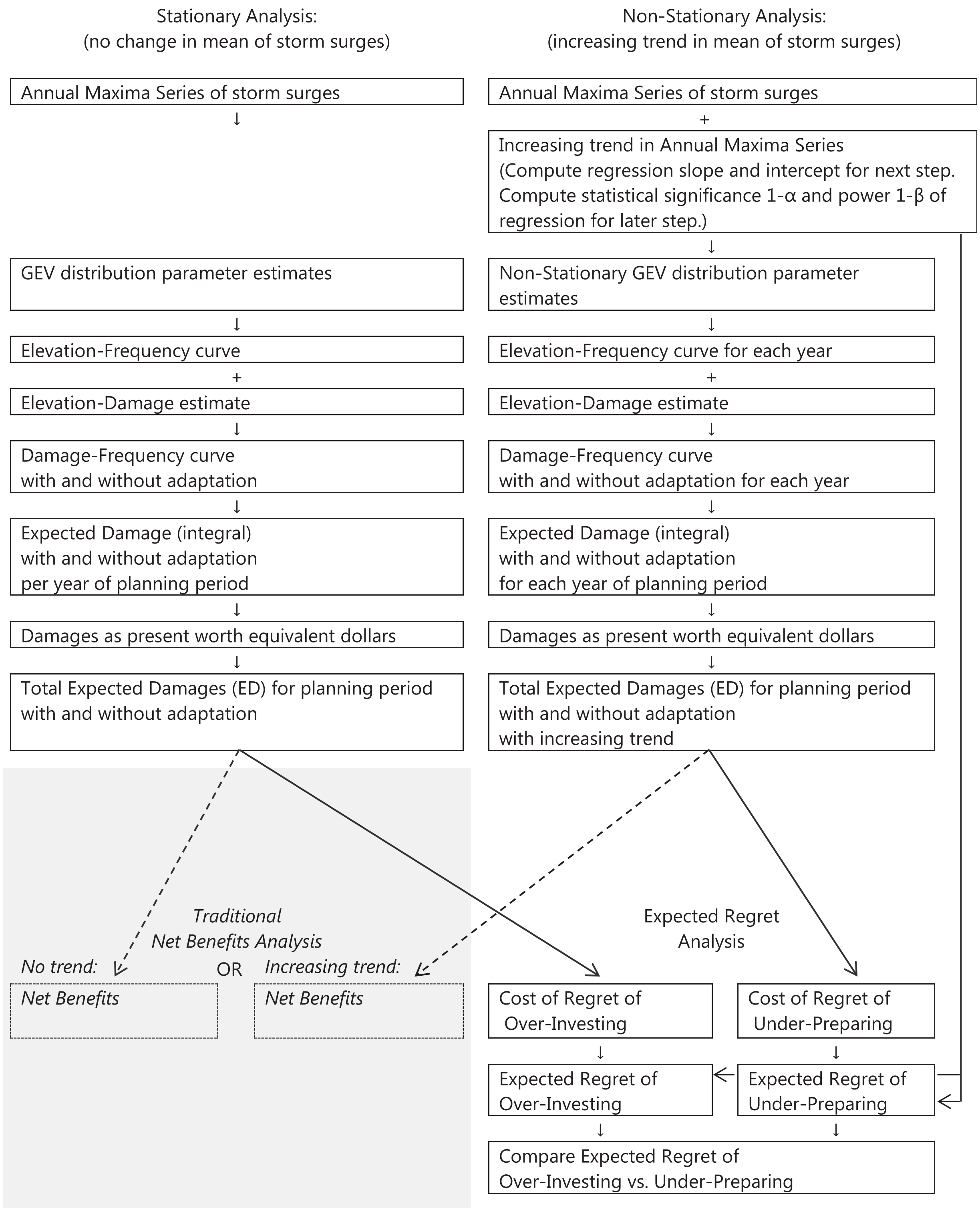

Figure 11. Steps for computation of expected regret for adaptation decisions.

Nevertheless, our proposed methodology should help decision-makers better understand the information requirements needed to make efficient and sensible flood management adaptation decisions.

The strength of this method is its ability to deal with situations in which a particular adaptation option is economically attractive under nonstationary conditions, but, when stationarity is assumed, the cost of the adaptation outweighs the damages avoided, and therefore appears not to be economically viable. These are precisely the adaptation situations in which a complete risk-based, decision-oriented approach is needed to aid decision-makers. Such situations consist of a combination of high uncertainty, substantial investment, and large risks associated with future damages. Our proposed method helps guide precisely these types of decisions, by bringing together the separate assessments of uncertainty and of damage costs in a meaningful and rational way.

By weighing the expected regret of over-investment against the expected regret associated with under-preparation, decision-makers can evaluate their various adaptation (investment) options. Our recommended approach based on expected regret led to dramatically different results than a traditional risk-based analysis based on net benefits for the case study considered. In our Mystic, CT case study, we showed that the traditional risk-based decision-making analysis using net benefits would dismiss the trend because the statistical hypothesis test yields $\alpha > 0.05$. Such an analysis would have led to a strong recommendation against building the barrier due to the negative net benefits under stationary conditions, and would fail to calculate the net benefits under nonstationary conditions. In our expected regret based analysis, which integrates the hypothesis testing with the economic analysis, the expected regret of over-preparation was only $249,000, while the expected regret of under-preparation is almost three times greater at $685,000. Thus, our approach concludes with a strong recommendation to adapt, despite the uncertainty associated with the increasing trend in future sea levels.

Our approach is predicated upon our ability to estimate the type I and type II error probabilities associated with a trend test. This initial study employed analytic expressions given in equations (1) and (2) based on linear regression. This is an attractive approach for numerous reasons outlined by *Vogel et al*. [2013] including the fact that it enables approximations of under- and over-preparation errors based on either observed or projected data and could also provide prediction errors associated with such trend extrapolations. Future work should consider nonparametric trend detection methods, such as the Mann-Kendall statistic. *Yue et al*. [2002], *Yue and Pilon* [2004], *Onoz and Bayazit* [2003], and *Morin* [2011] have examined the power of the Mann-Kendall test and other nonparametric techniques using Monte Carlo analysis. To our knowledge, no published study has provided analytical expressions for the power of the Mann-Kendall test, thus Monte Carlo estimates are probably necessary. It is interesting to note, however, that *Morin* [2011] found that for his particular application, the results of the linear regression and the Mann-Kendall were nearly identical.

The trend detection analysis outlined here is overly simplistic and would benefit from further analyses which attempt to distinguish future trends from natural climatic persistence, cycles or other "natural" climate behavior. For example, *Cohn and Lins* [2005], *Jain and Lall* [2001], and others have raised numerous concerns regarding our ability to detect trends in natural time series, and especially our inability to distinguish trends from the natural stochastic persistence, periodicities, and change points inherent in most hydro climatic processes. It is our hope that recent innovations in trend detection summarized by *Helsel and Hirsch* [2002], *Khalid et al*. [2009], *Kropp and Schellnhuber* [2011], *Sonali and Kumar* [2013], and others which account for numerous such complications associated with detection of trends will be integrated into the type of decision-oriented trend analyses introduced here. This study along with the study by *Vogel et al*. [2013] emphasize the need for further work to more fully evaluate our ability to estimate the type I and type II error probabilities associated with the various hypothesis tests which are commonly used in hydroclimatic investigations.

Acknowledgments
The second author is indebted to the U.S. Army Corps of Engineers, Institute for Water Resources (IWR), for their support of this research project. The first and third authors were funded under a grant from the Sectoral Applications Research Program (SARP) of the National Oceanic and Atmospheric Administration (NOAA) Climate Program Office, grant NA07OAR4310373. The views expressed represent those of these authors and do not necessarily reflect the views or policies of either IWR or NOAA. Research quality hourly sea level data for New London, CT, based on NOAA observations, was acquired from the Joint Archive for Sea Level (http://uhslc.soest.hawaii.edu/data/download/rq). The authors would like to thank to Elena Naumova for her valuable comments.

References

ASCE (2007), *The New Orleans Hurricane Protection System: What Went Wrong And Why*, Am. Soc. of Civ. Eng., Reston, Va.

Bowling, L. C., P. Storck, and D. P. Lettenmaier (2000), Hydrologic effects of logging in western Washington, United States, *Water Resour. Res.*, *36*(11), 3223–3240.

Chao, P. T., and B. F. Hobbs (1997), Decision analysis of shoreline protection under climate change uncertainty, *Water Resour. Res.*, *33*(4), 817–829.

Cohen, J. (1994), The earth is round ($p < .05$), *Am. Psychol.*, *49*(12), 997–1003.

Cohn, T. A., and H. F. Lins (2005), Nature's style: Naturally trendy, *Geophys. Res. Lett.*, *32*, L23402, doi:10.1029/2005GL024476.

Dupont, W. D., and W. D. Plummer Jr. (1990), Power and sample size calculations: A review and computer program, *Control. Clin. Trials*, *11*, 116–128.

Dupont, W. D., and W. D. Plummer (1998), Power and sample size calculations for studies involving linear regression, *Control. Clin. Trials*, *19*, 589–601.

Easterling, D. R., J. L. Evans, P. Y. Groisman, T. R. Karl, K. E. Kunkel, and P. Ambenje (2000), Observed variability and trends in extreme climate events: A brief review, *Bull. Am. Meteorol. Soc.*, *81*(3), 417–425.

Fiering, M. B., and N. C. Matalas (1990), Decision-making under uncertainty, in *Climate Change and U.S. Water Resources*, edited by P. E. Waggoner, pp., 75–84, John Wiley, New York.

Helsel, D. R., and R. M. Hirsch (2002), *Statistical Methods in Water Resources*, U.S. Geol. Surv. Tech. Water Resour. Invest., Book 4, Chap. A3, 510 pp.

Hobbs, B. F., P. T. Chao, and B. N. Venkatesh (1997), Using decision analysis to include climate change in water resources decision making, *Clim. Change*, *37*, 177–202.

Huang, W., S. Xu, and S. Nnaji (2008), Evaluation of GEV model for frequency analysis of annual maximum water levels in the coast of United States, *Ocean Eng.*, *35*(11–12), 1132–1147.

Huntington, T. G. (2006), Evidence for intensification of the global water cycle: Review and synthesis, *J. Hydrol.*, *319*(1–4), 83–95.

Jain, S., and U. Lall (2001), Floods in a changing climate: Does the past represent the future?, *Water Resour. Res.*, *37*, 3193–3205.

Khaliq, M. N., T. B. M. J. Ouarda, P. Gachon, L. Sushama, and A. St-Hilaire, (2009), Identification of hydrological trends in the presence of serial and cross correlations: A review of selected methods and their application to annual flow regimes of Canadian rivers, *J. Hydrol.*, *368*, 117–130.

Kirshen, P. H., C. Watson, and E. Douglas (2008), Coastal flooding in the Northeastern United States due to climate change, *Mitigation Adapt. Strategies Global Change*, *13*, 437–451.

Kropp, J. P., and H. J. Schellnhuber (2011), *In Extremis: Disruptive Events and Trends in Climate and Hydrology*, 400 pp., Springer, Berlin.

Kwon, H. H., C. Brown, and U. Lall (2008), Climate informed flood frequency analysis and prediction in Montana using Hierarchical Bayesian Modeling, *Geophys. Res. Lett.*, *35*, L05404, doi:10.1029/2007/GL032220.

Lettenmaier, D. P. (1976), Detection of trends in water quality data from records with dependent observations, *Water Resour. Res.*, *12*, 1037–1046.

Manning, L. J., J. W. Hall, H. J. Fowler, C. G. Kilsby, and C. Tebaldi (2009), Using probabilistic climate change information from a multimodel ensemble for water resources assessment, *Water Resour. Res.*, *45*, W11411, doi:10.1029/2007WR006674.

Matalas, N. C. (2012), Comment on the announced death of stationarity, *J. Water Resour. Plann. Manage.*, *138*, 311–312.

Milly, P. C. D., et al. (2008), Stationarity is dead: Whither water management?, *Science*, *319*, 573–574.

Morin, E. (2011), To know what we cannot know: Global mapping of minimal detectable absolute trends in annual precipitation, *Water Resour. Res.*, *47*, W07505, doi:10.1029/2010WR009798.

Mosteller, F., and J. W. Tukey (1977), *Data Analysis and Regression*, 588 pp., Addison-Wesley, Menlo Park, Calif.

National Research Council (2000), *Risk Analysis and Uncertainty in Flood Damage Reduction Studies*, Natl. Acad. Press, Washington, D. C.

Nicholls, N. (2000), The insignificance of significance testing, *Bull. Am. Meteorol. Soc.*, *81*(5), 981–986.

Onoz, B., and M. Bayazit (2003), The power of statistical tests for trend detection, *Turk. J. Environ. Eng. Sci.*, *27*, 247–251.

Rosner, A. (2012), Risk based trend detection for climate change adaptation, MS thesis, Tufts Univ., Medford, Mass.

Rothenberg, T. J. (1988), Approximate power functions for some robust test of regression coefficients, *Econometrica*, *56*(5), 997–1019.

Sarewitz, D., R. Pielke Jr., and M. Keykhah (2003), Vulnerability and risk: Some thoughts from a political and policy perspective, *Risk Anal.*, *23*(4), 805–810.

Sonali, P., and D. N. Kumar (2013), Review of trend detection methods and their application to detect temperature changes in India, *J. Hydrol.*, *476*, 212–227.

Stedinger, J. R. (1982), Design events with specified flood risk, *Water Resour. Res.*, *19*(2), 511–522.

Trenberth, K. E. (2011), Attribution of climate variations and trends to human influences and natural variability, *WIREs Clim. Change*, *2*, 925–930.

Tung, Y.-K. (2005), Flood defense systems design by risk-based approaches, *Water Int.*, *30*(1), 50–57.

U.S. Army Corps of Engineers (2000), Planning Guidance Statement, *ER 1105-2-100*, Dep. of the Army, Washington, D. C.

U.S. Army Corps of Engineers (2013), Federal Interest Rates for Corps of Engineers Projects for Fiscal Year 2014, *Econ. Guidance Memo. 14-01*, Dep. of the Army, Washington, D. C.

van den Brink, H. W., G. P. Konnen, and J. D. Opsteegh (2003), The reliability of extreme surge levels estimated from observational records of order hundred years, *J. Coastal Res.*, *19*(2), 376–388.

Villarini, G., J. A. Smith, F. Serinaldi, J. Bales, P. D. Bates, and W. F. Krajewski (2009), Flood frequency analysis for nonstationary annual peak records in an urban basin, *Adv. Water Resour.*, *32*, 1255–1266.

Vogel, R. M., C. Yaindl, and M. Walter (2011), Nonstationarity: Flood magnification and recurrence reduction factors in the United States, *J. Am. Water Resour. Assoc.*, *47*(3), 464–474.

Vogel, R. M., A. Rosner, and P. H. Kirshen (2013), Likelihood of societal preparedness for global change—Trend detection, *Nat. Hazards Earth Syst. Sci.*, *13*, 1–6.

Warner, N. N., and P. E. Tissot (2012), Storm flooding sensitivity to sea level rise for Galveston Bay, Texas, *Ocean Eng.*, *40*, 23–32.

Yue, S., P. Pilon, and G. Cavadias (2002), Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series, *J. Hydrol.*, *259*, 254–271.

Ziegler, A. D., J. Sheffield, E. P. Maurer, B. Nijssen, E. F. Wood, and D. P. Lettenmaier (2003), Detection of intensification in global- and continental-scale hydrological cycles: Temporal scale of evaluation, *J. Clim.*, *16*(3), 535–547.

Ziegler, A. D., E. P. Maurer, J. Sheffield, B. Nijssen, E. F. Wood and D. P. Lettenmaier (2005), Detection time for plausible changes in annual precipitation, evapotranspiration, and streamflow in three Mississippi river sub-basins, *Clim. Change*, *72*, 17–36, doi:10.1007/s10584-005-5379-4.

Ziliak, S. T., and D. N. McCloskey (2008) *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives*, Univ. of Michigan Press, Ann Arbor, Mich.

