Fondazione Eni Enrico Mattei

An Information-Theoretical Analysis of Budget-Constrained Nonpoint Source Pollution Control

Jonathan D. Kaplan, Richard E. Howitt and
Y. Hossein Farzin

NOTA DI LAVORO 31.2001

Corso Magenta, 63, 20123 Milano, tel. +39/02/52036934 – fax +39/02/52036946
E-mail: letter@feem.it
C.F. 97080600154

An Information-Theoretical Analysis of
Budget-Constrained Nonpoint Source Pollution Control*

Jonathan D. Kaplan**
Resource Economics Division, Economic Research Service, United States Department of Agriculture, 1800 M Street NW, Washington, DC 20036

Richard E. Howitt
Department of Agricultural and Resource Economics, University of California, Davis, CA 95616

Y. Hossein Farzin
Department of Agricultural and Resource Economics, University of California, Davis, CA 95616

Abstract
This paper analyzes budget-constrained, nonpoint source (NPS) pollution control with costly information acquisition and learning. To overcome the inherent ill-posed statistical problem in NPS pollution data the sequential entropy filter is applied to the sediment load management program for Redwood Creek, which flows through Redwood National Park in northwestern California. We simulate the dynamic budget-constrained management model with information acquisition and learning, and compare the results with those from the current policy. The analysis shows that the manager can reallocate resources from treatment effort to information acquisition, which in turn increases overall treatment effectiveness, and reduces sediment-related damage.

*Financial support for this research was provide in part by a grant (Project No. W-887) from the University of California Water Resources Center. The first author was also supported by the United States Environmental Protection Agency, "Science to Achieve Results," Graduate Fellowship Program. Farzin thanks the Giannini Foundation for partial support of this research. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Economic Research Service or the US Department of Agriculture.

**To whom correspondence should be addressed. E-mail: jkaplan@ers.usda.gov.

An Information-Theoretical Analysis of Budget-Constrained Nonpoint Source Pollution Control

I. INTRODUCTION

Nonpoint source (NPS) pollution control is primarily an information, or conversely, an uncertainty problem, and as such, modeling NPS pollution control requires that the role of information and learning be explicitly specified and addressed in the model. In general, pollution control problems can be characterized by the degree of uncertainty or incomplete information about the location of polluting sources and the magnitude of their contribution. If there were complete certainty or perfect information on the location and magnitude of pollution for each source, then the problem would be a "pure" point source (PS) problem. If, however, the manager in charge of pollution control does not know the location of the sources or has no knowledge of the contribution of each source in the aggregate pollution, then the problem would be a "pure" NPS problem.[1]

These two extremes of information mark the ends of a spectrum that defines all pollution problems by the degree of uncertainty about the location and pollution generation for each source rather than the vague classification of either PS or NPS. Note that the extreme NPS problem when the location is unknown is not susceptible to direct policy controls because the location of the pollution sources needs to be known to implement treatment policies. That is, implementation of pollution control policies requires that the location of the sources is known with a high degree of confidence, or information can easily be acquired about the identity of the sources so that treatment efforts can be undertaken.

Another complication to NPS control is the budgetary restrictions that often limit the extent of NPS pollution control. These financial limitations have implications for pollution control in general and in the United States particularly, where the federal government spends $3 billion annually to control NPS water pollution, yet reports that the public control of NPS pollution was limited by a lack of financial resources [33]. The United States government is perhaps the largest controller of NPS pollution and the largest contributor of NPS pollution, contributing nearly half of all the NPS pollution in the 11 western states from (Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming) [33].

The analysis of NPS pollution control is further complicated by the statistical nature of NPS pollution data. In any given period, there are more polluting sources than observations (i.e., an undersized sample). This statistical problem of estimating with an undersized sample, sometimes referred to as an ill-posed data problem, makes traditional statistical approaches inappropriate and requires that a new approach be employed. Although the policy maker could wait for a sufficiently long time series to be collected, thereby avoiding the ill-posed estimation problem, interim damage will occur that result in irreversible losses, such as the destruction of critical habitat for endangered or protected resources. If the information contained in the data is incorporated as it is collected the manager can take concurrent action that reduce interim damage rather than wait until sufficient data is collected and face the risk of incurring irreversible losses.[2] In the empirical application to follow we adopt a sequential entropy filter (SEF) approach, which generates parameter estimates from ill-posed data and updates the estimates, as new data becomes available [18].

As an application of the theoretical and methodological aspects described above, we consider the current sediment control program for Redwood Creek, which flows into and through Redwood National Park, located in Orick, California. If perfect information were available on the location and sediment loading contribution of the pollution sources, park managers could allocate the entire sediment control budget to treatment effort. However, with incomplete information, the management of sediment loading requires an explicit or implicit allocation of sediment control resources between information acquisition and treatment. We focus on the tradeoff between information acquisition and treatment by explicitly modeling information acquisition and learning within a budget-constrained, pollution control model.

In the analysis two data collection regimes are constructed to facilitate the estimation of the unobservable sediment loading parameters. These estimates are then incorporated into the sediment control model. We simulate the model to derive optimal policy decisions regarding treatment and information acquisition (learning) strategies and provide a comparison of the current treatment policy with two alternative site-specific NPS pollution treatment policies. This comparison allows us to test the hypotheses that a flexible treatment strategy, which allows for site specific treatment effort, will lower sediment related damages if the manager can acquire information and shift resources toward dirtier pollution sources. The results drawn from this analysis can, in general, shed light on the role of information and budgetary constraints on the efficient management of pollution problems elsewhere.

This paper extends the literature by developing a dynamic model of NPS pollution control that explicitly characterizes the uncertainty about the pollution loading and the budget constraint that limits the ability of managers to minimize damages. Also, this paper contributes to the literature by introducing a new statistical approach well suited for empirical analysis of NPS pollution problems and dynamic models in general. The application of the sequential entropy filter in this paper is, to our knowledge, the first use of the entropy approach to sequential update parameters for a dynamic NPS pollution control model. Finally, this paper provides model simulation and policy analysis that allows us to test the hypothesis that flexible treatment strategies with information acquisition and learning can reduce the pollution related damages even though resources used to collection data are no longer available to treat the polluting sites.

In the next section we provide past research related to our paper. Section III introduces the entropy formalism. Section IV develops the budget-constrained pollution control model under pollution generation uncertainty. Section V describes the empirical application of the SEF to the sediment control program for Redwood Creek and the estimation results. The simulated model and policy analysis are presented in Section VI. Section VII concludes.

II. RELATED LITERATURE

The management of pollution has often been cast as a non-cooperative, asymmetric game because the private polluter has better information on the cost of abatement than the pollution control manager [8, 14, 15, 24, 25]. Recently, Amacher and Malik [1] introduced the pollution control problem as a cooperative game in light of the bargaining potential observed in actual management strategies. The manager reduces asymmetric information by bargaining the adoption of a cleaner technology for less compliance and ultimately lower social costs. In the pollution control model presented here, the manager does not rely on bargaining but rather on strategic data collection to reduce uncertainty. This does not preclude a bargaining approach, but rather, emphasizes that budget constraints should be optimized prior to bargaining solutions. When polluting activities have ceased, as is the case in Redwood National Park, there is no asymmetric information. The park managers know the cost of abatement but are not fully informed on each sources' contribution to the pollution load.

The research on information acquisition in pollution control dates back to Rausser and Howitt [23], who examine information acquisition in regulatory mechanisms. More recently, Garvie and Keeler [8] consider the fiscally constrained regulator who is responsible for minimizing the pollution level in excess of a given minimum standard by allocating a given budget between data collection and enforcement. The regulator relies on data collection as evidence necessary to enforce the minimum standard. In other words, the data collection improves the enforcement effectiveness. The problem addressed in our paper is similar to the Garvie and Keeler regulator problem where the budget constrained NPS manager collects data in order to improve treatment effectiveness.

In another related paper, Xepapedeas [34] examines the use of an effluent tax, in conjunction with an ambient tax, as an incentive for individual polluters to reveal information about the connection between generator and ambient emissions. Xepapedeas' problem is that of a social planner who is not constrained by a limited budget. In the case of sediment loading, which we examine, the budget-constrained manager cannot impose fees against nature to learn more about each polluting source's contribution to the total ambient load. The manager must spend limited resources to obtain the information that otherwise could have been obtained by the social planner through the use of effluent fees.

Likewise, Cabe and Herriges [3] consider the NPS pollution control problem in a social welfare framework, incorporating both uncertainty and information acquisition. In this setting acquired information reduces the social cost of setting a control mechanism through an ambient tax, as proposed by Segerson [26]. However, the behavioral objective of social welfare maximization will lead to a different outcome and different policy prescription than under damage minimization subject to budget constraints [2]. Farzin and Kaplan [5] explicitly consider the budget-constrained public manager who acquires information in a static setting to improve treatment effectiveness and reduce pollution damages below the levels under uninformed management. They focus their attention on the theoretical underpinnings of the management decision making process but do not explore the dynamic nature of pollution control nor do they address the statistical problem confronting empirical analysis of NPS pollution control.

Empirical analysis of pollution control has been sparse. LaPointe and Rilstone [20] and Magat and Viscusi [21] analyze the effect of inspections on pulp and paper mill emissions. Gray and Deily [11] extend this work to consider management behavior in a similar emissions problem. These studies fail to address the economic trade-off between information acquisition and pollution treatment in a budget-constrained environment and when learning takes place over time. Fernandez [7] employs an entropy approach to estimate the parameters for state and observation equations in wastewater treatment facilities. However, her approach does not employ the sequential updating potential of the entropy formalism.

III. THE ENTROPY FORMALISM

The use of information entropy in economic theory and econometric application is a recent phenomenon. Shannon [27] first introduced the concept of information entropy as a measure of uncertainty. Khinchin [19] formalized the mathematical foundations for the information-theoretical approach to uncertainty. The basic idea underlying the entropy approach is that the entropy of the uniform distribution (i.e., where any realization is as likely as another) is the maximum or upper bound on the information entropy metric. This represents complete uncertainty for a given random parameter and the case where there is no information except the upper and lower bounds of the uniform distribution. Conversely, a spiked distribution (i.e., where only one realization can occur) has an entropy of zero. This lower bound corresponds to the case of certainty or complete information. Thus, normalizing the entropy measure by the maximum entropy bounds the uncertainty between zero and one and provides a bounded ordinal ranking of uncertainty. We can, in principle, model the entire spectrum of possible pollution control problems by incorporating this interpretation of information entropy into the management model.

There is a second application of the information entropy as well. That is, the estimation of parameter distributions. Jaynes [16] shows that the entropy criterion maximizes the multiplicity in the sense that the selected distribution is the one that can be realized in the greatest number of ways, and consistent with what is known (i.e., the data). The particular problem we face is estimating sediment-loading parameters for more sources than there are total ambient sediment measures. This is similar to the problem of reconstructing images from sparse data where maximum entropy methods have been used in the past [12, 13,29].

There are several advantages in using the entropy approach over least squares or maximum likelihood estimation that make it ideal for estimating NPS pollution parameters, where the underlying distributions for the natural system is unknown and the pollution generation is from diffuse sources. First, the error distribution does not have to be specified in the entropy approach. It is, however, estimated from the data. Second, given that in any given period there are more pollution source parameters to estimate than observations, the problem of inverting the matrix of explanatory variables is avoided with the entropy approach. This frees the researcher to make inferences about the model parameters without the degree of freedom restrictions or knowledge of the underlying distribution that are necessary to employ least squares or maximum likelihood estimation. In addition, the SEF can be employed to empirically incorporate learning (Bayesian updating) into the analysis, and update parameter distributions, as new data become available. The formal modeling of Bayesian learning by integrating over joint density functions is often very cumbersome [4]. Fortunately, The SEF is equivalent to Bayes' theorem [9,18,35] and overcomes the burden of integrating over multiple dimensions. Given that the SEF is easier to empirically model and has desirable consistency properties, we employ it in estimating the unknown sediment loading parameters.

Previous empirical application of the entropy metric in state space modeling, which is used in the sequential entropy filter approach, has been limited. Golan, Judge and Karp [10], in the context of a dynamic discrete time model, present the generalized maximum entropy (GME) approach for estimating unobservable state space parameters. Similarly, Fernandez [7], as mentioned earlier, uses GME to estimate the state and observation parameters for a dynamic model. Neither of these studies considers the sequential updating potential of the entropy approach in dynamic models.

IV. THE SEDIMENT CONTROL MODEL

In the sediment control model, we assume the manager knows the location and size of the sources but there is uncertainty (incomplete information) about the generation of pollution from each source.[3] The pollution control manager decides how to optimally allocate a limited budget between treatment effort and information acquisition. Treatment directly reduces the pollution generation. However, the manager can improve the treatment effectiveness by acquiring information that reduces the pollution generation uncertainty. If information is sequentially acquired (i.e., learning), then the degree of uncertainty declines with each sequential update. In some situations, a NPS problem may evolve to a PS one given that sufficient information is acquired.

This new model generalizes the problem to any type of pollution. In the model, information about the pollution generation is acquired through a single channel, namely data collection. In the context of sediment loading, data collection is often limited to stream flow for dispersed sources and an aggregate sediment load measure for the main watercourse. If there is significant variability in and between the stream flow from the sediment loading sources, then changes in stream flow can be compared with changes in aggregate sediment load to estimate the unobservable pollution generated by each source [17].

To construct the model let $s_i^n(t)$ be a sediment loading state and $p_i^n(t)$ be the probability that $s^n(t) = s_i^n(t)$, where *i* denotes the state, *n* denotes the source and *t* denotes time. We set *i* = 2 to simplify the presentation and write the probabilities as $p_1^n(t)$ and $p_2^n(t) = 1 - p_1^n(t)$, which allows us to reduce the number of state equations from *n*i* to *n*. The controls are $a(t)$ and $R(t) = (R^1(t), R^2(t), \cdots, R^N(t))$, the data collection intensity and level of treatment effort at each of the *n* sources, respectively. The data collection intensity provides information that allows the manager to update the probabilities $p_1^n(t)$ and $p_2^n(t)$. Treatment effort at any of the diffuse sources decreases the sediment loading states for that source. The state equations of motion for the probabilities are

$$\dot{p}_1^n(t) = g^n(a(t); p_1^n(t)), \forall n, \ \partial g / \partial a \text{ free} \qquad (2)$$

The expression $\partial g / \partial a$ maybe positive or negative and will depend on whether the posterior probability for the given state of nature is greater than or less than previously believed.

The sediment loading state equations of motion are

$$\dot{s}_i^n(t) = -b^n R^n s_i^n, \forall i, \forall n \qquad (3)$$

Without loss of generality, we assume a linear specification for the sediment loading state equations of motion, where $-b^n$ is the marginal decline in the sediment loading state for a given treatment effort. To reduce the number of state equations, we sum the n sediment loading state equations over *i* and obtain

$$\dot{s}_1^n(t) + \dot{s}_2^n(t) = -b^n R^n [s_1^n(t) + s_2^n(t)], \forall n \qquad (4)$$

Using the post-data probabilities and post-treatment sediment loading states we can construct the expected sediment loading from each source

$$E(s^n(t)) = \sum_i s_i^n(t) p_i^n(t) \qquad (5)$$

The sediment loading uncertainty about the known sources at a given time is measured with the normalized information entropy metric over all sediment loading distributions or formally

$$I(p(t)) = -(1/MaxEnt) \sum_n \sum_i p_i^n(t) \ln p_i^n(t), \qquad (1)$$

where $MaxEnt$ is the maximum entropy associated with the uniform distribution over the *p*'s.

In every time period the public manager is constrained by a budget $B(t)$, which is allocated between $a(t)$ and $R(t)$. We simplify the cost functions to focus our attention on the tradeoff between information acquisition and treatment effort. We assume constant unit costs of data collection, *m*, and treatment, *c*, and that treatment costs are identical across sites. Thus, the budget constraint is

$$B(t) = c \sum_n R^n(t) + a(t) m \qquad (6)$$

where *c* is the per unit treatment cost and *m* is the per unit data collection cost.

The optimal control problem facing the pollution manager is to choose the optimal paths of data collection intensity and treatment effort so as to minimize the discounted present value of damages from expected sediment loading at known sites. The time horizon for this problem is assumed to be infinite since there is no mandatory future date by which the control problem is expected to cease. Therefore, assuming a discount rate of *r*, the optimization problem is,

$$\underset{a(t),R(t)}{Min} \int_0^\infty e^{-rt} ED[S(t), P(t), a(t), R(t)]dt \tag{7}$$

subject to

$$\dot{p}_1^n(t) = g^n(p_1^n(t), a(t)), \forall n, \; p_1^n(t) = p_{1,0}^n, \forall n, \; p_2^n(t) = 1 - p_{1,0}^n, \forall n$$

$$\lim_{t\to\infty} p_1^n(t) = \overline{p}_1^n, \forall n, \; \lim_{t\to\infty} p_2^n(t) = 1 - \overline{p}_1^n, \forall n, \tag{7a}$$

$$\dot{s}_1^n(t) + \dot{s}_2^n(t) = -b^n R^n(s_1^n + s_2^n), \forall n, \; s_i^n(t) = s_{i,0}^n, \forall i, \forall n, \; \lim_{t\to\infty} s_i^n(t) = \overline{s}_i^n, \forall i, \forall n, \tag{7b}$$

$$E(s^n(t)) = \sum_i s_i^n(t) p_i^n(t), \tag{7c}$$

$$B(t) = c\sum_n R_n(t) + m a(t). \tag{7d}$$

Note, the initial probability $p_{1,0}^n$, is a subjective prior probability on the n sediment loading state. The expected damage function $ED(\cdot)$ is assumed to be twice continuously differentiable and convex in sediment loading.

To solve this optimal control problem, we first construct the Hamiltonian,

$$\begin{aligned} H &= e^{-rt} ED(S(t), P(t), a(t), R(t)) \\ &\quad + \sum_n l_p^n g^n(p_1^n(t), a(t)) - \sum_n l_s^n b^n R^n(t)[s_1^n(t) + s_2^n(t)] \end{aligned} \tag{8}$$

where l_p^n and l_s^n are the costate variables for the probabilities and sediment loading states at the n sources, respectively. Now we form the Lagrangian to incorporate the budget constraint, such that

$$L = e^{-rt} ED(S(t), P(t), a(t), R(t)) + \sum_n l_p^n g^n(p_1^n(t), a(t)) \\ - \sum_n l_s^n b^n R^n(t)[s_1^n(t) + s_2^n(t)] + m\left[c\sum_n R^n(t) + a(t)m - B(t)\right] \tag{9}$$

The necessary conditions to maximize the Lagrangian are (we suppress t as the argument of functions whenever no confusion arises)

$$\frac{\partial L}{\partial a} = e^{-rt}\frac{\partial ED}{\partial a} + \sum_n l_p^n \frac{\partial g^n}{\partial a} + mm = 0, \tag{10a}$$

$$\frac{\partial L}{\partial R^n} = e^{-rt}\frac{\partial ED}{\partial R^n} - l_s^n b^n\left[s_1^n(t) + s_2^n(t)\right] + mc = 0, \forall n, \tag{10b}$$

$$\dot{l}_p^n = -e^{-rt}\frac{\partial ED}{\partial p_1^n} - l_p^n \frac{\partial g^n}{\partial p_1^n}, \forall n,\ \lim_{t\to\infty} l_p^n(t) = \overline{l}_p^n, \forall n, \tag{10c}$$

$$\dot{l}_s^n = -e^{-rt}\frac{\partial ED}{\partial[s_1^n(t) + s_2^n(t)]} + l_s^n b^n R^n(t), \forall n,\ \lim_{t\to\infty} l_s^n(t) = \overline{l}_s^n, \forall n, \tag{10d}$$

$$\dot{p}_1^n(t) = g^n(p_1^n(t), a(t)), \forall n,\ p_1^n(t) = p_{1,0}^n, \forall n,\ p_2^n(t) = 1 - p_{1,0}^n, \forall n, \tag{10e}$$

$$\lim_{t\to\infty} p_1^n(t) = \overline{p}_1^n, \forall n,\ \lim_{t\to\infty} p_2^n(t) = 1 - \overline{p}_1^n, \forall n, \tag{10f}$$

$$\dot{s}_1^n(t) + \dot{s}_2^n(t) = -b^n R^n[s_1^n(t) + s_2^n(t)], \forall n,\ s_i^n(t) = s_{i,0}^n, \forall i, \forall n,$$

$$\lim_{t\to\infty} s_i^n(t) = \overline{s}_i^n, \forall i, \forall n, \tag{10g}$$

$$B(t) = c\sum_n R_n(t) + a(t)m,\ m \geq 0. \tag{10h}$$

The control problem does not allow us to solve for an analytical solution; however, we can still derive some qualitative properties from the necessary conditions. First, equation (10a) and (10b) provide the optimal conditions for choosing data collection intensities and treatment effort. These optimal conditions state the familiar story of selecting inputs so that the expected present value of marginal benefits from employing an input equals it's marginal cost. Farzin and Kaplan [5] and others noted that information acquisition is a collective good. Similarly, equation (10a) defines the optimal data collection intensity as the intensity where the expected value of acquired information (i.e., the increased treatment effectiveness in sediment control over all sources) is equal to the marginal opportunity cost of acquiring information. There is another interesting result obtained from (10a) and (10b) as well. We solve (10a) and (10b) for the shadow value on the budget constraint (m) and then equating these expressions for m yields

$$
\begin{aligned}
1/m\left(e^{-rt}\frac{\partial ED}{\partial a}+\sum_{n} l_p^n\frac{\partial g^n}{\partial a}\right) &= 1/c\left(e^{-rt}\frac{\partial ED}{\partial R^1}-l_s^1 b^1\left[s_1^1(t)+s_2^1(t)\right]\right) \\
&= \cdots = 1/c\left(e^{-rt}\frac{\partial ED}{\partial R^N}-l_s^N b^N\left[s_1^N(t)+s_2^N(t)\right]\right)
\end{aligned} \qquad (11)
$$

The economic interpretation of equation (11) is that, at any time, the present value of expected benefits from a dollar spent on data collection equals the present value of expected benefits from the same dollar spent on treatment at any of the sites. This must be so in an efficient allocation of the budget between (i) data collection and treatment efforts, and (ii) allocation of treatment efforts across the sites, for otherwise, either a reallocation between a and R^n or between R^n and R^j can further lower the present value of expected damage. In other words, the manager chooses data collection intensity and the levels of treatment effort at the polluting sources such that the adjusted marginal benefits obtained from data collection (a collective good) equals the adjusted marginal benefits from the treatment effort (site-specific goods).

V. EMPIRICAL APPLICATION

Redwood National Park was established in 1968 to preserve the coastal redwood (*Sequoia sempervirens*) forest ecosystem. The Tall Trees Grove, which includes several of the world's tallest trees, is located in the Park adjacent to Redwood Creek. Erosion from upstream logging activity and the associated road network increased sediment delivery to streams, resulting in destabilized stream channels, dramatic geomorphic changes, and the subsequent loss of aquatic and riparian habitat. Most importantly, the Tall Trees Grove was subject to increased flooding, bank erosion and an elevated water table [30].

To protect the grove, the United States Congress expanded the park in 1978 to include 36,000 additional acres upslope of the original Redwood Creek corridor. In 1981, Congress directed the United States National Park Service (USNPS) to minimize erosion from past land uses, re-establishing native patterns of vegetation, and protecting aquatic and riparian resources within tributaries and along the main stem of Redwood Creek [30]. The primary focus of erosion control would be to prevent or reduce erosion from logging roads within the Park [31].[4]

To date nearly 100 road rehabilitation projects have been completed within the Park boundary. We collected data on treatment levels and costs, and data collection costs from unpublished project reports for the period spanning 1981 to 1988. This time period provides the most complete data set. Daily stream flow and ambient sediment load measures were obtained from United States Geological Survey (USGS) gaging stations located along Redwood Creek (See Fig. 1). The Park arranged with the USGS to collect daily sediment load measures only at the Orick and Blue Lake gaging stations. The difference between the upstream and downstream sediment measures constitutes the total sediment entering Redwood Creek between Blue Lake and Orick. Daily stream flow, however, is measured at all seven gaging stations. In total, there are 957 daily observations of stream flow and sediment loading spanning five rain seasons. The data from the gaging stations delineates six catchment regions as polluting sources (see Table I).

In the empirical estimation the single ambient sediment measure is disaggregated among the six-catchment regions and then sequentially updated, as new data becomes available (i.e., with each observation). We also estimate and sequentially update the random state equation parameters. Two estimation scenarios are considered. In the first case, defined as high intensity data collection, we use the disaggregated stream flow data. In the second case, defined as low intensity data collection, the stream flow data for the lower reach of the creek (i.e., downstream from the Redwood Creek at Panther Creek gaging station) are combined into one stream flow measure, and the stream flow data for the upper reach of the creek are combined into a separate single measure. This delineation of Redwood Creek corresponds to the reaches of the Creek inside and outside of the Park boundary, respectively. We then use the relative area of each catchment region to reapportion the stream flow data to each region. These different data collection scenarios provide estimates of the sediment loading distributions, which are used later in the simulated sediment control model.

The empirical state and observation equations are defined as

$$E(s^n(t+1)) = E(s^n(t))[g^n(a(t+1)) - b^n(a(t+1))R^n(t)] + v^n(t), \forall n \qquad (12)$$

$$\ln q^n(t) = E(s^n(t))\ln(flw^n(t)) + w^n(t), \forall n \qquad (13)$$

where $q^n(t)$ is the actual unobservable sediment generated from the *nth* source, $flw^n(t)$ is the known stream flow from the *nth* source, and v^n and w^n are random errors that are also estimated. Equation (12) is a simplified combination of equation (2) and (3) with an additive error term.

We derive equation (13) by adding an error term to the log transformation on a known hydrological relationship between streamflow and sediment loading [28], which is

$$Q(t) \equiv \sum_n q^n(t) \equiv \sum_n \overline{q}^n(t) = \sum_n (flw^n(t))^{E(s^n(t))} . \tag{14}$$

Q is the ambient sediment load measure, and $\overline{q}^n(t)$ is the expected sediment generated from the nth source.[5]

To estimate probability distributions for the random model parameters, the following reparameterization from parameter space to probability space is necessary.

$$E(s^n(t)) = \sum_i s_i^n(t) p_i^n(t), \forall n,\ g^n(t) = \sum_i z_i^{n,g} g_i^n(t), \forall n,\ b^n(t) = \sum_i z_i^{n,b} b_i^n(t), \forall n,$$

$$w^n(t) = \sum_i z_i^{n,w} p_i^{n,w}(t), \forall n, \text{ and } v^n(t) = \sum_i z_i^{n,v} p_i^{n,v}(t), \forall n, \tag{15}$$

where $s_i^n(t), z_i^{n,g}, z_i^{n,b}, z_i^{n,w}$, and $z_i^{n,v}$ are support values for the respective distributions that represent the constraints on the probability space. The probability distributions sum to unity;

$$\sum_i p_i^n(t) = 1, \forall n,\ \sum_i g_i^n(t) = 1, \forall n,\ \sum_i b_i^n(t) = 1, \forall n,$$

$$\sum_i p_i^{n,w}(t) = 1, \forall n \text{ and } \sum_i p_i^{n,v}(t) = 1, \forall n \tag{16}$$

Substituting (13) into (10) and (11) yields the new reparameterized updating rules.

$$\sum_i s_i^n(t+1) p_i^n(t+1) = \left(\sum_i z_i^{n,g} g_i^n(t) - R^n \sum_i z_i^{n,b} b_i^n(t) \right) \left(\sum_i s_i^n(t) p_i^n(t) \right) + \sum_i z_i^{n,v} p_i^{n,v}(t), \forall n \tag{17}$$

and

$$\ln q_t^n = \left(\sum_i s_i^n(t) p_i^n(t) \right) \ln(flw_t^n) + \sum_i z_i^{n,w} p_i^{n,w}(t), \forall n \qquad (18)$$

In each time period t, the entropy specification of the objective function is

$$\begin{aligned} \underset{p^n, b^n, g^n, p^{n,w}, p^{n,v}}{Min} \quad & \sum_n \sum_i p_i^n \ln\left(p_i^n \big/ \tilde{p}_i^n \right) + \sum_n \sum_i b_i^n \ln\left(b_i^n \big/ \tilde{b}_i^n \right) \\ & + \sum_n \sum_i g_i^n \ln\left(g_i^n \big/ \tilde{g}_i^n \right) + \sum_n \sum_i p_i^{n,w} \ln p_i^{n,w} + \sum_j p_i^{n,v} \ln p_i^{n,v} \end{aligned} \qquad (19)$$

subject to equations (16), (17), and (18), where $\tilde{p}_i^n \equiv p_i^n(t-1), \forall n$, $\tilde{b}_i^n \equiv b_i^n(t-1), \forall n$, and $\tilde{g}_i^n \equiv g_i^n(t-1), \forall n$ are the prior probabilities.[6]

We assume the initial prior probabilities $\tilde{p}$, $\tilde{b}$, and $\tilde{g}$, are uniformly distributed. In other words, any state or value is as likely as any other and thus the probabilities will be equal to $\frac{1}{\sum i}$, where $\sum i$ is the total number of support values.[7] For tractability, and consistency with the earlier theoretical discussion, the number of support values is limited to two. The support values for s^n and $z^{n,b}$ range from 0.0 to 2.0. This implies that rainfall has a non-negative impact on sediment loading and treatment has a non-positive impact on sediment loading. Further, the upper bounds on s^n are set to contain estimates derived for other river systems as reported in Singh and Krstanovic [28]. The estimation tells us which catchment region deviates from the original uniform prior distribution. The values for $z^{n,g}$ range from 0.0 to 2.0 as well. This range allows the information effect (g^n) to increase or decrease the expected sediment loading but does not allow the loading to decrease with increases in rainfall. The values for z^w, and z^v, range from –3,000.0 to 3,000.0, and from –1.0 to 1.0 respectively. These ranges are conservatively set beyond the three-s rule for the estimated error terms [9, p. 88]. After estimating posterior distributions from the first observation, all subsequent prior probabilities are the previous period estimated posterior probabilities.

The estimation results for the sequential entropy filter appear in Fig. 2, 3 and 4. Fig. 2 illustrates the information acquired under the two data collection intensities. Recall that a normalized information entropy of one represents complete uncertainty. When the normalized information entropy approaches zero, the level of information approaches certainty. It appears that information is acquired (uncertainty is reduced) in the high intensity scenario to a greater extent than in the low intensity scenario. Over the time horizon represented in the data, roughly fifteen percent of the uncertainty associated with the pollution generation is reduced when data is collected with a high intensity. This result can also be interpreted as a fifteen-percent improvement in the predictability of the sediment loading parameters. The large and short-lived decrease in entropy seen just after the 800^{th} rain day represents a severe but short storm that occurred. We see that the entropy metric responds to this shock but then quickly returns to the earlier learning pattern.

Fig. 3 shows the low intensity scenario estimates for the six sediment loading parameters. All the parameters except for catchment region six remain at their subjective prior expected value of one. The change in the s6 parameter is most likely due to the fact that treatment effort only occurs in that region. We can infer from these results that catchment region 6's sediment loading is lower than previously believed. Fig. 4a and 4b show the result for the high intensity scenario. In these results we see greater movement away from the prior expected value of one for all the parameters. As with the low intensity result for s6, we see that the sediment loading from this region is lower than previously believed. We also see that sediment loading, on average is greater than previously believed for catchment region two, the upper Redwood Creek region. This result suggests that the majority of the sediment loading is more likely generated upstream from the park boundary and thus treatment effort may be more wisely spent treating roads outside the park. Overall, these results for the low and high intensity scenarios suggest that the level of aggregation of the stream flow measure affects the managers' ability to acquire information and produce more predictable parameter estimates. The results for b, the treatment parameter are not shown because they did not change from the prior expected value of 0.001. The estimated marginal information acquisition parameters (g) that are defined in equations (13) and (15) are shown in Fig. 5, 6a, and 6b**,** for region 6 in the low intensity scenario and regions 2 and 6 in the high intensity scenario, respectively. The remaining g parameters are not shown because they did not vary from the prior expected value of one.

VI. MODEL SIMULATION AND POLICY ANALYSIS

The control model defined in equation (7) was calibrated with the estimation results reported above and additional data provided by the Redwood National Park staff. The cost function coefficient was estimated with least squares using a linear specification without an intercept term, which resulted in a value for the estimated cost coefficient of 2178.1 with a t-statistic of 13.683. The damage function was estimated based on the assumption that the sediment related damage is greater than or equal to the resources spent on sediment control. The estimated damage function in dollars per year is

$$\ln(\text{Damage}) = -\underset{(-2.839)}{7.253} + \underset{(8.054)}{1.674}\ \ln(\text{sediment}),\ R^2 = 0.84,$$

where sediment is measured in cumulative average daily tons per rain season and the t-statistic are reported below the estimated parameters. The data collection cost under the varying intensities are *m*(low intensity) = \$41,154 and *m*(high intensity) = \$93,712, and the annual budget is B = \$200,000.

The policy analysis involves simulating the model under three scenarios. First, a uniform treatment policy is simulated, where all catchment regions are assumed to load sediment in proportion to their relative area. This first scenario coincides with the current policy employed in the Park which assumes that the largest catchment region (region 6) is the largest contributor to the sediment loading observed at the Orick gaging station. This assumption implies that all the treatment effort will be concentrated in this catchment region and no data are collected. Under these assumptions, that the optimal budget allocation requires that 92 roads are removed from region 6. Next, two scenarios depicting low and high intensity data collection are simulated. Here, the manager incorporates the information acquired through data collection to reallocate resources from relatively cleaner regions to relatively dirtier regions.

The results from the policy simulations are shown in Table II and III, and Fig. 7. First, we see in Table II that the optimal treatment for high intensity data collection has no treatment effort occurring in region 6 and under the low intensity data collection treatment effort in the later years moves out of region 6 into the dirtier regions. Given that the high intensity data collection allows the manager to treat the relatively dirtier regions we see in Fig. 7 the expected damages are lowest for this data collection scenario. Similarly, since the low intensity data collection allows the manager to eventually stop treating the relatively cleaner region, we also see that the low intensity scenario lowers expected damage below the levels for the status quo scenario. Table III lists the discounted expected damage levels for each scenario, assuming a three- percent discount rate. We clearly see that the present value of the stream of expected damages is greatest for the no data scenario and lowest for the high intensity scenario.

VII. SUMMARY

In this paper, we have presented an empirical application of the public management of NPS pollution under conditions of incomplete and costly information. The application was the case of sediment loading in Redwood Creek, which flows into and through Redwood National Park, located in Orick, California. The empirical estimation of the sediment loading parameters was conducted using the sequential entropy filter. This application shows the practical use of the entropy formalism for estimating ill-posed problems. We also employed the entropy metric to characterize the level of incomplete information about the pollution loading that provides a more tractable measure for determining whether a pollution problem is PS or NPS.

In addition, this paper simulated various policy scenarios, where three policy options were considered. These include a treatment policy where no data is collected and all polluting sources are assumed to produce the same sediment per square mile, and low and high intensity data collection policies. Data collection implies that the manager reallocates the budget to collect data in order to improve the treatment effectiveness and thus reduce sediment-related damage within the creek. The simulation results show that when sufficient information is acquired, the manager can more effectively allocate limited resources to reduce sediment-related damages. In other words, despite the budgetary restrictions on public expenditures, a greater intensity of data collection, and thus fewer resources available for pollution treatment, results in greater reductions in uncertainty and damages. This suggests that diverting some resources from treatment effort to information acquisition in order to improve the overall treatment effectiveness may enhance NPS control.

There are several limitations to interpreting these results, however. First, the analysis is based on a single sediment control program. We may wish in the future to consider other watersheds and, in particular, watersheds where economic activity is still occurring. Further, we have evaluated the program only over a short time period (5 years). This is a shortcoming for many resource problems. Time series on pollution related problems are just now being collected. It will be some time before long series are available. This condition demonstrates the advantages of sequential updating with the SEF, allowing policy makers to use new information, as it becomes available and not have to wait to learn where the pollution is generated. If we wait too long without we may find ourselves beyond the point of no return.

REFERENCES

1. Amacher, G.S., and A.S. Malik. "Bargaining in Environmental Regulation and the Ideal Regulator." *Journal of Environmental Economics and Management* **30**(1996):233-253.

2. Barrett, J., and K. Segerson. "Prevention and treatment in environmental policy design," *Journal of Environmental Economics and Management* **33**(1997):196-213.

3. Cabe, R., and J.A. Herriges. "The Regulation of Non-point Source Pollution under Imperfect and Asymmetric Information," *Journal of Environmental Economics and Management* **22**(1992):134-146.

4. Fahrmeir, L. "Posterior Mode Estimation by Extended Kalman Filtering for Multivariate Dynamic Generalized Linear Models," *Journal of the American Statistical Association*, **87**(1992): 501-509.

5. Farzin, Y.H. and J.D. Kaplan. "Nonpoint Source Pollution Control under Incomplete and Costly Information," *Fondazione Eni Enrico Mattei (FEEM)* Working Papers, No.32.99, May 1999.

6. Farzin, Y. H., "Optimal Pricing of Environmental and Natural Resource Use with Stock Externalities," *Journal of Public Economics* **62**(1996):31-57.

7. Fernandez, L. "Estimation of Wastewater Treatment Objectives through Maximum Entropy." *Journal of Environmental Economics and Management* **32**:2(1997):293-308.

8. Garvie, D., and A. Keeler. "Incomplete Enforcement with Endogenous Regulatory Choice." *Journal of Public Economics* **55**(1994):141-62.

9. Golan, A, Judge, G and D. Miller. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. John Wiley and Sons: New York, 1996.

10. Golan, A. Judge G., and L. Karp. "A Maximum Entropy Approach to Estimation and Inference in Dynamic Models or Counting Fish in the Sea Using Maximum Entropy." *Journal of Economic Dynamics and Control* **20**(1996):559-582.

11. Gray, W.B., and M.E. Deily. "Compliance and Enforcement: Air pollution Regulation in the U.S. Steel Industry." *Journal of Environmental Economics and Management* **31**(1996):96-111.

12. Gull, S.F. and J. Skilling. "Maximum Entropy Method in Image Processing." *IEEE Proceedings* **131**:6(October 1984) 646-659.

13. Gull, S.F., and G.J. Daniell. "Image Reconstruction from Incomplete and Noisy Data." *Nature* **272**:20(April 1978) 686-690.

14. Harford, J.D. "Measurement Error and State-Dependent Pollution Control Enforcement." *Journal of Environmental Economics and Management* **23**(1991):67-81.

15. Harford, J.D., and W. Harrington. "A Reconsideration of Enforcement Leverage when Penalties are Restricted." *Journal of Public Economics* **45**(1991):391-95.

16. Jaynes, E.T. "Information Theory and Statistical Mechanics." *Physics Review* **106**(1957):620-30.

17. Kaplan, J.D. “Nonpoint Source Pollution Control, Incomplete Information and Learning: An Entropy Approach,” Ph.D. Thesis, Department of Agricultural and Resource Economics, University of California, Davis (1999).

18. Kaplan, J.D. and R.E. Howitt, “Estimating Nonpoint Source Pollution: An Application of a Sequential Bayesian Entropy Approach.” Working Paper, Department of Agricultural and Resource Economics, University of California, Davis January 2000.

19. Khinchin, A.I. *Mathematical Foundations of Information Theory*. Dover Publications: New York,1957.

20. LaPointe, B. and P. Rilstone. "Environmental Inspections and Emissions of Pulp and Paper Industry in Quebec." *Journal of Environmental Economics and Management* **31**(1996):19-36.

21. Magat, W.A. and W.K. Viscusi. "Effectiveness of the EPA’s Regulatory Enforcement: The Case of Industrial Effluent Standards." *Journal of Law and Economics* **33**(1990):331-60.

22. Mount, J.F. *California Rivers and Streams: The Conflict Between Fluvial Process and Land Use*. University of California Press: Berkeley, 1995.

23. Rausser, G.C., and R.E. Howitt (1975), 'Stochastic Control of Environmental Externalities,' *Annuals of Economics and Social Measurement* **4**, 271-292.

24. Romstad, E., and O. Bergland. "Inducing Individual Firm Compliance to Emission Quotas When Abatement Costs are Private Knowledge." Department of Economics and Social Sciences, The Agricultural University of Norway, 1993.

25. Russell, C.S., Harrington, W., and W.J. Vaughan. *Enforcing Pollution Control Laws.* Resources for the Future: Washington, DC, 1986.

26. Segerson, K. "Uncertainty and Incentives for Nonpoint Source Pollution Control." *Journal of Environmental Economics and Management* **15**(1988):87-98.

27. Shannon, C.E. "A Mathematical Theory of Communications." *Bell System Technical Journal* **27**(1948):379-423.

28. Singh, V. P. and P.F. Krstanovic. "A Stochastic Model for Sediment Yield Using the Principle of Maximum Entropy." *Water Resource Research* **23**:5(1987):781-793.

29. Skilling, J. Ed., *Maximum Entropy and Bayesian Methods*. Kluwer: Cambridge, 1989.

30. Spreiter, T.A., Franke, J.F. and D.L. Steensen. "Disturbed Lands Restoration: The Redwood experience." Paper Presentation at the Eight Biennial Conference on Research and Resource Management in Parks and Public Lands, George Wright Society, Hancock, Michigan, April 17-21, 1995.

31. United States Department of the Interior (DOI). *Watershed Rehabilitation Plan: Redwood National Park*, United States National Park Service, US Government Press, Washington DC, 1981.

32. United States Environmental Protection Agency (EPA). Chapter 3: Management Measures for Forestry-Introduction, http://www.epa.gov/OWOW/NPS/MMGI/Chapter3/ch3-1.html, 1999.

33. United States General Accounting Office (GAO). *Water Quality: Federal Role in Addressing-and-Contributing to-Nonpoint Source Pollution*, GAO/RCED-99-45, US Government Press, Washington DC, 1999.

34. Xepapedeas, A.P. "Observability and Choice of Instrument Mix in the Control of Externalities," *Journal of Public Economics* **56**(1995):485-498.

35. Zellner, A. "Optimal information Processing and Bayes' Theorem." *American Statistician* **42**(1988):279-84.

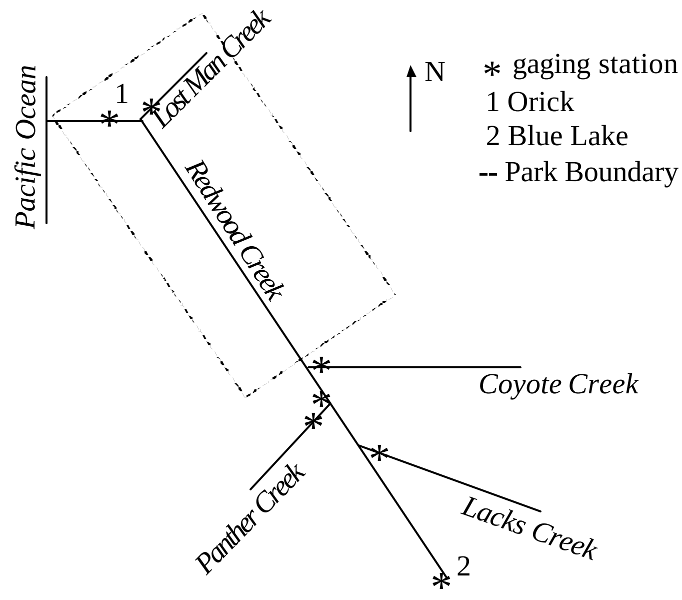


FIG. 1. A Simple Representation of Redwood Creek Watershed.

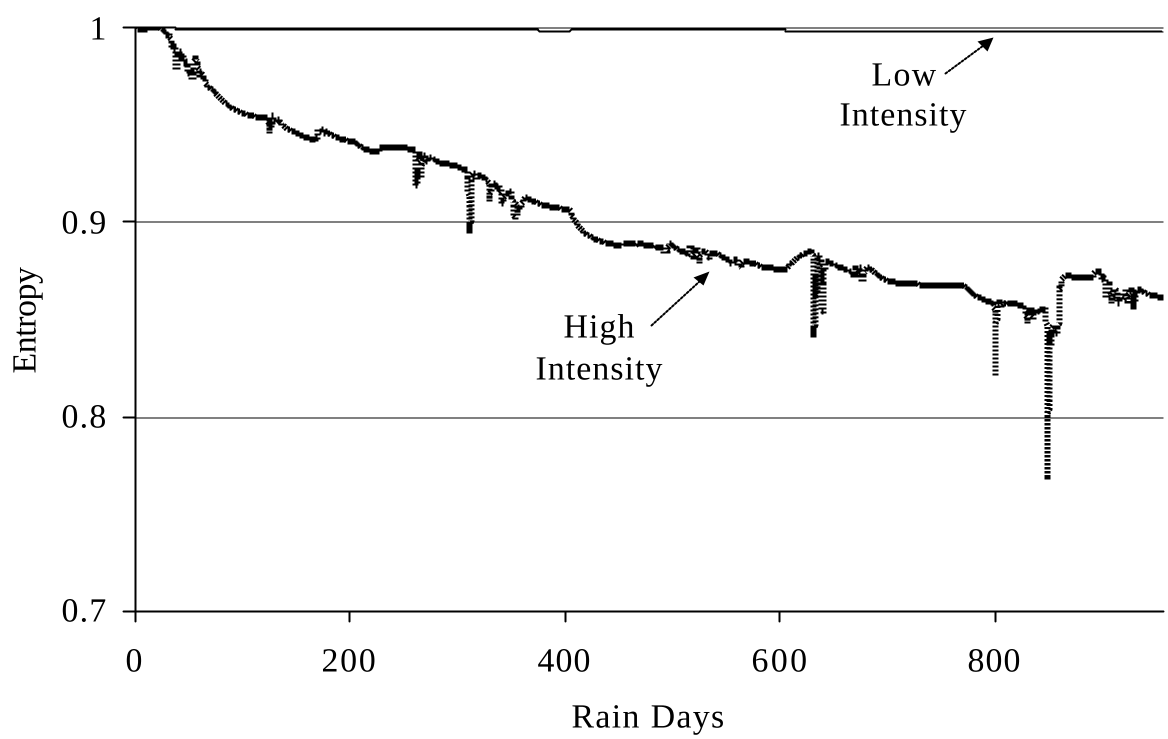


FIG. 2. Normalized Entropy for Low and High Intensity Data Collection.

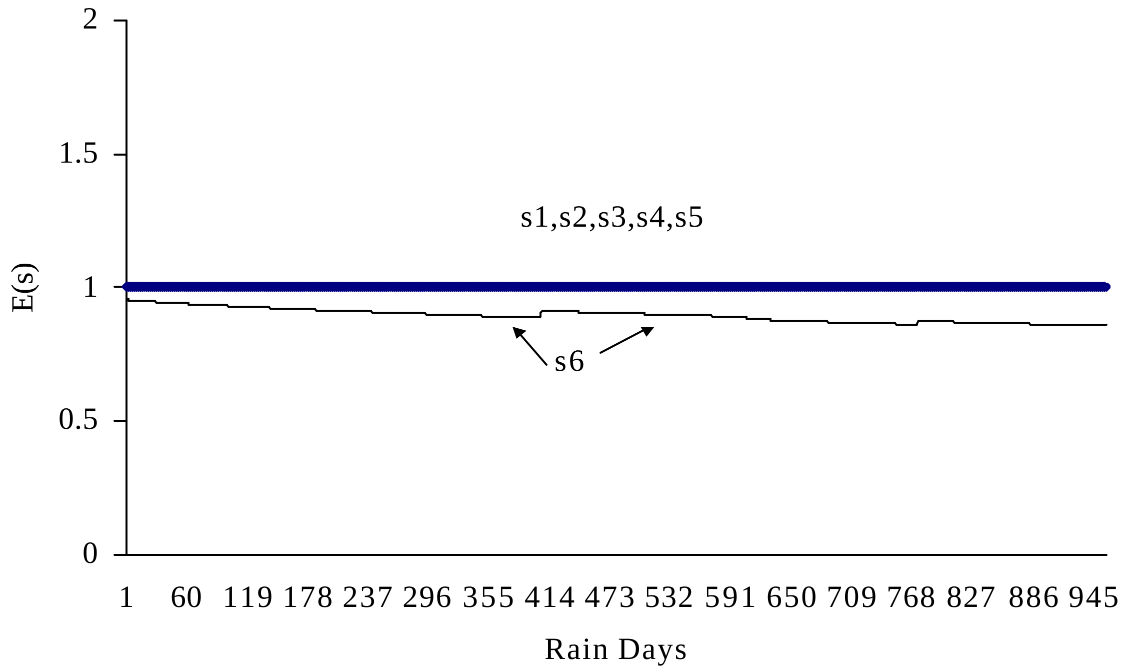


FIG. 3. Low Intensity Scenario Estimated Sediment Loading Parameters.

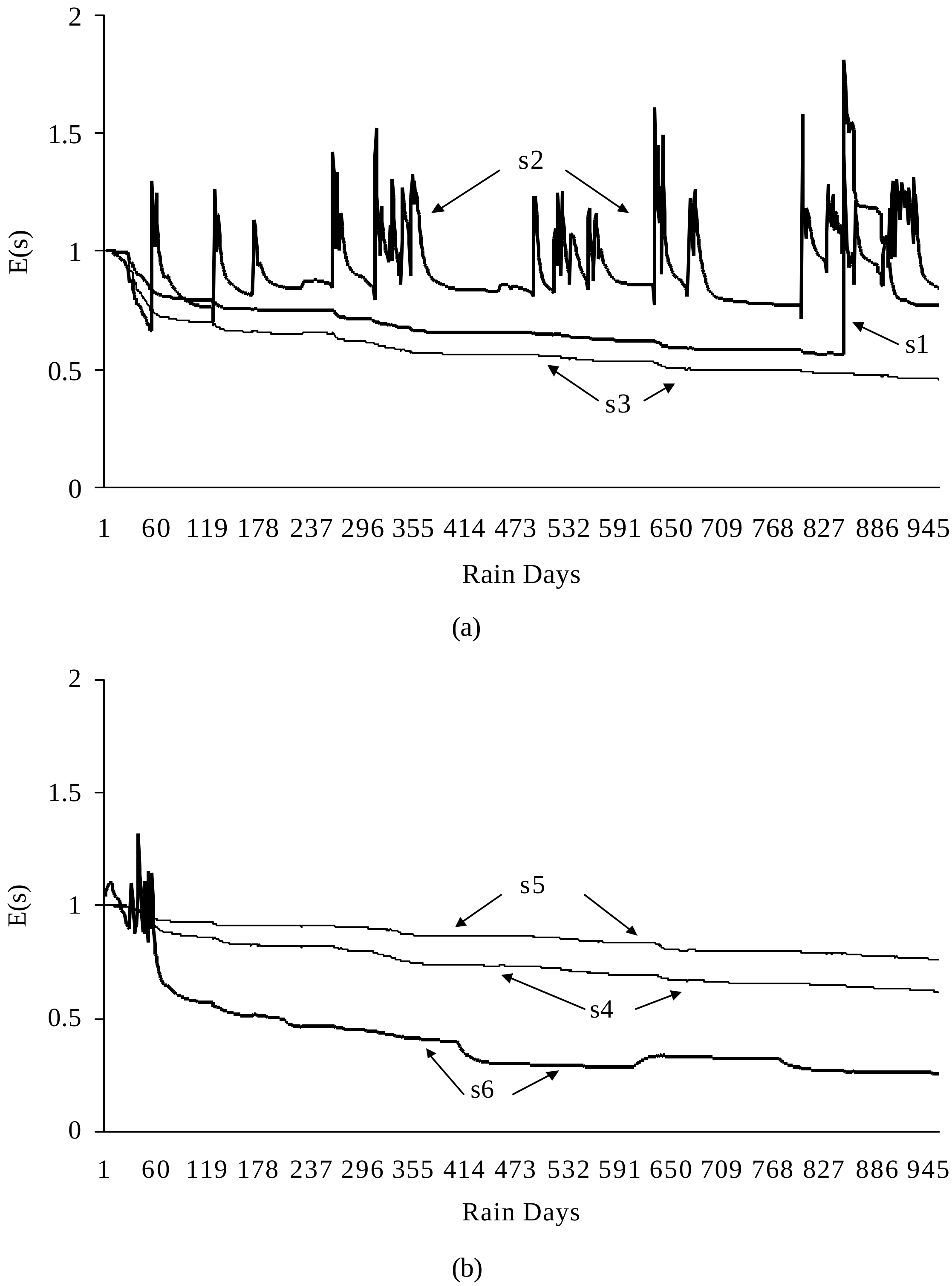


FIG. 4. High Intensity Scenario Estimated Sediment Loading Parameters (a) s1,s2, and s3; (b) s4,s5, and s6.

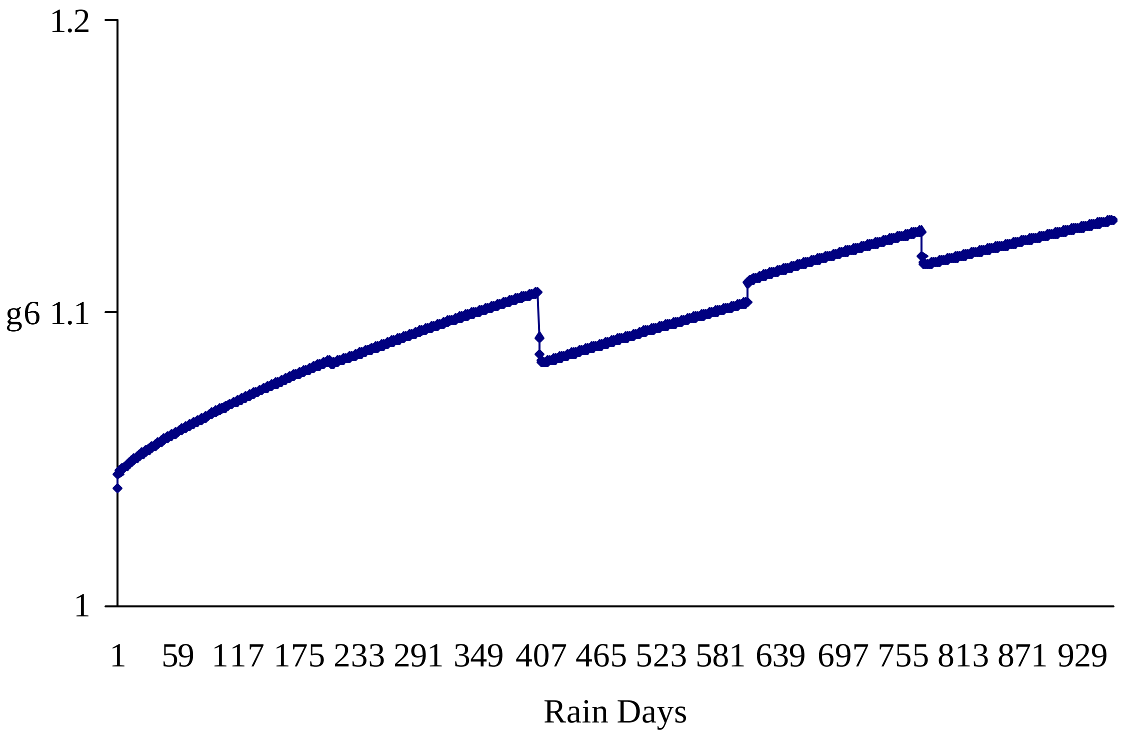
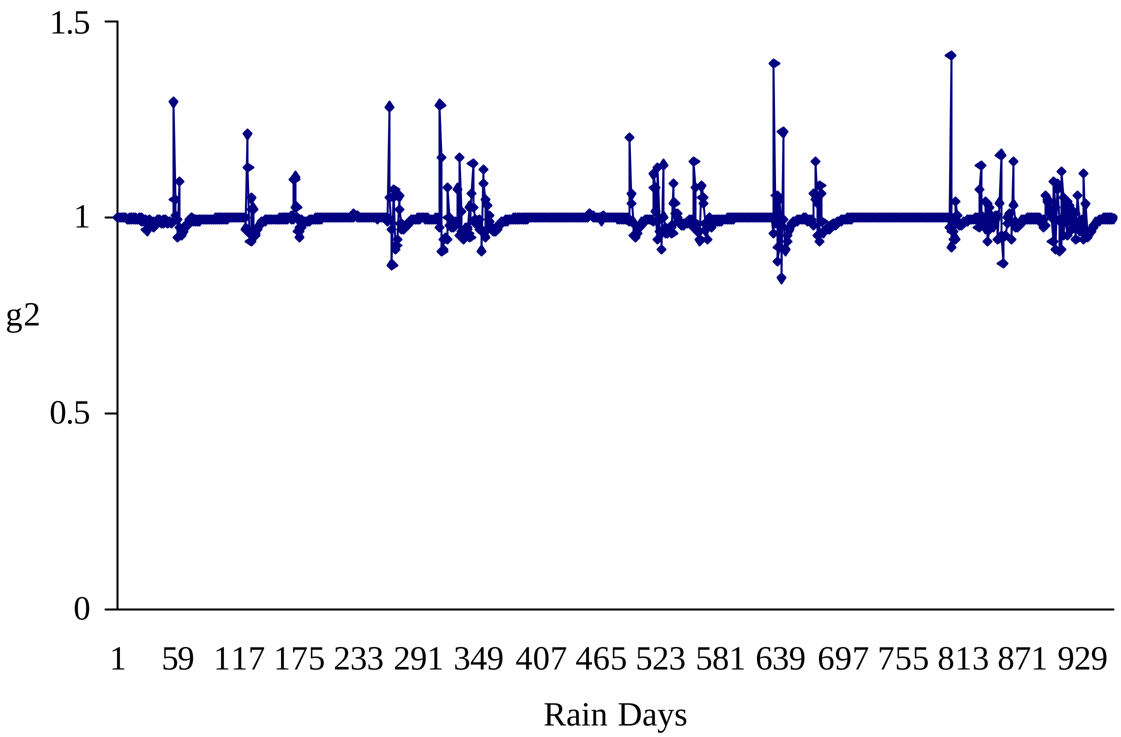
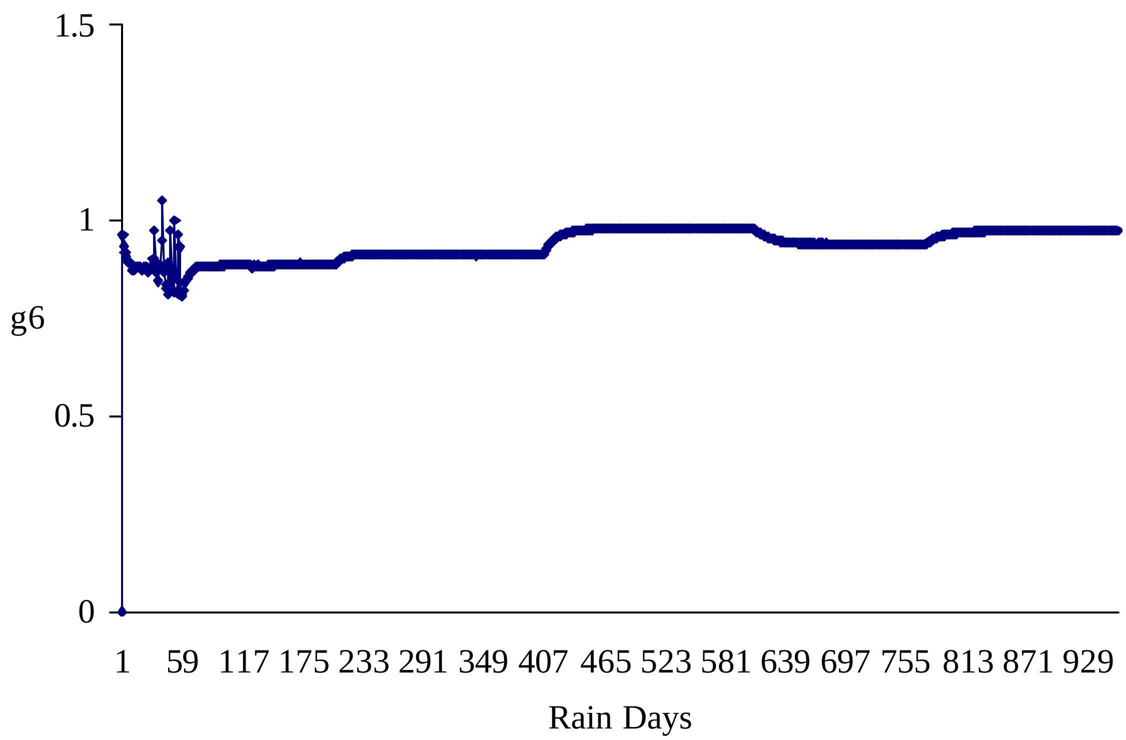


FIG. 5. Estimated g6 Parameter for Low Intensity Scenario.



(a)



(b)

FIG. 6. High Intensity Scenario Estimated Information Acquisition Parameters (a) g2; (b) g6.

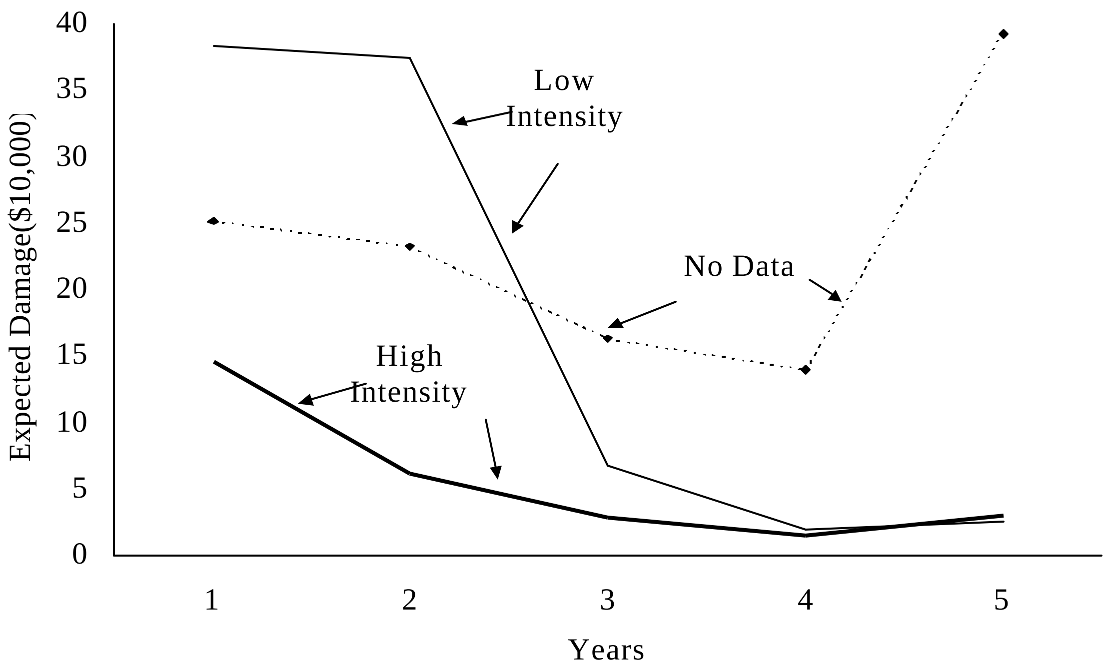


FIG. 7. Expected Damages under No Data, Low and High Intensity Data Collection.

Table I
List of Catchment Regions and Respective Area[8]

Variable	Catchment Region	Area (sq. mi.)
s1	Coyote Creek	7.78
s2	Upper Redwood Creek	65.4
s3	Lacks Creek	16.9
s4	Panther Creek	6.07
s5	Lost Man Creek	3.46
s6	Lower Redwood Creek	115.8

Table II
Optimal Treatment under Low and High Intensity Data Collection

High Intensity Treatment Decisions

Year	R1	R2	R3	R4	R5	R6
1	0	49	0	0	0	0
2	0	49	0	0	0	0
3	0	43	6	0	0	0
4	0	13	36	0	0	0
5	25	4	21	0	0	0

Low Intensity Treatment Decisions

Year	R1	R2	R3	R4	R5	R6
1	0	0	0	0	0	73
2	0	0	0	0	0	73
3	0	29	0	0	0	44
4	0	16	14	0	0	43
5	12	10	12	4	0	35

Table III
Discounted Expected Damage ($10,000) under No Data, Low and High Data Collection

	No Data	Low	High
1	24.44	37.18	14.19
2	21.89	35.23	5.75
3	14.91	6.20	2.60
4	12.40	1.77	1.36
5	33.82	2.22	2.60
sum	107.46	82.60	26.49

Footnotes:

1. If the manager cannot identify the sources, then it follows that she cannot know the magnitude of pollution generated by each source either. However, even if she knows the location, due to the effects of a mix of random events, she may not be able to identify the contribution to the total pollution from each source.

2. Farzin [6] illustrates the importance of accounting for interim damages in irreversible stock pollution problems.

3. We also assume that the economic activity, which initiated the generation of sediment loading, has ceased. However, the pollution loading continues. This type of problem is common in publicly owned forests and in mining regions where abandon mines become the responsibility of public entities.

4. Managers in the park identified the road network from past logging as the most probable sediment contributors to Redwood Creek. Logging roads are generally considered the principal contributor to sediment loading in logged forests [22, 32,33].

5. The identity on the left-hand side of equation (14) formalizes the fact that the observed ambient load must be equivalent to the sum of sediment from each source and to the sum of expected sediment from each source.

6. We drop the argument t from the objective function for convenience.

7. Although the managers have knowledge to differentiate the sources, the uniform distribution is chosen to reflect the knowledge prior to the enactment of the sediment control program. It is worth noting that the solution to the optimal control problem is sensitive to the initial prior subjective probability distribution given that the marginal benefit from data collection is

sensitive to this distribution. For instance, if the prior distribution is close to the actual sediment loading distribution then the marginal benefit of data collection will be low, and vice versa.

8. The Upper Redwood Creek region consists of Redwood Creek above Panther Creek and below Blue Lake accounting for Lacks Creek. The Lower Redwood Creek region is downstream of Panther Creek accounting for Coyote Creek and Lost Man Creek.

1. If the manager cannot identify the sources, then it follows that she cannot know the magnitude of pollution generated by each source either. However, even if she knows the location, due to the effects of a mix of random events, she may not be able to identify the contribution to the total pollution from each source.

2 Farzin [6] illustrates the importance of accounting for interim damages in irreversible stock pollution problems.

3 We also assume that the economic activity, which initiated the generation of sediment loading, has ceased. However, the pollution loading continues. This type of problem is common in publicly owned forests and in mining regions where abandon mines become the responsibility of public entities.

4. Managers in the park identified the road network from past logging as the most probable sediment contributors to Redwood Creek. Logging roads are generally considered the principal contributor to sediment loading in logged forests [22, 32,33].

5. The identity on the left-hand side of equation (14) formalizes the fact that the observed ambient load must be equivalent to the sum of sediment from each source and to the sum of expected sediment from each source.

6. We drop the argument t from the objective function for convenience.

7. Although the managers have knowledge to differentiate the sources, the uniform distribution is chosen to reflect the knowledge prior to the enactment of the sediment control program. It is worth noting that the solution to the optimal control problem is sensitive to the initial prior subjective probability distribution given that the marginal benefit from data collection is

sensitive to this distribution. For instance, if the prior distribution is close to the actual sediment loading distribution then the marginal benefit of data collection will be low, and vice versa.

8. The Upper Redwood Creek region consists of Redwood Creek above Panther Creek and below Blue Lake accounting for Lacks Creek. The Lower Redwood Creek region is downstream of Panther Creek accounting for Coyote Creek and Lost Man Creek.