APPLICATION OF BAYESIAN DECISION THEORY IN WELL FIELD DESIGN

by

Charles A. Bostock[1] and Donald R. Davis[2]

INTRODUCTION

Bayesian decision theory is a method for comparing expected utilities (costs and/or benefits) of alternative actions given various possible states of nature. The method treats uncertainty as to the true state of nature by determining the expected utility of each action in terms of the probabilities of the various possible states. The decision rule is to choose the action having the best expected utility.

One alternative may be to perform an experiment so as to obtain additional information about the true state of nature. Bayes theorem combines the additional information with information already available in order to improve the probability assessments over the various possible states. By considering the possible outcomes of the experiment and their probabilities, one can calculate the expected saving with the additional information. The expected saving is a measure of worth of the additional information.

Application of Bayesian decision theory has been demonstrated for a variety of uses involving surface water hydrology (Davis et al., 1972; Davis and Dvoranchik, 1971; and Lenten et al., 1974). The decision problem concerns the optimal level of flood protection in a stream channel in view of uncertainty in the magnitude and frequency of extreme flood events. Few applications in groundwater hydrology have been demonstrated. Gates and Kisiel (1974) investigated the problem of what type of new data collected in the Tucson basin would yield the most improvement in a digital model of the basin aquifer system. Uncertainty is in the values of the model parameters at each node, which best represent the true groundwater system. The decision problem is to determine in which parameters will errors have the greatest effect on water level predictions by the model.

The present paper illustrates one application of Bayesian decision theory in a particular kind of well field design problem. The decision to be made is what capacity-density combination to choose for wells located in an extensive, uniform grid. Uncertainty lies in anticipating the frequencies of transmissivity values among the wells. The well field design problem and method of solution were defined and discussed by Bostock (1975).

The application is described in three sections. A hypothetical example illustrates the application. The first section describes the conditions of the well field design problem, the assumptions necessary for its solution, and the preliminary cost calculations which are used later to determine utilities of action-state pairs in the Bayesian decision theory application. The second section describes the Bayesian decision theory application. The Bayes risk is the expected utility of any alternative action. Bayes theorem enables one to combine new information with old information in order to make better decisions and to estimate the worth of obtaining more information. In the third section, the application of Bayesian decision theory to this well field design problem is discussed.

WELL FIELD DESIGN PROBLEM

AQUIFER CONDITIONS

The well field design problem described here involves installing a large number of new wells over extensive, flat, delta regions, such as occur in Bangladesh. The aquifer is unconfined. With the exception of its permeability, the aquifer's properties are treated as being uniform laterally. Groundwater is mined during the annual dry season. An annual wet, or monsoon, season fully recharges the aquifer between pumping seasons.

---

1. Research Associate, Dept. of Hydrology & Water Resources, Univ. of Ariz., Tucson, 85721.
2. Asst. Professor of Hydrology & Water Resources, and Systems and Industrial Engineering, Univ. of Ariz., Tucson, 85721.

The illustrative example uses the following numerical values for aquifer properties; 1) specific yield is .15; 2) at the start of each pumping season, the water table is 20[ft] below ground surface; 3) although the aquifer thickness may be greater than 200[ft], the wells are all to be drilled to a uniform depth of 220[ft]. The maximum available drawdown is assumed to be 160[ft].

WELL FIELD OPERATION

Water demand is expressed as a depth, uniform over the ground surface. The well field is to be operated so as to produce a total demand, D[ft], during the dry season whose length is specified as t_{max}[sec]. The maximum production rate by the well field is to be D_i[ft/sec].

In the example, the operation requirements of the well field are as follows: 1) the total demand, D, for the pumping season is 5[ft]; 2) the maximum length of the pumping season, t_{max}, is 1.725×10^{7}[sec] or 200 days; and 3) the maximum instantaneous demand rate, D_i, is 8.70×10^{-7}[ft/sec].

WELL CAPACITY-DENSITY COMBINATION

It is desired to proceed directly with well field construction based on a uniform, square grid design. (See figure 1a) In the uniform grid, all of the wells have the same design and are equally spaced. The problem is to decide what capacity-density combination for the wells will minimize the costs of well construction, pumping, and losses due to well failures. Each capacity-density combination represents an alternative action. Note that the product of well capacity, Q_c[ft^3/sec-well], multiplied by well density, N_w[well/ft^2], must equal the maximum-instantaneous-demand rate, D_i.

Compared on the basis of costs per unit production rate, large wells are cheaper to construct than small wells. Thus, from the construction point of view, it is cheaper to construct fewer, larger wells spaced farther apart, in order to meet D_i, than to construct a larger number of smaller wells spaced closer together. From the pumping energy and well failure points of view, the reverse is true: pumping lift costs are less and well failures are fewer with smaller wells spaced more closely together. To obtain a measure of the utility of any well capacity-density combination, the total of these costs is estimated.

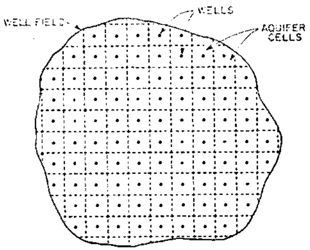
UTILITY OF A DESIGN

Homogeneous aquifer assumption. In order to develop the method, for evaluating the utility of a design, the aquifer is at first assumed to be homogeneous. In addition, all of the identically designed wells in the uniform grid are assumed to operate simultaneously. Under these conditions, a no-flow boundary surrounds each well as shown by the dashed lines in figures 1a and 1b. The squares enclosed by the no-flow boundaries can be thought of as aquifer cells. The no-flow boundaries are indicated as cell divisions in figure 1b.

Aquifer cell response model. All of the aquifer cells are of the same size and contain an identically designed well at their centers (See figure 1c.) Under the homogeneous aquifer and simultaneous pumping assumptions, aquifer response at each well is the same. Consequently, an aquifer response model representing a typical aquifer cell can be used to predict the following three quantities at each well: 1) the maximum drawdown at the wells due to any specified pumping rate and duration, 2) the energy consumed in pumping, and 3) the amount of water deficit, if well failure occurs before the end of the specified pumping duration.

Well failure occurs when drawdown in a well reaches the level of the pump intake. The lowest level for the pump intake corresponds to the level of maximum available drawdown. From the time of well failure until the end of the specified pumping duration, a failed well is assumed to produce at whatever rate groundwater will enter the well with well drawdown held at the level of the pump intake.

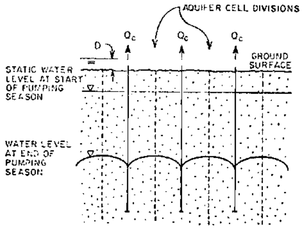
Costs per unit area and choice of designs. Costs at each well belong in one of the following four categories: 1) well construction and replacement costs, 2) pump purchase and replacement costs, 3) pumping energy costs, and 4) water deficit costs. Other maintenance and supervision costs are excluded from this analysis.



i) ALL WELLS ARE OF THE SAME DESIGN AND ARE LOCATED IN A UNIFORM GRID.

ii) N_w = DENSITY OF WELLS $\left[\frac{\text{WELL}}{\text{ft}^2}\right]$.

(a) PLAN VIEW OF WELL FIELD.



i) THE AQUIFER IS UNCONFINED AND HORIZONTAL.

ii) AREALLY UNIFORM RECHARGE OCCURS DURING AN ANNUAL WET SEASON.

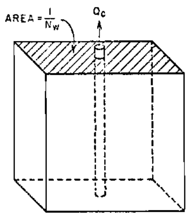
iii) PUMPING OCCURS DURING AN ANNUAL DRY SEASON.

iv) PERMEABILITY WITHIN EACH CELL IS CONSTANT, BUT VARIES AMONG DIFFERENT CELLS.

v) WATER DEMAND PER PUMPING SEASON, $D\left[\text{ft}\right]$, IS A UNIFORM DEPTH OF WATER OVER THE GROUND SURFACE.

vi) Q_c = DESIGN CAPACITY OF WELL $\left[\frac{\text{ft}^3}{\text{sec}}\right]$.

(b) SECTION OF AQUIFER THROUGH LINE OF WELLS



i) ALL CELLS ARE THE SAME EXCEPT FOR AQUIFER PERMEABILITY.

ii) INSTANTANEOUS DEMAND RATE = $D\left[\frac{\text{ft}}{\text{sec}}\right] = Q_c\left[\frac{\text{ft}^3}{\text{sec-well}}\right] \times N_w\left[\frac{\text{well}}{\text{ft}^2}\right]$.

iii) WELL FIELD DESIGN MODEL CONSISTS OF ONE WELL AND ITS APPROPRIATELY SIZED CELL OF AQUIFER.

(c) WELL WITH CELL OF AQUIFER.

FIGURE 1. Well field design model

In the example, well construction and replacement costs are expressed as a function of well design capacity, as follows: $560.\ Q_c^{.780}$[\$/year-well]. Pump purchase and replacement costs are expressed as a function of well design capacity and maximum drawdown in the well, as follows: $400. + 2.25\{Q_c - .713\}\{s_{max} + 12.\}$[\$/year-well]. The costs of well and pump components are amortized to equal annual payments over their estimated lifetimes.

Pumping energy costs are assumed to be directly proportional to the volume of water pumped multiplied by the pumping lift to ground surface in any time increment. The constant of proportionality used in the illustrative example was $3.06 \times 10^{-7}[\$/ft^4]$. Pump outlet pressure is assumed to be atmospheric, distribution of water from the wells being by open ditches.

Water deficit costs are equal to the value of the water deficit, had it not occurred. In the example, this cost was assumed to be directly proportional to the volume of the deficit, the constant of proportionality being $2.00 \times 10^{-4}[\$/ft^3]$.

For each well, these four costs are summed and divided by the ground surface area of the aquifer cell. (Figure 1c) The resulting annual cost per unit area is independent of the size of the well field area and the number of wells it would contain with different well densities. It provides a basis for comparing costs (i.e., utilities) of alternative well capacity-density combinations for a well field design.

Figure 2 shows the results of cost calculations which were obtained using this method (detailed by Bostock, 1975). These annual costs, expressed per unit area, are plotted against well capacity-density combination, Q_c - N_w, over the range of cell permeability values which are represented by the K's. Note that for each Q_c value shown as the abscissa in figure 2, there exists a corresponding well density, $N_w = D_1/Q_c$. The costs shown in figure 2 are the starting point in the application here of Bayesian decision theory. For each cell permeability value, the blacked-in circles indicate the well capacity-density combination with minimum cost. If permeability were homogeneously distributed in the aquifer and the true value were known, then the well capacity-density combination that minimizes costs could be chosen from the graph.

Heterogeneous aquifer assumption. In real aquifers, transmissivity is distributed heterogeneously. Common practice is to assign to each well, in an existing well field, a transmissivity value determined by pump testing it. In the well field design method described here, aquifer heterogeneity is accounted for by assuming that each aquifer cell will have a transmissivity value measured by a future pump test of its well. Uncertainty exists in the frequency distribution of transmissivity values to be realized in the completed well field.

Since the aquifer is dewatered in part by pumping the wells, transmissivity changes in time particularly in the immediate vicinity of each well. Therefore, of interest is "field permeability" which is equal to initial transmissivity at a well divided by initial saturated thickness of the aquifer at the well.

The heterogeneity of field permeability is accounted for in the method by the assumptions that 1) permeability within each cell is constant, but 2) permeability among different cells may be different.

Effects of intercell flow across the assumed bourdaries of no-flow around each cell were examined by Bostock (1975). For the conditions of this example, they were found to be negligeable. Also, intercell flow due to non-simultaneous well operation was considered. It was found that a maximum period of non-simultaneous pumping, such that effects are insignificant, can be determined. However, the total volume of water pumped from each well during this period must be the same.

PDF defines heterogeneity. The total well field cost (expressed per unit area) is of interest rather than the cost at each individual well. Therefore, it is the frequencies of the various cell-permeability values that is important, rather than the spatial distribution of field permeability. The frequency of each cell-permeability value can be defined by a probability density function (pdf).

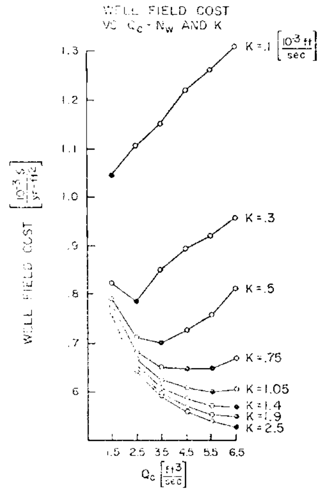


FIGURE 2. Annual well costs per unit area versus well capacity-density combination (Q_c-N_w) and aquifer cell permeability (K).

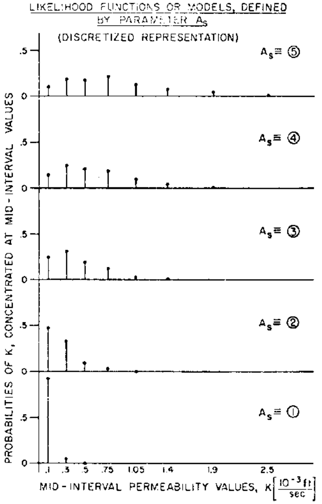


FIGURE 3. Possible likelihood functions or models, which represent frequencies of field permeability values among the aquifer cells.

The critical uncertainty to be examined in this paper, is in the eventual frequency distribution of permeability to be realized in the completed well field. In this Bayesian decision theory application, the state of nature (distribution of permeability in the aquifer) is represented by a pdf.

BAYESIAN DECISION THEORY APPLICATION

Bayesian decision theory is a method for comparing expected utilities of alternative actions, given possible states of nature about which there is uncertainty. In the previous section, the alternative actions were defined as choices for well capacity-density combination. The state of nature was defined to be represented by a pdf. In the present section, we define the probabilities of various possible states of nature, and the utility, or risk, of each action-state pair. Next, the Bayes risk is defined as the expected utility of any action in view of uncertainty about the true state of nature. The decision rule is to choose the action with optimum Bayes risk. Then, Bayes theorem is defined and used in combining new information with old information to obtain better assessments of the model probabilities. Finally, a method for calculating the expected worth of new information is described.

LIKELIHOOD FUNCTIONS

The unknown state of nature is the distribution of field permeability in the aquifer as defined by a pdf. Since the well field will have many wells, it is not the distribution in space that matters, but the frequency distribution of values among the wells. A likelihood function is a pdf that defines one possible state of nature. A likelihood function can also be thought of as defining the probability, or likelihood, of finding any permeability value at a point chosen randomly over the well field area. Since there is uncertainty as to what the relative frequency distribution of field permeabilities among the completed wells will be, a set of likelihood functions is defined in order to cover the range of variations considered possible for the particular well field area. Figure 3 shows the five likelihood functions which were used in the example, to model the possible states of nature.

RISK

Assuming a given state, represented by one of the likelihood functions on Figure 3, the annual average cost per unit area for the well field can be calculated for each capacity-density combination by using the cost data on figure 2. That is, the next step is to weight the costs, located vertically above each of the actions denoted by Q_c on figure 2, by the relative frequency of their K value in accord with the particular likelihood function from figure 3. The total of these weighted costs is the average cost for that action-state pair, and is called the risk. Thus, the risk is the average utility, or cost in this case, for a given action-state pair. Figure 4 shows these risks plotted as circles.

PRIOR PDF

The various likelihood functions model the range of possible states of nature. We are uncertain as to which of the likelihood functions represents the true state of nature. To account for this uncertainty, we use a "prior" pdf to define the probability of each likelihood function modelling the true state of nature. The prior pdf results from the assessment of the probabilities of the models based on presently available knowledge, i.e., available prior to obtaining more information. The left side of figure 5 shows the prior pdf used in the example.

BAYES RISK

Up to this point, we have a risk, or annual average cost per unit area, for each action-state pair, and a prior pdf which defines the probabilities as to which likelihood function models the true state of nature. Now, for each action, we can calculate the expected risk in accord with the prior probabilities of the models. The expected risk (expected utility) for a given action (well capacity-density combination) is called the Bayes risk. The Bayes risk is calculated in the same manner as the risk, but the costs for the heterogeneous aquifer models, shown by the circles on figure 4, replace the costs for the homogeneous aquifer cell-permeabilities, shown on figure 2. In addition, the prior probabilities on the left side of figure 5 replace the relative frequencies

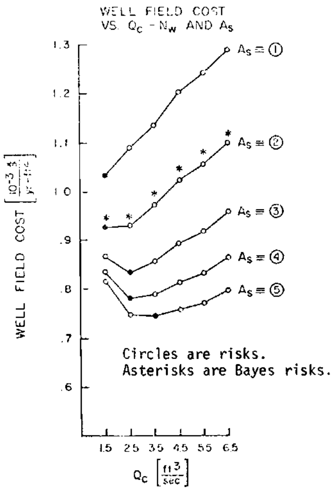


FIGURE 4. Risk of each action (Q_c-N_w)-state (A_s) pair, and Bayes risk of each action.

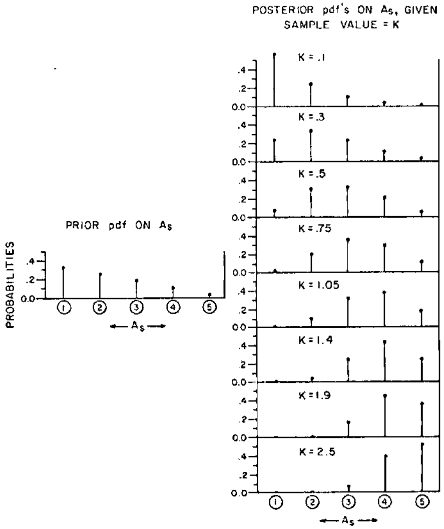


FIGURE 5. Prior and posterior pdf's.

from a particular likelihood function. That is, to evaluate the Bayes risk for any action, the costs located vertically above that action on figure 4 are weighted by their model probabilities according to the prior pdf on figure 5. The Bayes risk is the total of these weighted costs, and is shown for each well capacity-density combination by the asterisks on figure 4. The close correspondence on figure 4, between the Bayes risks and the risks for model As2, is coincidental.

The decision rule is to choose the action having the optimum Bayes risk, which in this case is the minimum expected cost. In the example, the choice indicated by the asterisks on figure 4 is somewhere between Q_c = 1.5 and 2.5.

BAYES THEOREM AND THE POSTERIOR PDF

Bayes decision theory is used for decision making when there is uncertainty about the true state of nature. If the degree of uncertainty could be reduced, then in many cases, better decisions could be made. One way of obtaining more information, about the frequency distribution of field permeability in an aquifer, is to conduct an experiment as follows: Drill a test well at some location in the aquifer and measure field permeability there by means of a pump test. The result is a sample permeability value from the aquifer, and this constitutes more information.

Bayes theorem enables combining new information from the sample with old information expressed in the prior probabilities of the models. Accordingly, the posterior pdf is conditional on the sample value. Figure 5 shows the range of posterior pdf's from the possible sample values in the example.

The appropriate posterior pdf becomes a new prior pdf from which a new set of Bayes risks can be calculated. The new set of Bayes risks will indicate an optimal action based on both the new and the old information, and which may or may not be different from the action indicated with only the old information.

Bayes theorem (see figure 6) states that the probability of a particular model given the sample (experiment outcome) equals the probability of the sample given the model (likelihood function for the model) times the probability of the model (prior pdf) divided by the probability of the sample. The probability of the sample is derived from the likelihood functions and the prior pdf as follows: the probability of the sample equals the probability of the sample given the model times the probability of the model, summed over all models. This denominator in Bayes theorem is a normalizing factor. It makes all the probabilities of the models add up to one. The pdf defined over all possible sample values is called the predictive pdf (see figure 7).

EXPECTED WORTH OF ADDITIONAL INFORMATION

Before deciding to acquire more information, we would like to know whether or not the worth of the information will justify the cost of obtaining it. We can estimate this worth by considering each possible sample value from the proposed experiment, and calculating the optimal Bayes risk associated with it using Bayes theorem. Knowing the optimal Bayes risk for each possible sample, the expected optimal Bayes risk can be calculated using the predictive pdf of the sample. The difference between the expected optimal Bayes risk and the Bayes risk is the expected worth of the additional information. The decision rule is that if the expected worth of the information from the experiment exceeds the cost of the experiment, then the experiment should be done.

DISCUSSION

EXPECTED VALUE VIEWPOINT

The logical choice for the well field design is the capacity-density combination giving the optimal Bayes risk, i.e., the minimum expected cost. Because the Bayes risk is based on the expected value viewpoint, a small portion of the time poor results will be obtained in the form of a high frequency of low permeabilities giving well failures. The probability of this can be reduced to any degree by sufficiently increasing the cost of the water deficit resulting from well failures.

BAYES' THEOREM

$$\overbrace{P[\text{MODEL}|\text{SAMPLE}]}^{\substack{\text{POSTERIOR p.d.f.}\\\text{(CONDITIONAL}\\\text{ON SAMPLE)}}} = \frac{\overbrace{P[\text{SAMPLE}|\text{MODEL}]}^{\substack{\text{LIKELIHOOD}\\\text{FUNCTION}\\\text{(DESCRIBES MODEL)}}} \times \overbrace{P[\text{MODEL}]}^{\substack{\text{PRIOR p.d.f.}\\\text{(PROBABILITY}\\\text{OF THE}\\\text{MODELS)}}}}{P[\text{SAMPLE}]}$$

$$\text{WHERE: } \underbrace{P[\text{SAMPLE}] = \sum_{\text{ALL MODELS}} P[\text{SAMPLE}|\text{MODEL}] \times P[\text{MODEL}]}_{\text{PREDICTIVE p.d.f. OF THE SAMPLE}}$$

FIGURE 6. Bayes theorem.

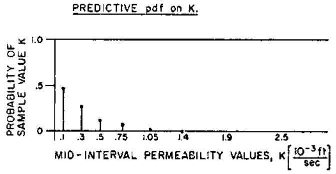


FIGURE 7. Predictive distribution of the sample, that is, a permeability value at any random location on the aquifer.

JUDICIOUS CHOICE OF SAMPLING POINTS

The locating of sample points in the aquifer, for conducting measurements of field permeabiltiy by test wells, requires judgment. There is usually some correlation between permeabilities at different locations in the aquifer, but generally, the correlation tends to decrease with increasing distance between sample points. If several samples are to be obtained, their locations should be distributed uniformly over the aquifer area to which the likelihood functions apply. Clustering of sample points will tend to reduce the value of the information obtained because of the possible correlation of values between nearby points and the redundancy of new information obtained. With clustered sample points, one would in effect be taking the same sample repeatedly which could lead to misleading results.

CHOICE OF MODELS AND PRIOR DISTRIBUTION

The results of the analysis are influenced by the choice of pdf models to represent the possible permeability distributions that might be present in the well field area, and by the prior probabilities assigned to them. The choosing of the models and the delimiting of the land surface area over which they apply, as well as the assessment of prior probabilities, should be based on the best available information and understanding of the hydrogeology of the proposed well field area. Note that it is implicit in the method that one of the likelihood functions is a close approximation of the true state of nature. If none of the models resemble the true state of nature, any good decisions resulting from application of Bayes decision theory will be fortuitous.

OTHER APPLICATIONS

The example given in this paper has considered uncertainty in the spatial distribution of field permeability, with regard to a particular kind of well field design problem. Other uncertainties exist which may or may not be as important. For example, the quantity of groundwater stored in the aquifer at the beginning of the pumping season depends on a stochastic process involving the frequency and intensity of rainfall during the wet seasons when the aquifer receives recharge. The decisions to be optimized in this case are 1) to what depth to drill the wells, and 2) at what depth to set the pump intakes. The objective is to minimize the drilling costs and pump column costs in view of uncertainty in the initial water table elevation at the start of each pumping season.

The more general well field design method is a sequential procedure, involving drilling test holes and test wells. The results of each test influence how and where to test next, until a well field design eventually evolves from the information acquired. The authors are currently exploring prospects for employing Bayes decision theory to this more general procedure for well field design.

ACKNOWLEDGMENTS

The well field design method referred to in this paper is the result of doctoral research carried out by C. A. Bostock and directed by Professors E. S. Simpson and T. G. Roefs of the Department of Hydrology and Water Resources, University of Arizona. The doctoral research was supported in part by the Canadian Department of Environment, Inland Waters Branch, and in part by the U.S. Office of Water Resources, Grant A-0240-ARIZ. Research on the Bayesian decision theory application was partially supported by the Office of Water Resources Research and Technology Grant 14-31-0001-5056, B-043-ARIZ, entitled "Practical Use of Decision Theory to Assess Uncertainties in Actions Affecting the Environment."

REFERENCES CITED

Bostock, Charles A. 1975. Minimizing costs in well field design in relation to aquifer models. Ph.D. dissertation, University of Arizona, Tucson, Arizona, 144 p.

Davis, Donald R. and William M. Dvoranchik. 1971. Evaluation of the worth of additional data. Water Resources Bulletin 7(4):700-707.

Davis, Donald R., Chester C. Kisiel and Lucien Duckstein. 1972. Bayesian decision theory applied to design in hydrology. Water Resources Research 8(1):33-41.

Gates, Joseph S. and Chester C. Kisiel. 1974. Worth of additional data to a digital computer model of a groundwater basin. Water Resources Research 10(5):1031-1038.

Lenton, Roberto L., Ignacio Rodriguez-Iturbe, and John C. Schaake, Jr. 1974. The estimation of ρ in the first-order autoregressive model: a Bayesian approach. Water Resources Research 10(2):227-241.