Spatial Complexity and Reservoir Optimization Model Results

BY

BRIAN JOHN VAN LIENDEN
B.S. (University of California, Davis) 1998

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

CIVIL AND ENVIRONMENTAL ENGINEERING

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

____

____

____

Committee in Charge

2000

ABSTRACT

The complexity of a model can have important implications for its costs of development, ease of use, and the reliability of its output. However, there is no standard definition of model complexity in the literature and little acknowledgement of its many possible forms. Furthermore, it is difficult to determine the appropriate levels of complexity for a particular model because there are few quantitative studies of the effects of model complexity on results. In this paper, an attempt is made to classify various types of model complexity and to determine an appropriate indicator for quantifying each type of complexity. Using three of these indicators, network flow models of Northern California's water system are formulated and compared at six levels of spatial aggregation. The results show that some spatial aggregration is possible with little change in the overall results but that high levels of aggregation can cause significant errors. In addition, changes in complexity are shown to have more significant effects at the local level. Additional research is needed to more completely understand the effects of model complexity on results.

TABLE OF CONTENTS

INTRODUCTION

"One should not increase, beyond what is necessary, the number of entities required to explain anything." - William of Occam (1285-1349)

Recent advances in computer technology and in data management and acquisition have made possible the development of increasingly complex models. The high level of complexity incorporated in many recent models has renewed a debate over the merits of simple versus complex models. Advocates of complex models (Nihoul 1994) argue that they are more reliable, represent the system more comprehensively, and are less likely to be used inappropriately. However, simpler models are said to be less time-consuming and costly to develop, to require less data, and to produce results that are easier to understand and interpret (Wood 1989). The most common sentiment (Jackson 1975; Palmer and Cohan 1986; BDMF 2000) seems to be that modelers should attempt to develop models that contain just enough complexity to accurately accomplish the project objectives, but no more.

These arguments underscore the importance of selecting an appropriate level of complexity for each model. The level of complexity incorporated into a model has important implications for the costs and availability of input and calibration data, for running the model (i.e., run time), and for interpreting model output. Increased model complexity also is presumed to improve the reliability of the results. However, while it is relatively easy to measure the costs of increased complexity, very little data exists to evaluate the expected benefits of incorporating additional complexity into a model.

The purpose of this study is to evaluate the effect of varying spatial complexity on the results of CALVIN, an economically based network flow optimization model of California's water supply system (Howitt et al. 1999). The paper begins with a review of previous studies of the effects of complexity on model results, followed by a definitional discussion of model complexity. Next, the formulation of network flow optimization models of Northern California at six different levels of spatial complexity is described and the results of these models are presented and discussed. Finally, some conclusions are made on the impact of model complexity on results and their interpretation.

STUDIES OF MODEL COMPLEXITY

Considering the perceived importance of model complexity, relatively few studies compare the results of models at different levels of complexity. This is especially true for reservoir systems models. The only study found which directly compared the results of a simpler reservoir model with those of a more complex one is by Palmer and Cohan (1986), who modeled the hydropower system of the Columbia River using both a single reservoir and a multi-reservoir model. They found that the simple model returned monthly hydropower production values 7-8% higher than those of the more complex model from July to December and 3-4% lower from January to June, resulting in a net annual difference of only about 1%.

A common technique for modeling multi-reservoir hydropower systems is to use aggregation methods to simplify the solution of a stochastic dynamic program. The simplification of the problem is accomplished either by solving each of the reservoirs sequentially while aggregating the potential energies of the remaining reservoirs into one or two composite reservoirs (Turgeon 1981; Archibald et al. 1997) or by aggregating all reservoirs into a single reservoir and then disaggregating the solution to obtain the operations of each individual reservoir (Saad et al. 1996; Turgeon and Charbonneau 1998). Turgeon (1981) and Archibald et al. (1997) compared the solutions obtained with their aggregation methods with those obtained by solutions of the entire system using deterministic optimization models. Turgeon found a difference in average annual value of less than one percent for a system of reservoirs in series, while Archibald et al. found a difference in average annual value of 0.3-0.4% for systems of 3 to 4 reservoirs, 2.2% for a system of 8 reservoirs, and 3.1% for a system of 17 reservoirs. However, neither paper attempted to validate the results by operating the same systems using a simulation model. Furthermore, these numbers do not truly describe the differences between simple and complex model formulations because the simplification gained by the spatial aggregation is offset by the added complexity of incorporating stochastic series of inflows. Thus, it is difficult to evaluate the significance of these comparisons in terms of model complexity.

Loague and Freeze (1985) and Jakeman and Hornberger (1993) studied the effects of complexity on rainfall-runoff model results. Loague and Freeze modeled 269 rainfall events on three upland catchments using a regression model, a unit hydrograph model, and a quasi-physically based model. They found that the regression and unit hydrograph models, which are simpler, provided as good or better runoff predictions than the more complex physically based model. Jakeman and Hornberger tested the impact of the number of model parameters by modeling seven catchments using six different storage configurations. Although the available configurations included three three-reservoir configurations, the optimal configuration for every catchment was either one storage or two storages in parallel. Thus, these rainfall-runoff studies found that the most complex model formulation is not necessarily the most accurate. In these cases, the superior performance of the simpler formulations is due in part to an absence of sufficient data to accurately characterize the more complex formulation.

Fontaine (1995) provides evidence that a more complex rainfall-runoff model produces better results for extreme flood situations. The same extreme flood was modeled using both a simple and a complex modeling approach and the results of each were compared to observed data. The simple approach used the HEC-1 model applied as an event-mode model without calibration. The more complex approach used the Hydrologic Simulation Program-Fortran (HSPF) in continuous mode with extensive calibration. The results showed that the modeled results exceeded observed data in peak daily discharge by 40% with HSPF and by 79% with HEC-1. The modeled 5-day runoff volumes were 20% higher than observed data using HSPF and 29% higher using HEC-1. Although the more complex modeling approach yielded better results than the simpler approach its results are so different from the observed data that its implementation may not be worth the additional effort required.

The issue of modeling scale can be very important in hydrologic modeling. While at small scales the patterns of topography, soil, and rainfall are important in governing runoff hydrology, increases in scale cause corresponding increases in the variability of distributions that are sampled within the watershed area. Very large scales, however, contain inhomogeneities brought about by large-scale geologic formations. To solve this problem Wood et al. (1988) proposed the use of an intermediate scale, which they call the Representative Elementary Area (REA), at which the average hydrologic response is invariant or only slowly varies with increasing catchment area. Wood et al. (1988) and others (Wood 1995; Woods et al. 1995) have attempted to determine the size of the REA and found values ranging from 0.5 to 5 km^2.

For unsteady open channel flow, the nonlinear Saint Venant equations apply. However, as there are no exact solutions of these equations for any but the simplest problems, numerical solution schemes of varying complexity have been developed to obtain approximate solutions. Numerous studies compare the accuracy and computational effort of two or more such schemes. Ponce et al. (1978) compared the performance of a kinematic wave model, which neglects the inertia and pressure terms in the equations of motion, and a diffusion wave model, which neglects only the inertia terms, and compared the results with those of a full dynamic model. Both schemes gave good results in most cases, but the diffusion wave model better described the subsidence of the flood wave. Keskin and Agiralioglu (1997) compared the results of a general dynamic model with those of a simplified dynamic model, in which the derivative of the friction slope with respect to space is assumed to be negligible, and found similar results. Thus, these studies of unsteady open channel flow have found good correlation between the results of various schemes to approximate the Saint Venant equations in most cases.

Sinha et al. (1995) studied the effects of time and space scales on the results of a flood routing model. A finite-difference spectral method based on the Chebyshev collocation technique and a finite-difference Preissmann scheme were applied to route a log-Pearson Type III hydrograph through a wide rectangular channel. The results showed that the order of accuracy for time discretization is more important than for space discretization. The spectral scheme, which is almost second-order accurate in time and second-order accurate in space, performed worse with small time steps and better with larger time steps than did the Preissmann scheme, which is first-order accurate in time and infinite-order accurate in space.

Water quality modelers also have devoted some attention to evaluating the effects of modeling complexity. Costanza and Sklar (1984) classified 87 models of freshwater wetlands according to their degrees of articulation, which is a measure of complexity that takes into account the time and space scales and the number of components used by the model. For each model tested, an index was calculated for each of these factors and the model articulation equaled the average of the three indices. Each model was evaluated for descriptive accuracy by comparing the model output with a set of historical data. They found that the models with lower articulation (the simpler models) tended to have higher descriptive accuracy. These results also seem to run counter to the common assertion that more complex models are more reliable, although they may reflect errors caused by limited calibration data.

Warwick and Cale (1987) proposed a method for water quality models that uses Monte Carlo techniques to estimate the probability of achieving a desired level of reliability. The desired reliability is achieved by balancing the errors caused by choosing a model of inappropriate complexity (Type I error) and the errors caused by uncertainties in parameter characterization (Type II error) in order to minimize overall modeling error. Warwick (1989) found that reducing one type of error often causes an increase in the other kind of error, and often an overall reduction in model reliability. This indicates that more complex models are only valuable if adequate data exists to describe their parameters.

Many other studies of model complexity exist in a variety of modeling fields. These include: a comparison of two point snowmelt models under different weather and snowpack conditions (Bloschl and Kirnbauer 1991); a study of models describing the decrease of galactic cosmic rays applied at both one and two dimensions (Le Roux and Potgieter 1991); Stockle's (1992) study of the performance of plant canopy models at different levels of compelxity; Palsson's and Lee's (1993) study of red blood cell metabolism models; and a study of the effects of model complexity on the performance of automated vehicle steering controllers (Smith and Starkey 1995). With the exception of Stockle (1992), all of these studies concluded that simpler models yield inadequate results in some situations. For example, Le Roux and Potgieter found that the two-dimensional model matched observed data much better at very large radial distances. Smith and Starkey concluded that low-order vehicle steering controller models are inadequate for high-g manuevers. Thus, these models differ from reservoir, hydrologic, and water quality models in that the accuracy of their results is greatly influenced by model complexity. Perhaps more complete data sets are typically available in these fields to characterize more complex model formulations.

To summarize, several studies attempt to evaluate the effects of model complexity on results. However, these studies are distributed over a wide range of modeling applications, and an intensive investigation of the impacts of model complexity on results has not been performed for any particular class of models. In general, studies of reservoir, hydrologic, flood routing, and water quality models have shown that simpler model formulations are very often more accurate than more complex formulations. Typically, this result is attributed to a lack of available data to properly characterize the additional parameters present in the more complex model. Studies in other fields have shown complex models to be more accurate, which may be due to the availability of more complete data in those fields.

WHAT IS MODEL COMPLEXITY?

There is no single accepted definition of model complexity. Most authors who discuss the subject do not even attempt a definition. Applying the dictionary definition of complexity, Brooks and Tobias (1996) define model complexity as "a measure of the number of constituent parts and relationships in the model." With this definition, they distinguish a model's complexity with its level of detail. This is the definition used in this paper.

Numerous indicators can be used to measure a model's complexity (Palmer and Cohan 1986). It would be impossible to develop a single measure of complexity that incorporated all aspects of complexity. What is more useful is to evaluate alternative models for each indicator individually to get a general feeling for each model's level of complexity. Thus, model complexity can be divided into several distinct types of complexity and numerical indicators can be developed to quantify each type of complexity. Some types of model complexity, with example indicators, appear in Table 1.

Table 1. Types of Model Complexity

Complexity Type	Example Indicator
Spatial	Number of spatial variables and the degree to which they interact
Temporal	Number of time steps incorporated into the model
Input	Amount of input required to run the model
Uncertainty	Number of stochastic variables incorporated into the model
Programming	Length of the model's programming code
Interface	Complexity of the user's interaction with the model
Run-time	Amount of time required to run the model
Interpretation	Amount of time required to interpret the model results
Calibration	Amount of data needed to calibrate the model

In this study, the CALVIN model of California (Howitt et al. 1999) is used to model each test case - only the model input data is changed. Thus, the indicators for the uncertainty, programming, and interface classes of complexity are the same for each case. Temporal complexity is also the same for all alternatives because the same 72-year period of record has been run in monthly time steps for all six test cases. Input complexity is neglected because the test cases are simplifications of an existing model and so it is difficult to evaluate how much effort would have been required to assemble the data for each individual alternative. In some cases, data were already available in disaggregated form, so the input data cost may actually increase slightly for the less complex models. Finally, calibration complexity is neglected because this study is not concerned with the accuracy of any of the individual model runs, but only with the differences between them. This paper focuses on three types of complexity, spatial, run-time, and interpretation, which are discussed below.

1. Spatial Complexity

In this study, the spatial complexity is measured as the sum of the number of inflow links, reservoirs, and demand nodes contained in the system. A disadvantage of this measure is that certain aspects of spatial complexity, such as the representation of conveyance facilities, are neglected. However, in general the overall schematic complexity of each test case approximately corresponds to the relative values of spatial complexity. Input and calibration complexity are likely to increase with spatial complexity.

2. Model Run-time

In this study, model run-time complexity is measured in terms of the number of decisions required of the optimization model and the number of iterations needed to find a solution. These are considered more reliable measures of complexity than the actual run-time because they are not influenced by the computer's processing speed. With recent advances in computer technology, model run-time complexity is becoming less important as a practical consideration in model selection.

3. Interpretation Time

The interpretation time is the amount of time needed by the modeler to analyze the results produced by the model and to interpret their practical meaning. An important factor to consider when making such a comparison is what output should be generated from each model. A more complex model usually generates more detailed results than a simpler model. The time needed to generate and evaluate all of the potential results from a complex model can be much greater than is required to interpret more aggregated results similar to those produced by a simpler model. However, it is often necessary to look at a model's detailed results to ensure that the model is behaving reasonably. In this study, it would be difficult to accurately gauge the amount of time required to evaluate each case individually because all six test cases involve the same water system. Therefore, the interpretation time will be measured as the number of time series required to analyze each case and understand the results. This measure counts the deliveries to each demand, storage in each reservoir, surplus Delta outflow, marginal values of each inflow, and shadow values of each minimum flow, refuge demand, and surface water reservoir.

THE CALVIN MODEL

CALVIN is an optimization model developed at the University of California at Davis to describe the California inter-tied water system (Howitt et al. 1999). CALVIN uses the network flow reservoir optimization model HEC-PRM (USACE 1994) to maximize economic benefits by allocating water over a 72-year period of historical inflows. Economic benefits are characterized by piece-wise linear economic value functions at each demand location. The entire CALVIN model contains 56 surface water reservoirs, 30 groundwater reservoirs, 29 urban demand regions, and 26 agricultural demand regions. The model results contain monthly time series of flow and storage for every element in the system. The alternative used as a base case for this study is that of an unconstrained water market, with water allocations fettered only by physical and environmental constraints. Because preliminary CALVIN runs have shown few shortages under this scenario, the amount of water available has been artificially reduced for this study by reducing the external inflows by 20% and by increasing the losses on demand return flows by 30%. While these assumptions are acceptable for studying model complexity, the optimization results presented here are not intended to accurately represent the system's current or potential operation.

For this study, only the northern portion of the CALVIN model (the Sacramento Valley and Sacramento-San Joaquin Delta) is used (see Figure 1). The portions of the state south of the Banks and Tracy pumping plants and of the Calaveras River are assumed to be operated identically for each alternative and are not modeled. The northern portion of CALVIN contains 17 surface water reservoirs, 9 groundwater reservoirs, 7 urban demand regions, and 9 agricultural demand regions. The groundwater storage in each agricultural region is depicted by a single reservoir. This representation is considered the base case (Test Case A) for the present study and is the most complex model tested. Simpler model formulations are aggregated versions of this base case.

Figure 1 (Part 1). Case A Schematic

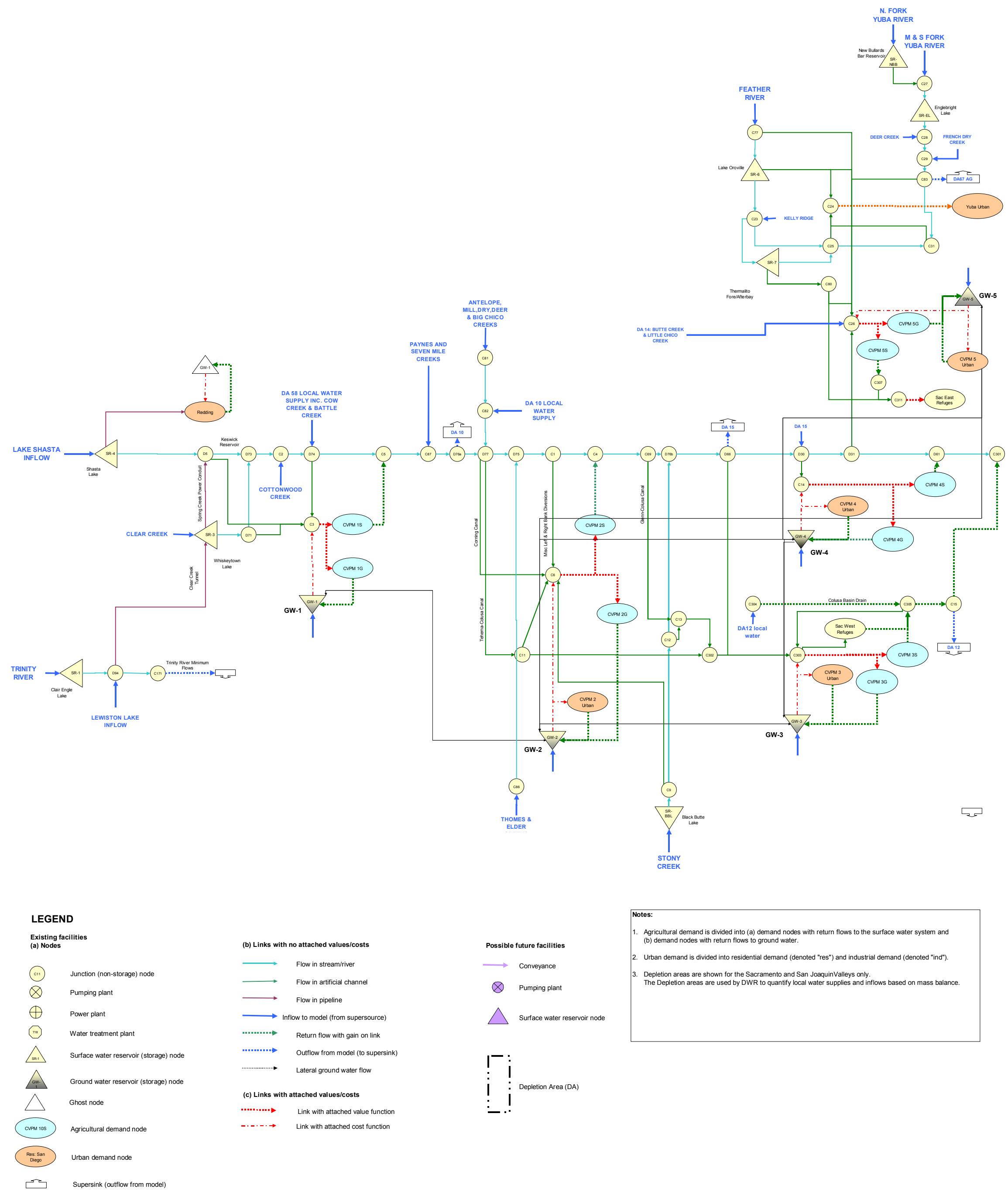
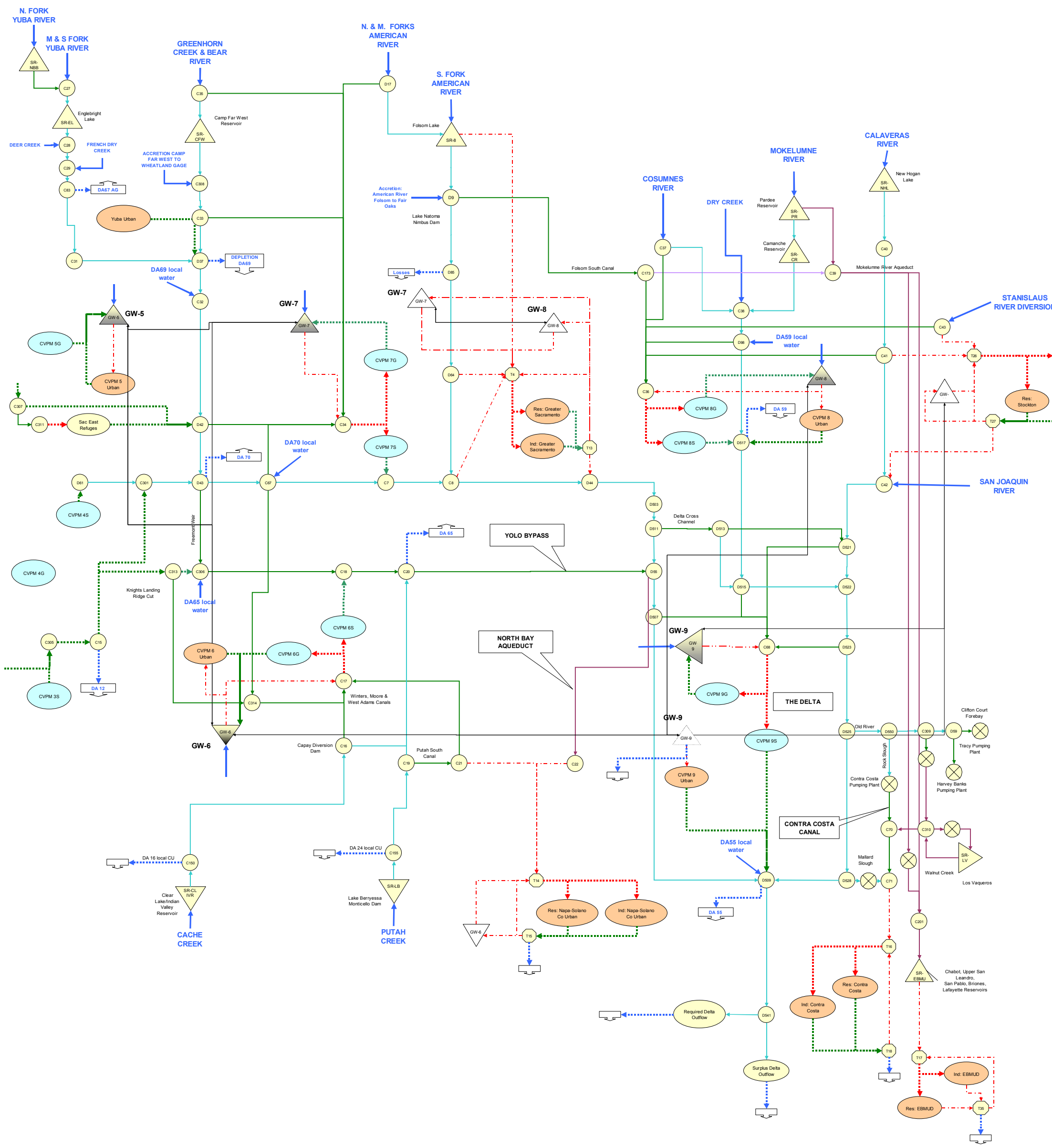


Figure 1 (Part 2). Case A Schematic



TEST CASES

Six test cases have been developed at different levels of aggregation. Table 2 shows the relative complexity of each alternative's physical representation. The sum of the number of inflows, reservoirs, and demand regions is used as an overall index of spatial complexity. A brief description of each case is given below. Schematic representations of Cases D-F can be found in the appendix.

Table 2. Test Case Spatial Complexities

Case	# Inflow Links	# Reservoirs	# Demand Regions	Spatial Complexity
A	40	25	15	80
B	28	19	13	61
C	17	13	12	42
D	12	7	7	26
E	4	2	2	8
F	3	1	1	5

Test Case A: Full CALVIN Representation

Case A is the full CALVIN representation of the northern portion of the California water system as described above and depicted in Figure 1.

Test Case B: Local Aggregation

Case B differs from Case A only in the aggregation of selected regional elements. The following pairs of reservoirs are combined to form individual composite reservoirs:

- Clair Engle Lake and Whiskeytown Lake
- Lake Oroville and the Thermalito Afterbay
- New Bullards Bar Reservoir and Englebright Lake
- Pardee Reservoir and Camanche Reservoir
- Groundwater Storage in agricultural regions 1 and 2
- Groundwater Storage in agricultural regions 3 and 4

In addition, agricultural regions 1 and 2 have been combined into a single demand region, as have agricultural regions 3 and 4. On the American River, the inflow for the north and middle forks has been combined with that of the south fork. Finally, the conveyance paths to the Contra Costa Water District via the Rock Slough, the Old River, and the Mallard Slough have been combined into a single link.

Test Case C: Aggregation by River System

In this case, the system is divided into nine regions. Within each region, the surface water storage nodes, groundwater storage nodes, agricultural demand nodes, and urban demand nodes are combined into single aggregate nodes of each type. In addition, the conveyance of water and the external inflow locations are greatly simplified. The nine regions are divided according to river system or local region as follows:

- Upper Sacramento River, including the Trinity River and Stony Creek
- Feather, Yuba, and Bear Rivers
- American River
- Cache and Putah Creeks
- Mokelumne River
- Calaveras River
- Contra Costa Water District
- Sacramento-San-Joaquin Delta
- San Francisco Bay Area portion of the EBMUD system

All of the system's water storage capacity and water demand is represented in the reservoir and demand nodes for these regions. However, the Mokelumne River region does not contain any urban demand, the Calaveras River region does not contain any agricultural demand, the Delta region does not contain any surface water storage or urban demand, and the Contra Costa Water District and EBMUD systems do not contain any groundwater storage or agricultural demand.

Test Case D: Aggregation of Eastern and Western Sacramento Valley

Case D is similar to Case C except that the river regions emanating from the Eastern portion of the Sacramento Valley (the Feather, Yuba, Bear, American, Calaveras, and Mokelumne Rivers), and those emanating from the Northern and Western portions (the Sacramento and Trinity Rivers and Cache, Putah and Clear Creeks), are each combined into aggregate regions representing the Eastern and Western portions of the Sacramento Valley. Within each of these regions, the surface water storage, groundwater storage, agricultural demand, and urban demand are combined into single aggregate nodes. The representation of the conveyance and external inflows into these regions are also simplified relative to Case C. The Contra Costa Water District, Sacramento-San Joaquin Delta, and EBMUD regions are represented as in Case C.

Test Case E: Aggregation by Group Types

In Case E, surface water storage, groundwater storage, agricultural demand, urban demand, and the environmental refuges are each aggregated into individual nodes representing the entire contents of the modeled system. Groundwater pumping links are constrained so that only a limited amount (the Redding demand) can be pumped to urban areas. The external inflows have been combined into three links, one entering each storage node and one entering downstream of both nodes.

Test Case F: Full Aggregation

In Test Case F, all agricultural and urban demand is combined into a single demand node. While surface and groundwater storage remain separate, the configuration of the groundwater reservoir has been altered so that groundwater can supply both urban and agricultural demands up to the pumping limits.

ISSUES AFFECTING CASE COMPARISONS

Certain aspects of spatial aggregation make it difficult to compare the results of particular runs of each case. These limitations can be divided into two classes - those caused by the aggregation of spatial elements and data and those caused by limitations of the network flow optimization model.

Aggregation Limitations

When developing Cases B-F, the data in Case A was duplicated as thoroughly as was possible given the simplified spatial formulations. In doing so, however, assumptions had to be made to combine data from several different links and storage nodes in Case A onto individual elements in the aggregated systems. Of these, the most problematic involved the aggregation of demand regions and of the constraints and costs on conveyance facilities.

The aggregation of the demand regions was done such that the total demand of each aggregated region would equal the sum of the total demands of each of the component demand regions. While this assumption allows for one-to-one comparisons of deliveries between different cases, it does not account for the reduction in demand caused by reuse between the demand regions. To account for reuse, reuse links were added to each aggregated demand region that allowed for return flows to be re-routed back to the region's delivery node. The amount of return flow available for reuse was conservatively limited to the amount available to the downstream demand if the upstream demand were fully supplied. While this exaggerates the amount available for reuse in virtually all months for most of the aggregated agricultural regions, the actual implementation of reuse was very low. For example, while Case F has a reuse capacity of 800 TAF/year the addition of this reuse link increased the average annual deliveries by only about 10 TAF/year.

The aggregation of conveyance facilities affected the representation of canal capacities, minimum flow requirements, and pumping costs in the simpler cases. Because aggregated elements essentially have infinite and costless capacity between them, many canal capacity constraints are neglected altogether, and many of those that do appear are limited by the demand of a particular region within an aggregated region rather than by the size of a particular canal. While most minimum flow constraints and pumping costs are represented in all cases, they are aggregated in the simpler cases and therefore do not reflect local conditions. In a few cases it was necessary to eliminate certain data altogether. For example, in Case F the pumping costs to CCWD and EBMUD are neglected because the urban and agricultural demands are combined. This limitation may affect the economic results of Case F, and could explain why the Case F values shown in Tables 6 and 9 are consistently higher than those shown for Case E.

Limitations of Network Flow Optimization

For many water resource problems, there may be numerous near optima with very similar objective function values. Thus, with small variations in formulation (such as might arise in the aggregation of data), several different solutions may be possible (Rogers and Fiering 1986). While these possible solutions will have very similar objective function values, and most likely similar overall results, the values in any given year or in a given local region can be very different. It can therefore be difficult to interpret the differences in results between different cases in particular time periods because small changes in the input data can cause the model to arrive at a different optimum and yield different results. For example, although the average annual surplus Delta outflows for Case E are less than those for Cases A-D, during the 1982-83 flood Case E has much larger surplus Delta outflows (see Figure 12). To test the impact of such deviations, Case E was re-run with a cost of $0.10/AF added for every acre-foot of surplus Delta outflow above 11 MAF in any given month. The overall results were very similar, with the same average annual shortages and surplus Delta outflows, but the surplus Delta outflow in March 1983 was reduced from 25 MAF to 11 MAF.

MODEL RESULTS

The primary indicators used to evaluate the performance of the 6 cases are the average shortages and groundwater mining resulting from each model formulation. These results can be found for each case in Table 3. In CALVIN, the shortage is defined as the difference between the maximum economic demand and the actual delivery. Each case has 1,765 TAF/year of urban demand and 9,063 TAF/year of agricultural demand. Cases with higher complexity measures show larger urban and agricultural shortages and more groundwater mining. Because urban deliveries are valued much more highly than agricultural deliveries, all cases show much fewer urban shortages. Only Cases A-C show significant urban shortages. While all cases show significant agricultural shortages (due to the artificial reduction of available water), there is a gradual increase in shortage quantity as the complexity increases. The groundwater mining in Cases A-D is a function of the water balance in individual groundwater basins, explained later in this section.

Table 3. Comparison of Results for Cases A-F

Case	Total Shortages (TAF/year)	Agricultural Shortages (TAF/year)	Urban Shortages (TAF/year)	Groundwater Mining (TAF/year)
A	2,282	2,274	8	42.4
B	2,143	2,137	7	34.2
C	2,127	2,121	7	34.2
D	1,918	1,918	0	34.2
E	1,719	1,719	0	0
F	1,636	n/a	n/a	0

To evaluate the accuracy of each case, it is assumed that the results of Case A are 100% accurate and that any deviations between the results of Case A and any other case are the result of the spatial aggregation of that case. While in actuality Case A (as an optimization model) represents a major deviation from actual system operation, it is assumed that, given sufficient data, the most complex formulation will be the most accurate and therefore can be used as a benchmark for evaluating other formulations. Two measures used for comparison are the total shortages and the average annual net delivery, which is calculated as follows:

Net Delivery = Total Demand – (Total Shortages + Groundwater Mining).

The percent errors are calculated for each case by taking the percent difference between the net delivery or total shortage for the case in question and that for Case A. The percent errors in net delivery and shortage are shown for each case in Table 4. Table 5 contains the complexity measures for each case. Figure 2 shows the percent error in net delivery versus percent complexity for each complexity measure. These results show

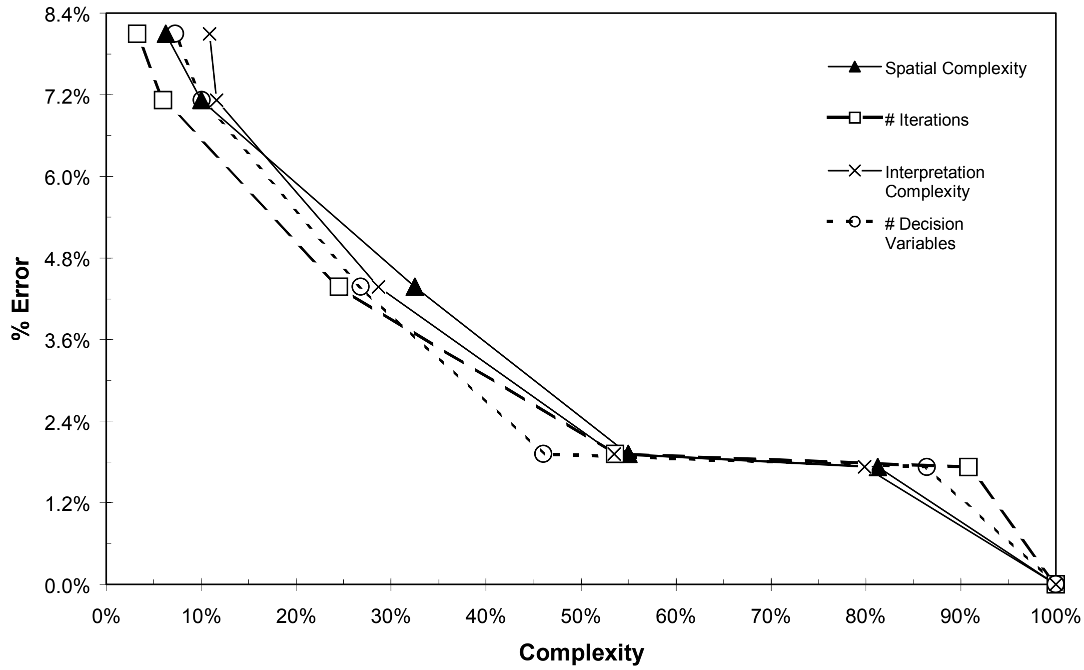
Table 4. Net Delivery and Errors Relative to Case A

Case	Net Delivery (TAF/year)	% Error in Net Delivery	% Error in Total Shortage	Average Net Difference in Monthly Shortage (TAF/month)	Average Absolute Difference in Monthly Shortage (TAF/month)
A	8,504	0.0%	0.0%	0.0	0.0
B	8,651	1.7%	6.1%	11.5	13.3
C	8,667	1.9%	6.8%	12.9	19.0
D	8,876	4.4%	16.0%	30.3	32.9
E	9,109	7.1%	24.7%	46.9	49.2
F	9,192	8.1%	28.3%	53.8	54.8

Table 5. Complexity Measures for Each Case

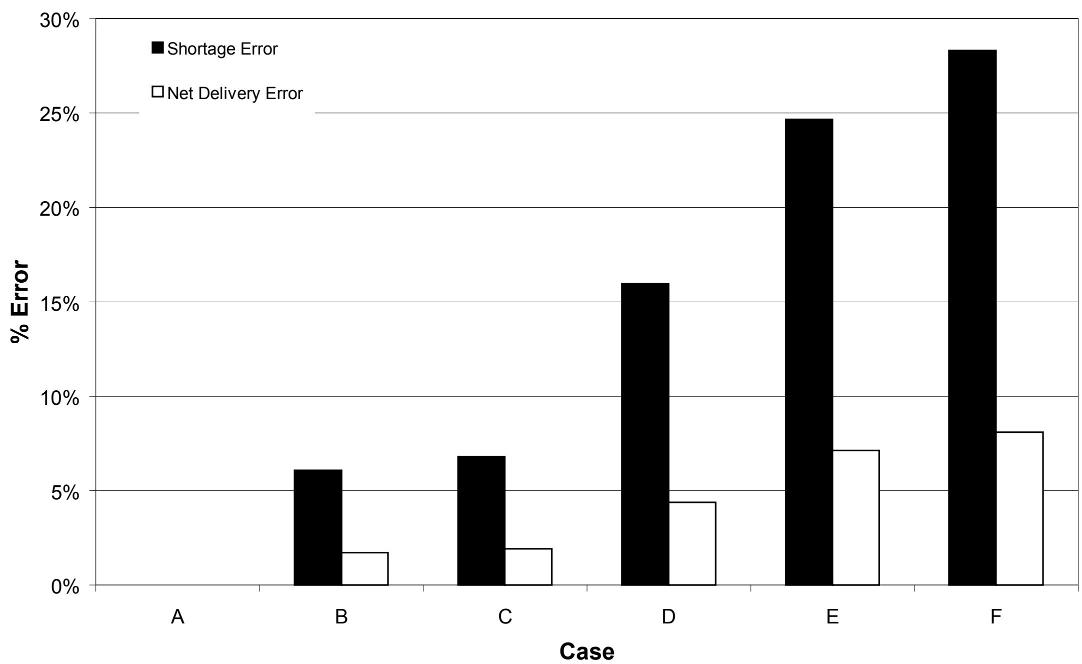
Case	Spatial Complexity	# Iterations	# Decision Variables	Interpretation Complexity
A	80	1,523,663	442.603	129
B	65	1,383,246	382,551	103
C	44	816,591	203,695	69
D	26	373,372	118,660	37
E	8	105,396	44,498	15
F	5	53,037	32,114	14

Figure 2
% Error versus Model Complexity (Relative to Case A)



good correlation between each of the complexity measures. As the values of each decrease, the percent error increases. The rate of increase in percent net delivery error does not increase linearly with the increase in complexity. The amount of error increases the most between Cases E and F but very little between Cases B and C. When measured in terms of shortage rather than net delivery, the percent errors of Cases B-F are much larger in relation to Case A. Figure 3 shows the percent errors in shortage and net delivery for each case. The percent shortage errors are less than 7% for Cases A-C but greater than 16% for Cases D-F, which indicates that a certain amount of aggregation is possible with minimal error but that greater aggregation may produce unacceptable errors. However, these measures underestimate the error of Case C because, while the other cases have less total shortage than Case A in almost every month, Case C has greater total shortage than Case A in many months. This is indicated by the differences in the average monthly net and absolute shortages for each case relative to Case A, which can be seen in Table 4. While the average net and absolute differences are very similar for the other cases, for Case C the average net difference is much lower.

Figure 3
Percent Error For Each Case

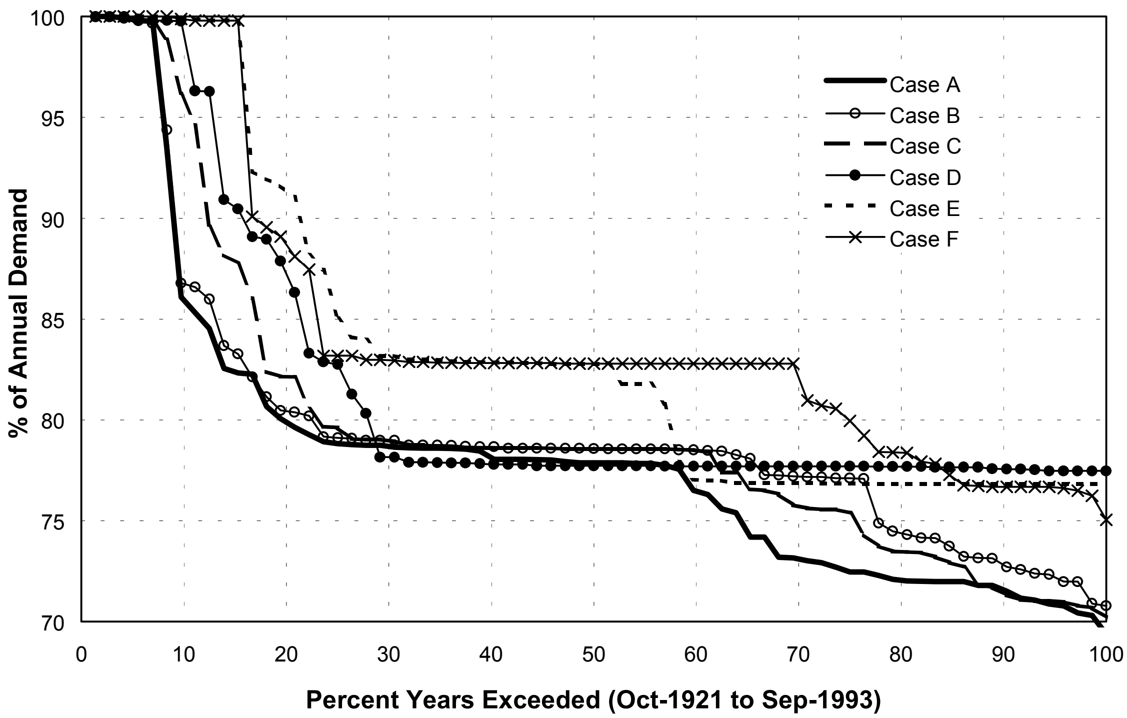


The remainder of this analysis focuses on more specific aspects of the results. First, the annual time series results will be analyzed for the entire system. Then, monthly time series analysis will be performed on the entire system for specific time periods. Finally, the differences in results for certain local regions will be analyzed.

System-Wide Analysis of Annual Time Series

Figure 4 shows the probabilities of exceedance of total annual deliveries for each case. All 6 cases show the same basic reliability, with a rapid drop in delivery to about 80-85 percent with 25 percent reliability and then a gradual decrease to about 70-75

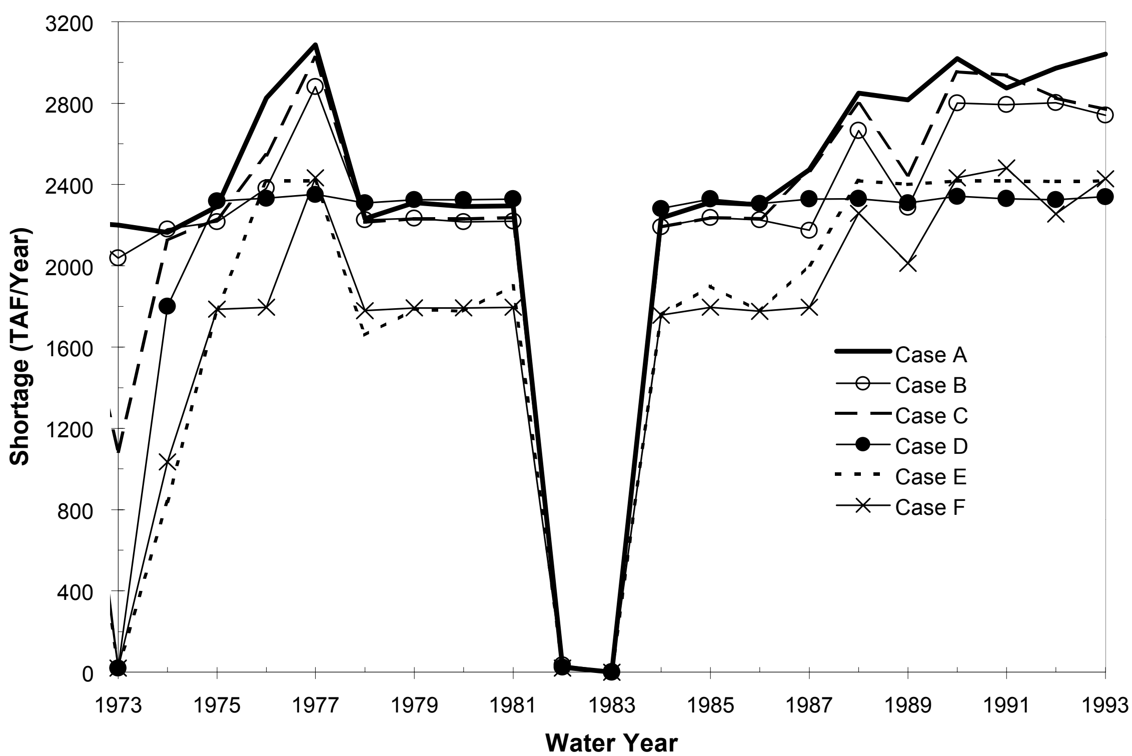
Figure 4: Total Deliveries
Annual Probability of Exceedence



percent delivery with 100 percent reliability. With a few exceptions, the curves are ordered as might be expected, with the more complex cases suggesting less reliability of delivery than the more simple cases. The greatest exception is Case B, which is less reliable than Case C for deliveries that are exceeded more than 60 percent of the time. This result seems to show that while spatial aggregation tends to allow greater delivery reliability during the years with more water available, these gains can be made at the cost of greater shortages during drier years. It may be that the aggregation of storage allows for more efficient water use during wet years but is not as much of a factor during drought years. It is unclear, however, why the model does not use the greater flexibility of the simpler formulations to alleviate the more severe droughts, in which the marginal cost of shortage would be higher. This may be caused by the problem of flat objective function surfaces in the CALVIN model, by which several different possible solutions give very similar objective function values. The spatial aggregation may cause minor deviations between the cases in individual years that may skew the shape of the reliability curves. This explanation is supported by the annual time series of shortages, which show frequent fluctuations in magnitude order between the cases from year to year. As an example,

Figure 5 shows the annual time series of shortages for each case from 1973 to 1993. This time period contains two significant droughts: 1976-77 and 1987-93. In 1974, Case B has slightly more shortage than Case A, which is followed by Cases C, D, F and finally E. By 1976, however, Case B has fewer shortages than Case A, C, or E. Between 1988 and 1989, both Case B and C show sharp drops in shortage while the other

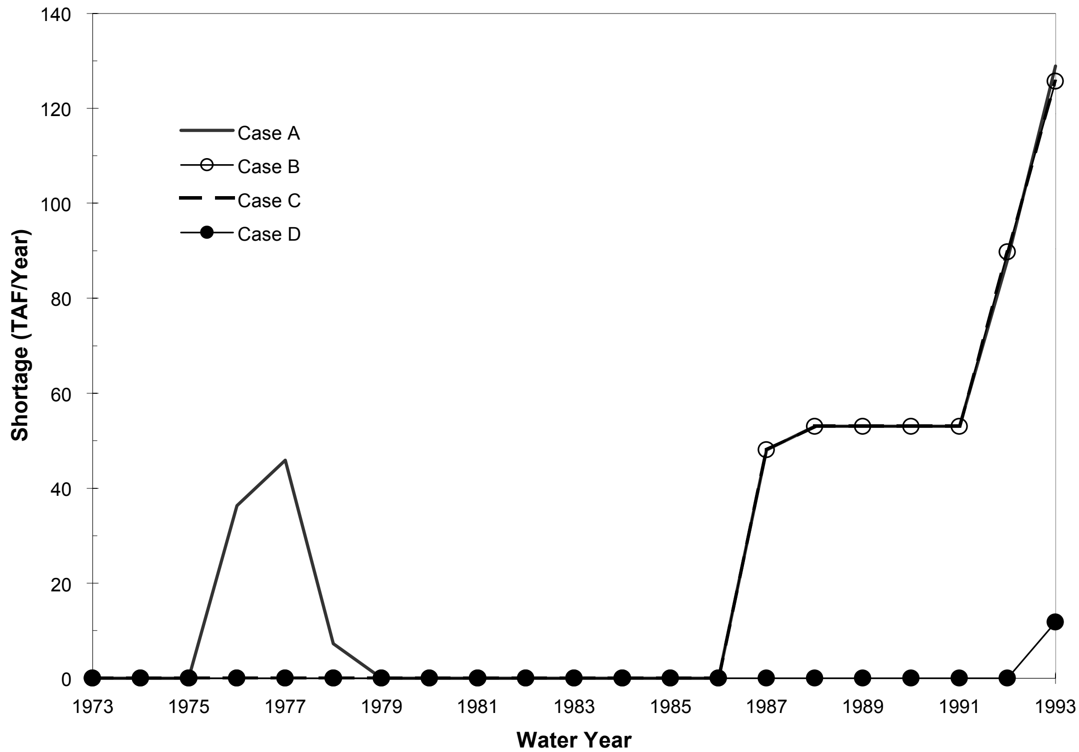
Figure 5
Total Annual Shortage



4 cases stay flat. Case D shows a unique response to both droughts in that its level of shortage rises the fastest of all 6 cases in both 1975 and 1984, to about 2,200 TAF, but then remains flat so that Case D has the lowest shortage in 1977 and from 1990-93. Fluctuations in curve order such as these make it difficult to draw definite conclusions about the differences between the models in individual years. However, the results in Figure 5 indicate that all 6 model formulations are reacting realistically to the annual changes in hydrology – all show increases in the amount of shortages during the drought years and virtually no shortages during the extremely wet years of 1982-83.

The drought years also experience the largest urban shortages, which appear in Figure 6. All of the cases which contain urban shortages (A-D) experience them during the 1987-93 drought period and Case A, which contains the largest total urban shortages, has additional shortages from 1976-78. Although the differences in urban shortage are small, they are significant because of the high value of urban demands. All of the urban shortages occurred in the EBMUD demand node, which is isolated from the rest of the system in that it can only receive water from the Pardee Reservoir on the Mokelumne River. In Cases B and C, the Pardee Reservoir is combined with the Camanche Reservoir, providing additional storage space to provide water for EBMUD. With this change, the shortage experienced from 1976-78 is eliminated. The further aggregation in Case D of the Mokelumne River with the rest of the Eastern Sacramento Valley eliminates almost all of EBMUD’s shortage, while the additional aggregation in Cases E and F eliminates the shortage altogether. Thus, there is a gradual reduction in EBMUD’s shortage as the amounts of storage and external inflows available to supply the region increase.

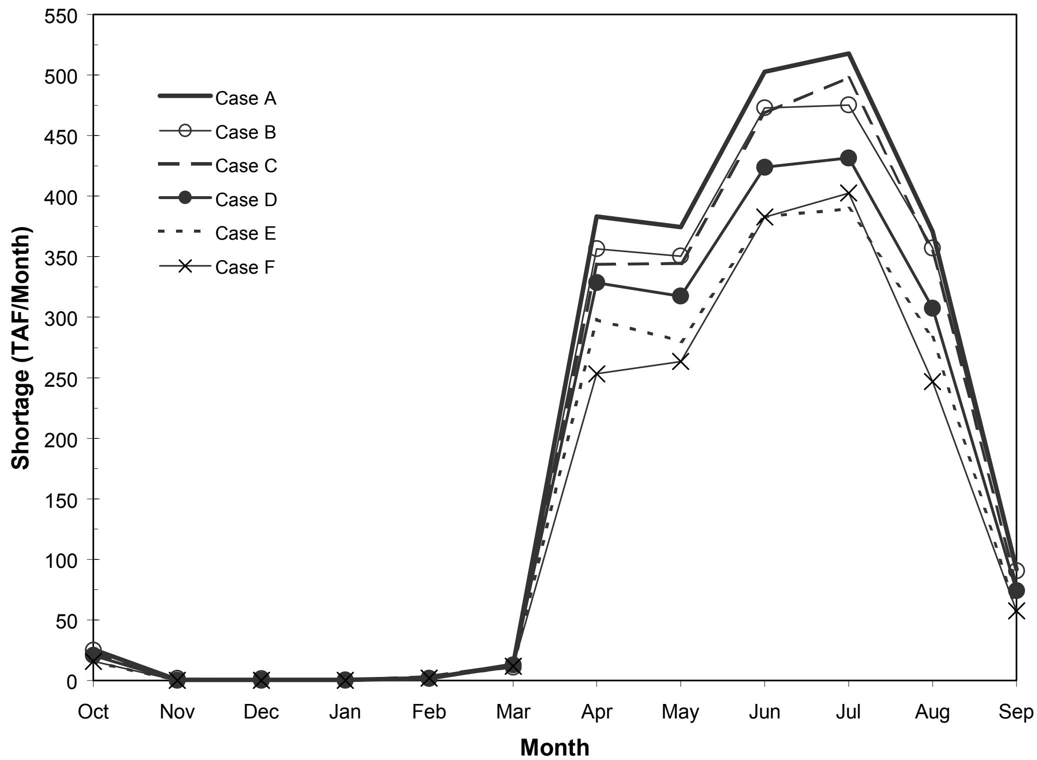
Figure 6
Total Annual Urban Shortage



System-Wide Analysis of Monthly Time Series

On an average monthly basis, there appears to be very good correlation between the model complexity and the level of deliveries. Figure 7 shows the average total monthly shortages for each case. All of the cases have little shortage from October through March, when there are very few agricultural water demands. During the summer months, all cases have higher shortages and, with a few exceptions, the more complex cases have higher shortages than the simpler cases during every month. In the next sections, the results for each case will be analyzed on a monthly basis for the 1976-77 dry period and for the 1982-83 wet period.

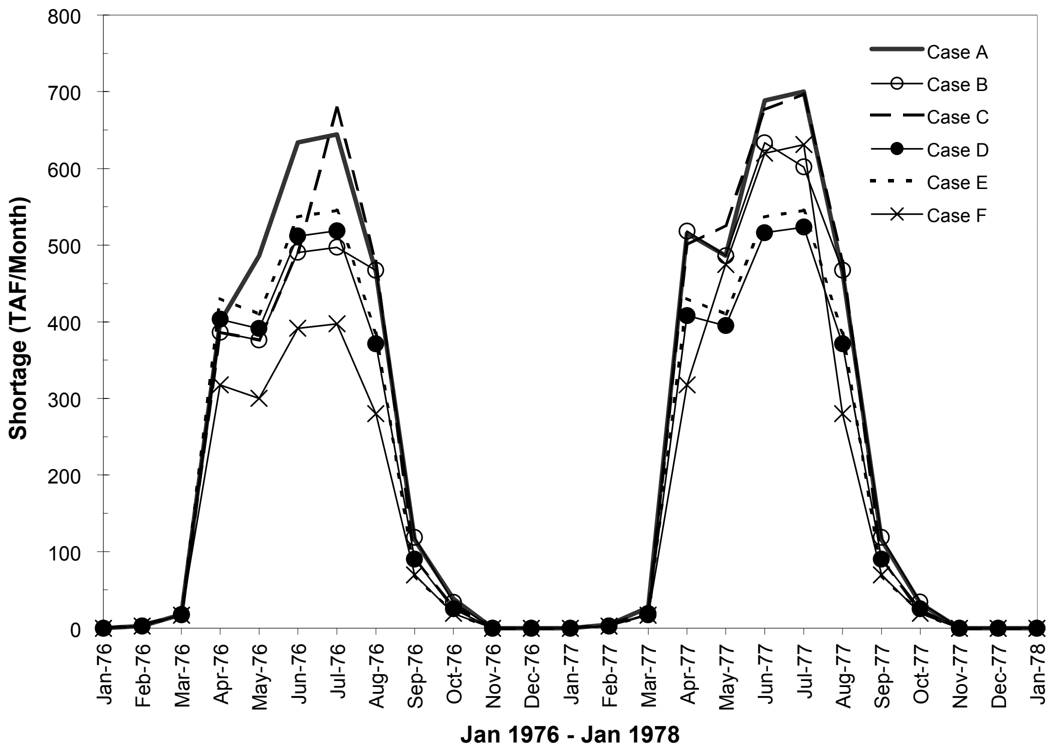
Figure 7
Average Total Monthly Shortages



1976-77 Drought

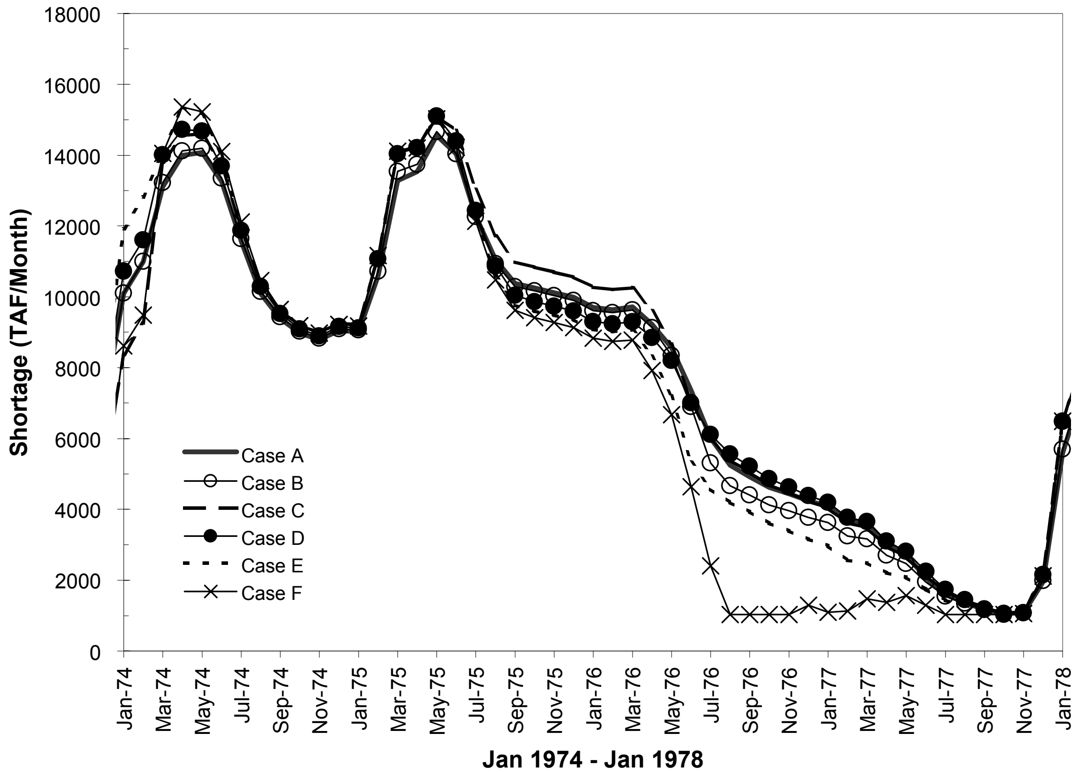
As shown in Figure 5, each case responded differently to the 1976-77 drought. While Cases A, B, C, and F experienced shortage peaks during 1976 and 1977, the shortages in Cases D and E tended to plateau at a smaller shortage rate but maintained that level for a longer time period. In addition, although Case B had larger overall shortages, Case C tended to have larger annual shortages during the drought years. These trends also can be seen in the plot of monthly agricultural shortages shown for each case from January 1976 to January 1978 in Figure 8. While all 6 cases had the largest shortages during June and July of each year, Cases D and E had the smallest peaks in both years with the exception of Case B in 1976. In addition, Case C had the largest peak in 1976 and virtually the same shortages as Case A in 1977. The magnitude of the particular shortages for each case in 1976-77 may help to explain the economic differences between the cases in those years.

Figure 8
Total Agricultural Shortages



Because HEC-PRM is a deterministic optimization model, it is able to anticipate a drought and fill the reservoirs to capacity to provide maximum possible deliveries. The model can also anticipate the end of the drought and completely drain the reservoirs during the last dry year. This operation is reflected in the total monthly surface water storage curves shown for each case in Figure 9. This plot is shown beginning in January 1974 to show the first critical peak in storage at which the reservoir capacity constraints are most costly. From this figure, it can be seen that all of the cases reach a peak in March 1974, reach another peak in May 1975, and then are depleted until a minimum is

Figure 9
Total Surface Water Storage



reached in November 1977. While all the cases are able to efficiently manage the drought, the simpler cases are able to operate the storage space more efficiently because the storage and reservoir inflows are aggregated into fewer storage nodes. Thus, in March 1974, Cases E and F reach the highest peak, followed by Cases D, C, B, and A. These differences occur because in the more complex cases the storage space and reservoir inflows are divided among several reservoirs and conveyance constraints make it impossible to employ the entire storage space in every reservoir in every month. Thus, in Cases A-D, some reservoirs are at capacity in March 1974 and others are not. During the reservoir depletion period, Cases A-E seem to follow very similar storage paths while Case F depletes its surface water storage more rapidly. This difference is probably caused by the increased flexibility of groundwater storage in Case F; it is the only case in which groundwater can be used to supply any demand in any month. Thus, the rapid depletion of surface water storage is not as critical.

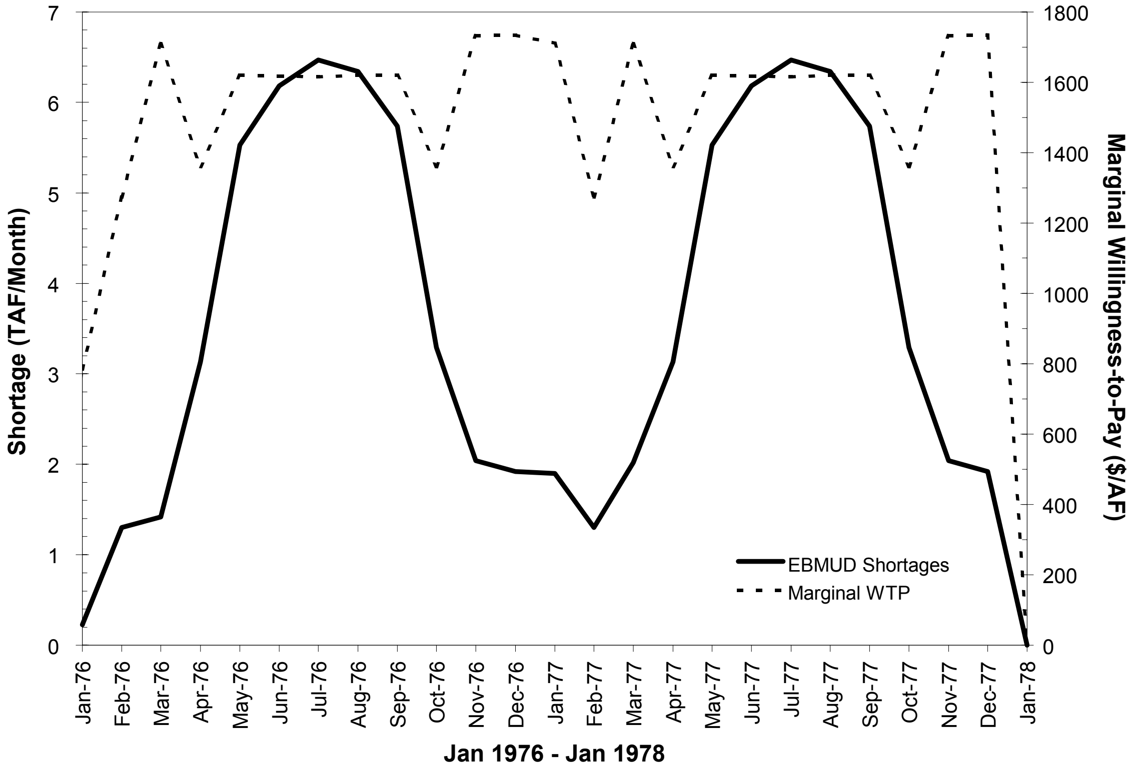
Table 6. Economic Output (in $/AF) For Each Case

Case	Lake Shasta Storage Shadow Value (Mar 1974)	Required Delta Outflow Shadow Value (Apr 1974)	Sac West Refuges Shadow Value (Apr 1974)	San Joaquin River Marginal Value (Apr 1974)	Agricultural Marginal Willingness-to-Pay (1976-77)
A	133.5	133.9	116.0	133.9	138.5
B	132.6	133.0	116.0	133.0	113.6
C	138.5	138.9	117.8	138.9	146.7
D	120.2	120.7	101.0	120.7	114.8
E	49.6	49.7	42.0	49.7	103.9
F	104.2	104.5	88.3	104.5	103.9

During March 1974, certain reservoirs in all 6 cases have a very high shadow value on the reservoir capacity constraint. This shadow value reflects the cost of water shortage during the coming drought. The reservoir containing Shasta Lake (SR-4) is at capacity in every case, and the shadow values in March 1974 for this reservoir can be seen for each case in Table 6. In addition, Table 6 shows the shadow values on the Required Delta Outflow and for inflows into the Sacramento West Refuges, and the marginal values of additional inflow from the San Joaquin River in April 1974. All of these economic values pertain to flow values that are represented individually in all 6 cases, with the exception of the Sacramento West Refuges, which are combined with the Sacramento East Refuges in Cases E and F. Each of the economic values are the highest in Case C, followed by Cases A, B, F, D, and E. The low values shown for Case E are caused by the unusually low shortages of Case E in 1974 (see Figure 5). While the other cases have agricultural marginal willingness-to-pay values in 1974 comparable to those in 1976-77, the maximum agricultural marginal willingness-to-pay for Case E is only $43.2/AF in 1974. With the exception of the Sacramento West Refuges, all of these values are very similar in each case. The increased economic values for Cases C and F relative to the other cases may be due to the unusually high shortages they experience during 1976 and 1977. The economic values for Case F may be higher than those for Case E because of the absence of the urban pumping cost to EBMUD and CCWD. The shadow value on the Required Delta Outflow exactly equals the marginal value of additional inflow from the San Joaquin River for all months in all cases, which reflects the absence of any costs or losses on water flowing from the San Joaquin River out the Delta. The Sacramento West Refuges shadow value is $8-22/AF less than the Storage shadow value in every case. This difference occurs because the refuge is typically in parallel with the agricultural demand region while the reservoir and the Delta outflow are in series with both. Thus, if the agricultural delivery were increased by one acre-foot because of reduced refuge demand, all of the return flow from that one acre-foot would still be needed to supply the Delta outflow. However, if the Delta outflow requirement were reduced by one acre-foot, all of the return flow from that one acre-foot would be available for reuse, thereby increasing the value of that unit of water.

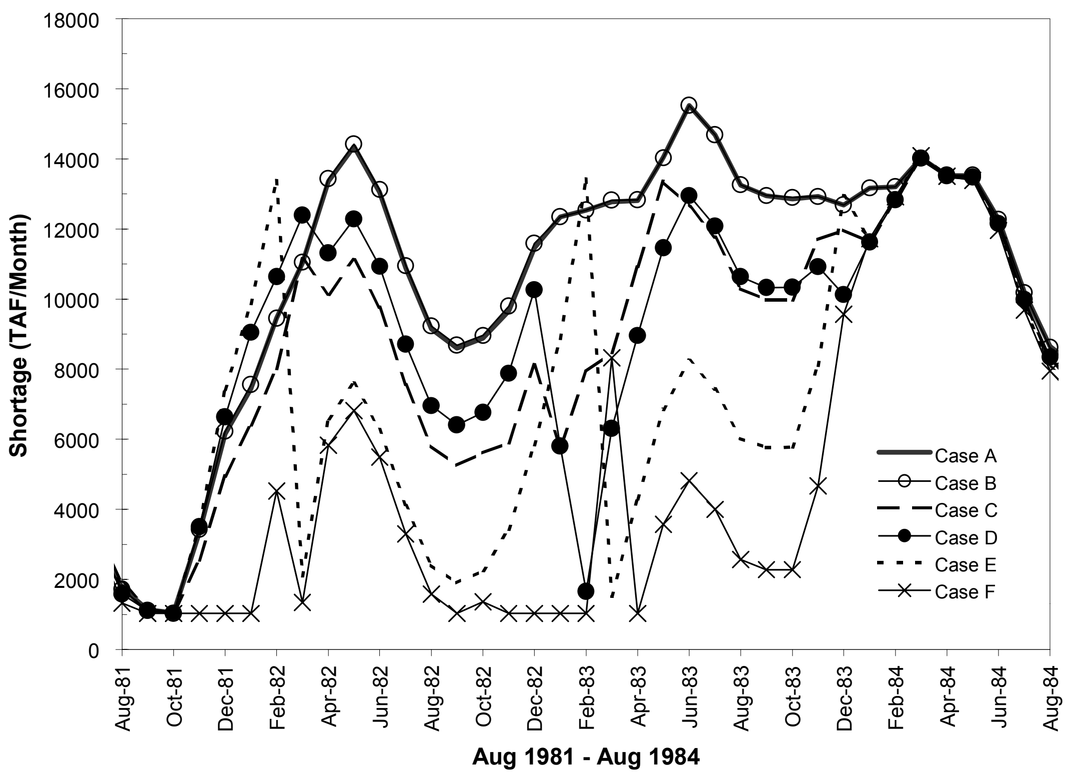
All of the marginal and shadow values seem to be related to the marginal willingness-to-pay values for the agricultural regions. The marginal willingness-to-pay is defined as the amount that the demand regions would be willing to pay to receive one additional acre-foot of water in a given month. Table 6 shows the largest marginal willingness-to-pay value for any agricultural region in the Western Sacramento Valley

Figure 10
EBMUD Shortages



for every case during 1976 or 1977. These values are shown primarily to demonstrate that they are of the same magnitude as the marginal values and shadow values. It is difficult to draw definite conclusions from a comparison of the marginal willingness-to-pay across different cases because for the more complex cases they represent the maximum value of many different demand regions, each of which has a different marginal willingness-to pay value. The marginal-willingness-to-pay is much smaller for agricultural regions than for urban regions. The EBMUD demand node in Case A was the only urban demand node in any case to have any shortage during the 1976-77 drought. Figure 10 shows the monthly shortages and the marginal willingness-to-pay for EBMUD from January 1976 to January 1978. While the magnitude of shortage is small, the marginal willingness-to-pay values are very large – they fluctuate around $1,500/AF during the drought period, which is much larger than the marginal willingness-to-pay values for the agricultural regions. The marginal willingness-to-pay for EBMUD does not affect the other economic values in Case A because the EBMUD system is isolated from the rest of the system and any additional water that may be available could not be conveyed to the EBMUD demand node.

Figure 11
Total Surface Water Storage



1982-83 Flood

Because they were extremely wet years, no case contains any shortages during 1982 or 1983. As in the drought years, the results during these wet years were influenced by the perfect foresight of the deterministic optimization. Figure 11 shows the total surface water storage values for Cases A-F from August 1981 through August 1984. In

October 1981, the total surface water storage is at the minimum in all 6 cases. In March 1984, all six cases store the maximum surface water storage possible (the maximum in March is less than in May or June because of changes in the amount of storage reserved for flood space). However, the storage paths between these dates vary widely. Cases A and B have nearly identical storage paths. During most months, the storages in Cases A and B are higher than that in the other cases. This is most likely caused by the disaggregation of the reservoirs, which makes it more difficult to refill all of the reservoirs in the later months. The storage paths of Cases C and D are both fairly similar to those of Cases A and B, while Cases E and F deviate widely. As during the drought period, Case F seems to maintain much lower shortages than do the other cases for many months. Again, this may be caused by the increased flexibility of groundwater storage in Case F. The availability of groundwater in later months makes the maintenance of surface water storage less critical.

Table 7. Average Annual Surplus Delta Outflows

Case	Surplus Delta Outflow (TAF/year)
A	2837.5
B	2773.8
C	2448.8
D	2277.2
E	2159.0
F	2072.1

The excess water present in 1982-83 is reflected in the very large surplus Delta outflows during those years. The Delta outflow is the only location in the model through which surplus flows can be removed from the system. The average annual surplus Delta outflows can be seen in Table 7. Over the entire 72-year period, the formulations with the larger complexities have larger surplus Delta outflows, which reflect the greater ability of the aggregated reservoirs to capture flood flows. Cases E and F, which both aggregate all of their surface water storage into a single reservoir, have far fewer surplus Delta outflows than Cases A-D.

Figure 12
Surplus Delta Outflow

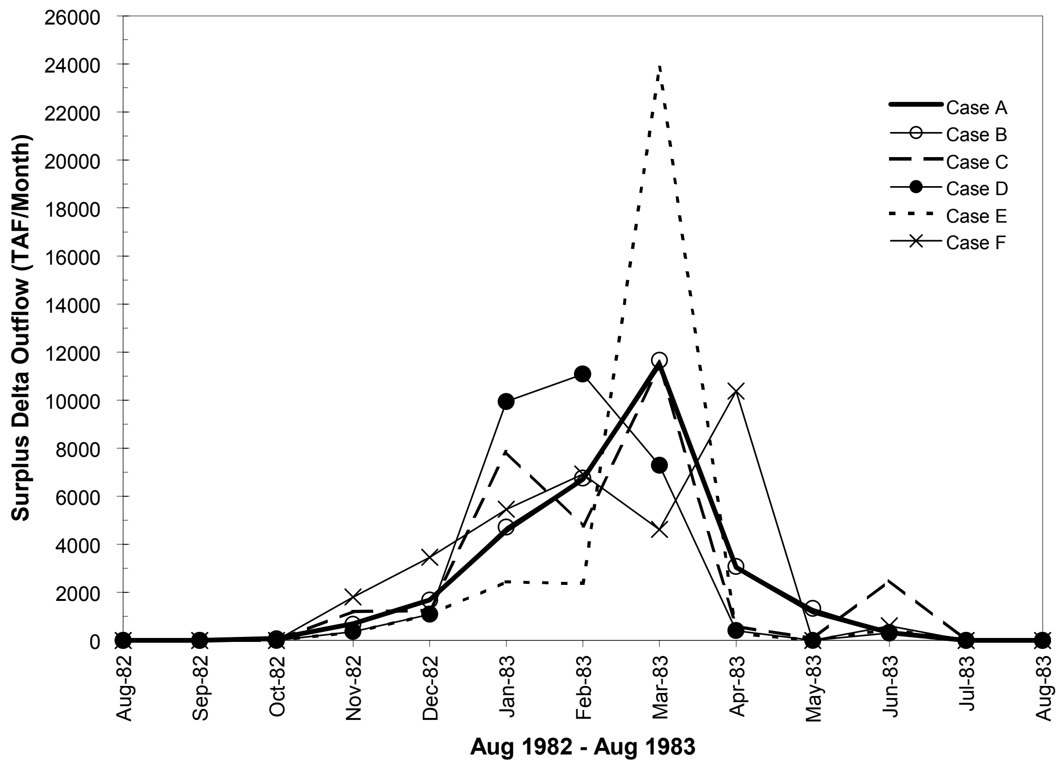


Figure 12 shows the surplus Delta outflows for all 6 cases from September 1982 through August 1983. As demonstrated earlier, the very large surplus Delta outflows in Case E during March 1983 are caused by a very small difference in the optimal objective function value. Evan a small persuasion penalty reduced the peak to 11 MAF with very little change in the other results. As with the storage curves, Cases A and B have nearly identical surplus Delta outflows and the largest peak flows of any of the cases. Case C seems to follow the paths of Cases A and B fairly well, but does not reach the same peak flow. Case D also seems to follow a similar path as Cases A and B, but reaches a peak flow two months earlier. Cases E and F seem to deviate significantly from Cases A and B. Thus, there seems to be a gradual progression away from the Case A curve as the complexity is decreased.

Analysis of Results From Local Regions

Thus far, all of the results presented have pertained either to the system as a whole or to those parts of the system which are modeled identically by all 6 cases. In this section, the results are examined for selected local regions to determine what differences, if any, have resulted from the aggregation of system components. In particular, the following components will be examined:

- Groundwater storage in the North Sacramento Valley (GW-1) and the Delta (GW-9)
- Agricultural deliveries in the Western Sacramento Valley (CVPM 1-4)
- Average reservoir shadow values

<u>Groundwater Storage</u>

The average annual groundwater mining was listed for each case in Table 3. Case A had an average of 42.4 TAF/year of groundwater mining, Cases B-D had an average of 34.2 TAF/year, and Cases E and F had no groundwater mining. All groundwater mining calculations are based upon the initial and ending groundwater storages for each groundwater basin. Months in which the groundwater level is raised count as negative groundwater mining over the entire 72-year period. The model is free to have any ending storage, but must pay a very large cost for every acre-foot less storage in any groundwater basin than that basin's initial storage. Therefore, groundwater mining will only occur if its constrained outflows are greater than its inflows over the entire 72-year period. The constrained groundwater basin outflows include pumping to local urban demands, infiltration losses, and losses to other groundwater basins. If the sum of these values exceeds the sum of the inflows, which include natural recharge, return flows and gains from other basins, groundwater mining will occur. While this will not normally be the case, groundwater mining was found in 2 groundwater basins in Case A because of the artificial decrease in groundwater inflows into the system. This decrease caused the inflow in these basins to be less than the outflow.

Of the 42.4 TAF/year of groundwater mining in Case A, 34.2 TAF/year occurred in GW-9 and 8.2 TAF/year occurred in GW-1. These results only persisted in the other cases if these groundwater basins continued to be disaggregated from other groundwater basins. GW-1 is aggregated with GW-2 in Case B and further aggregated in the other cases. The combined GW-1 and GW-2 groundwater basin in Case B has no groundwater mining, which indicates that there is enough excess inflow in GW-2 to make up for the water deficit in GW-1. Similarly, the GW-9 groundwater mining persists in Cases A-D when this region is left disaggregated but is eliminated when all groundwater is aggregated together in Cases E and F. Although these examples result from an artificial change made to the groundwater hydrology, they could be significant if found in a model intended to be accurate. They indicate that the implementation of a more aggregated model formulation can not only cause a lack of precision about the location of certain results within aggregated regions but also eliminate important results that could be derived from a more complex formulation.

Agricultural Deliveries

This analysis explores the effects of aggregating four agricultural regions in the West Sacramento Valley (CVPM 1-4) and focuses on Cases A-C. In Case A, all four regions are represented separately. In Case B, CVPM 1 and CVPM 2 are combined, as are CVPM 3 and CVPM 4. In Case C, all four agricultural regions are combined into a single demand region. The average percent deliveries for each grouping of CVPM regions in each case are shown in Table 8.

Table 8. Percentage of Total Agricultural Demand by CVPM Region for Cases A-C

Case	1-4	1-2	1	2	3-4	3	4
A	69.1%	77.8%	92.3%	69.9%	67.4%	44.4%	82.6%
B	72.3%	75.8%	n/a	n/a	72.2%	n/a	n/a
C	76.1%	n/a	n/a	n/a	n/a	n/a	n/a

These results indicate that the reliability of demand delivery generally increases as the demand regions are aggregated. The only exception is in CVPM 1-2, in which a higher percentage of demand is satisfied in Case A than in Case B. This is caused by the occurrence of groundwater mining in GW-1. Although GW-1 is compelled to not pump any groundwater to the agricultural region, the model attempts to maximize the agricultural return flow to the groundwater basin by maximizing the amount of surface water deliveries. Therefore, CVPM-1 has a much higher delivery than CVPM-2 and the total percent delivery is greater than in Case B. CVPM Regions 3 and 4 are similar in that CVPM 3 has much less reliable deliveries than the combined CVPM 3-4 demand in Case B while CVPM 4 has much more reliable deliveries. This result occurs because CVPM 4 is downstream of CVPM 3 and can therefore implement its surface water return flows. While Case B allows for the use of return flows within this region, no distinction is made between water used by either region. The much more reliable supply of CVPM 4 is therefore not indicated in the aggregated demand node of Case B. These two examples of differences between Cases A and B are indicative of the kind of local information that can potentially be absent in a more simplified model run.

Table 9. Storage Shadow Values

Case	Reservoir	Average Shadow Value ($/AF)
A	Englebright Lake	9.11
B	Black Butte Lake	5.68
C	Pardee/Camanche Reservoir	4.01
D	EBMUD Local Storage	3.41
E	Total Aggregated Reservoir	0.99
F	Total Aggregated Reservoir	1.12

Reservoir Shadow Values

When two or more reservoirs are aggregated together, the probability that the aggregated reservoir will have a storage equal to its maximum or minimum storage is less than the probabilities of any of the reservoirs represented individually. The reduced pressure on reservoir storage is reflected in the average storage shadow values. Table 9 shows the reservoir in each case with the largest average shadow value on its upper-bound constraint. In Case A, Lake Englebright has the highest average shadow value. However, in Case B Lake Englebright is aggregated with New Bullards Bar and the aggregated reservoir has an average shadow value of only $3.51/AF, much less than that of Black Butte Lake. In Case C, Black Butte Lake is aggregated with other reservoirs and therefore the Pardee/Camanche aggregated reservoir has the highest average shadow value. In Case D the Pardee/Camanche reservoir is aggregated with additional reservoirs and the EBMUD local storage has the highest shadow value. The EBMUD local storage is one of two disaggregated reservoirs in Case D – the two aggregated reservoirs have average shadow values of $0.98/AF and $1.50/AF, which are very similar to the average shadow values for the aggregated surface water reservoirs of Cases E and F. All of the reservoirs with the highest shadow values in Cases A-D are relatively small reservoirs. Englebright Lake has a capacity of 66 TAF, Black Butte Lake has a capacity of 150 TAF, the Pardee/Camanche reservoir has a capacity of 641 TAF, and the EBMUD local storage has a capacity of 153 TAF. It seems that the larger the aggregated storage capacity, the less likely the maximum constraint will be binding and the smaller the average shadow value will be. Because aggregated reservoirs tend to have relatively large capacities, they are likely to have smaller average shadow values. Thus, reservoir aggregation can have a significant effect on the valuation of surface water storage. For capacity expansion location and valuation purposes, more aggregate system representations are likely to reduce the estimated values of new capacity and change the locations of the most preferred expansion locations.

CONCLUSIONS

This study yields insight into the effects of spatial complexity on the results of a network flow optimization model. If the most complex formulation is assumed to be the most accurate, each level of simplification introduces some additional error to the results. The magnitude of this error, however, differs depending upon how one looks at the results. Although the net delivery and shortage for each case are inter-dependent, the degree of error appears much greater if viewed in terms of shortage rather than net delivery. Case F, for example, has an 8.1% error in net delivery but a 28.3% error in shortage. While 8.1% may be considered an acceptable margin of error given the magnitude of simplification, 28.3% clearly cannot. However, the results for Cases B and C demonstrate that a certain degree of simplification is possible with minimal error.

Even with greater spatial aggregation it may still be possible to develop a reasonable and useful model. Because the levels of both agricultural and urban shortages are reduced as the amount of aggregation is increased, modelers can compensate when developing a simpler model by making parameter and input assumptions that increase the demand, reduce aggregated reservoir storage capacities, or reduce the amount of water available. Perhaps, the reuse within aggregated demand areas can be neglected. In addition, in many cases the magnitude of error may not be a major concern because the study in question is concerned with the differences in results between two or more alternatives, each of which will be equally affected by the errors caused by spatial aggregation.

Choosing an appropriate level of complexity might be easier if more studies were conducted to broaden understanding of the effects of model complexity on results. This study is just the tip of the iceberg of what could be studied concerning model complexity, even for reservoir models. Such studies could be expanded to explore the effects of a proposed policy alternative at each level of aggregation. In addition, any of the types of model complexity listed in Table 1 could be examined for their relative effects, and these studies could be expanded to test the effects on many different systems. The next logical step may be to explore the effects of temporal complexity by running the model at different time steps.

APPENDIX 1: REFERENCES

Archibald T.W., McKinnon, K.I., and Thomas, L.C. (1997) "An Aggregate Stochastic Dynamic Programming Model of Multireservoir Systems." *Water Resources Research*, 33(2), 333-340.

BDMF (2000) *Protocols for Water and Environmental Modeling.* Bay Delta Modeling Forum, Richmond, CA, January 21, 2000. (www.sfei.org/modelingforum/)

Bloschl, G. and Kirnbauer, R. (1991) "Point Snowmelt Models with Different Degrees of Complexity – Internal Processes." *Journal of Hydrology*, 129, 127-147.

Brooks, R.J. and Tobias, A.M. (1996) "Choosing the Best Model: Level of Detail, Complexity, and Model Performance." *Mathematical and Computer Modeling*, 24(4), 1-14.

Costanza, R. and Sklar, F.H. (1984) "Articulation, Accuracy and Effectiveness of Mathematical Models: A Review of Freshwater Wetland Applications." *Ecological Modeling*, 27, 45-68.

Fontaine, T.A. (1995) "Rainfall-Runoff Model Accuracy for an Extreme Flood." *Journal of Hydraulic Engineering-ASCE*, 121(4), 365-374.

Howitt, R.E., Lund, J.R., Kirby, K.W., Jenkins, M.W., Draper, A.J., Grimes, P.M., Davis, M.D., Ward, K.B., Newlin, B.D., Van Lienden, B.J., Cordua, J.L., and Msangi, S.M. (1999) *Integrated Economic-Engineering Analysis of California's Future Water Supply*. State of California Resources Agency, Sacramento, CA. (http://cee.engr.ucdavis.edu/faculty/lund/CALVIN/)

Jackson, B.B. (1975) "The Use of Streamflow Models in Planning." *Water Resources Research*, 11(1), 54-63.

Jakeman, A.J. and Hornberger, G.M. (1993) "How Much Complexity is Warranted in a Rainfall-Runoff Model?" *Water Resources Research,* 29(8), 2637-2649.

Keskin, E.M. and Agiralioglu, N. (1997) "A Simplified Dynamic Model for Flood Routing in Rectangular Channels." *Journal of Hydrology*, 202, 302-314.

Le Roux , J.A. and Potgieter, M.S. (1991) "The Simulation of Forbush Decreases with Time-dependent Cosmic-ray Modulation Models of Varying Complexity." *Astronomy and Astrophysics*, 243, 531-545.

Loague K.M. and Freeze R.A. (1985) "A Comparison of Rainfall-Runoff Modeling Techniques on Small Upland Catchments." *Water Resources Research,* 21(2), 229-248.

Nihoul, J.C. (1994) "Do Not Use a Simple Model When a Complex One Will Do." *Journal of Marine Systems*, 5, 401-410.

Palmer, R.N. and Cohan, J.L. (1986) "Complexity in Columbia River Systems Modeling." *Journal of Water Resources Planning and Management-ASCE*, 112(4), 453-468.

Palsson, B.O. and Lee, I. (1993) "Model Complexity Has a Significant Effect on the Numerical Value and Interpretation of Metabolical Sensitivity Coefficients." *Journal of Theoretical Biology*, 161, 299-315.

Ponce, V.M., Li, R.M., and Simons, D.B. (1978) "Applicability of Kinematic and Diffusion Models." *Journal of the Hydraulics Division-ASCE*, HY3, 353-360.

Rogers, P.P., and Fiering, M.B. (1986) "Use of Systems Analysis in Water Management." *Water Resources Research*, 22(9), 146-158.

Saad M., Bigras, P., Turgeon, A., and Duquette, R. (1996) "Fuzzy Learning Decomposition for the Scehduling of Hydroelectric Power Systems." *Water Resources Research*, 32(1), 179-186.

Sinha, J., Eswaran, V., and Bhallamudi, S.M. (1995) "Comparison of Spectral and Finite-Difference Methods for Flood Routing." *Journal of Hydraulic Engineering-ASCE*, 121(2), 108-117.

Smith, D.E. and Starkey, J.M. (1995) "Effects of Model Complexity on the Performance of Automated Vehicle Steering Controllers: Model Development, Validation and Comparison." *Vehicle Systems Design*, 24, 163-181.

Stockle, C.O. (1992) "Canopy Photosynthesis and Transpiration Estimates Using Radiation Interception Models with Different Levels of Detail." *Ecological Modeling*, 60(1), 31-44.

Turgeon, A. (1981) "A Decomposition Method for the Long-Term Scheduling of Reservoirs in Series." *Water Resources Research*, 17(6), 1565-1570.

Turgeon, A. and Charbonneau, R. (1998) "An Aggregation-Disaggregation Approach to Long-Term Reservoir Management." *Water Resources Research*, 34(12), 3585-3594.

USACE (1994) *Hydrologic Engineering Center's Prescriptive Reservoir Model, Program Description*. U.S. Army Corps of Engineers, Hydrologic Engineering Center, Davis, CA, February 1994.

Warwick, J.J. (1989) "Interplay Between Parameter Uncertainty and Model Aggregation Behavior." *Water Resources Bulletin*, 25(2), 275-283.

Warwick, J.J. and Cale, W.G. (1987) "Determining the Likelihood of Obtaining a Reliable Model." *Journal of Environmental Engineering-ASCE*, 113(5), 1102-1119.

Wood, E.F. (1995) "Scaling Behaviour of Hydrological Fluxes and Variables: Empirical Studies Using a Hydrological Model and Remote Sensing Data." *Hydrological Processes*, 9, 331-346.

Wood, E.F., Sivapalan, M., Beven, K., and Band, L. (1988) "Effects of Spatial Variability and Scale with Implications to Hydrologic Modeling." *Journal of Hydrology*, 102, 29-47.

Wood, S.C. (1989) "Arguments for Constructively Simple Models." *Journal of the Operational Research Society*, 40(2), 141-153.

Woods, R., Sivapalan, M., and Duncan, M. (1995) "Investigating the Representative Elementary Area Concept: An Approach Based on Field Data." *Hydrological Processes*, 9, 291-312.

APPENDIX 2: TEST CASE SCHEMATICS

Figure 13. Case D Schematic

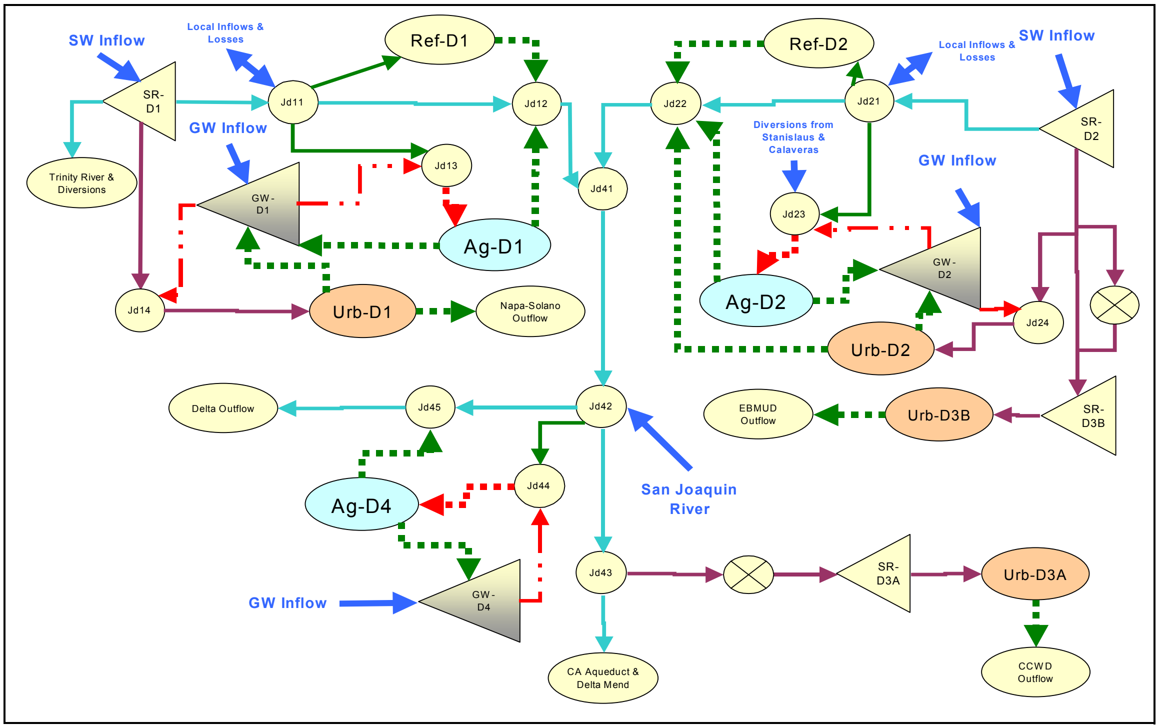


Figure 14. Case E Schematic

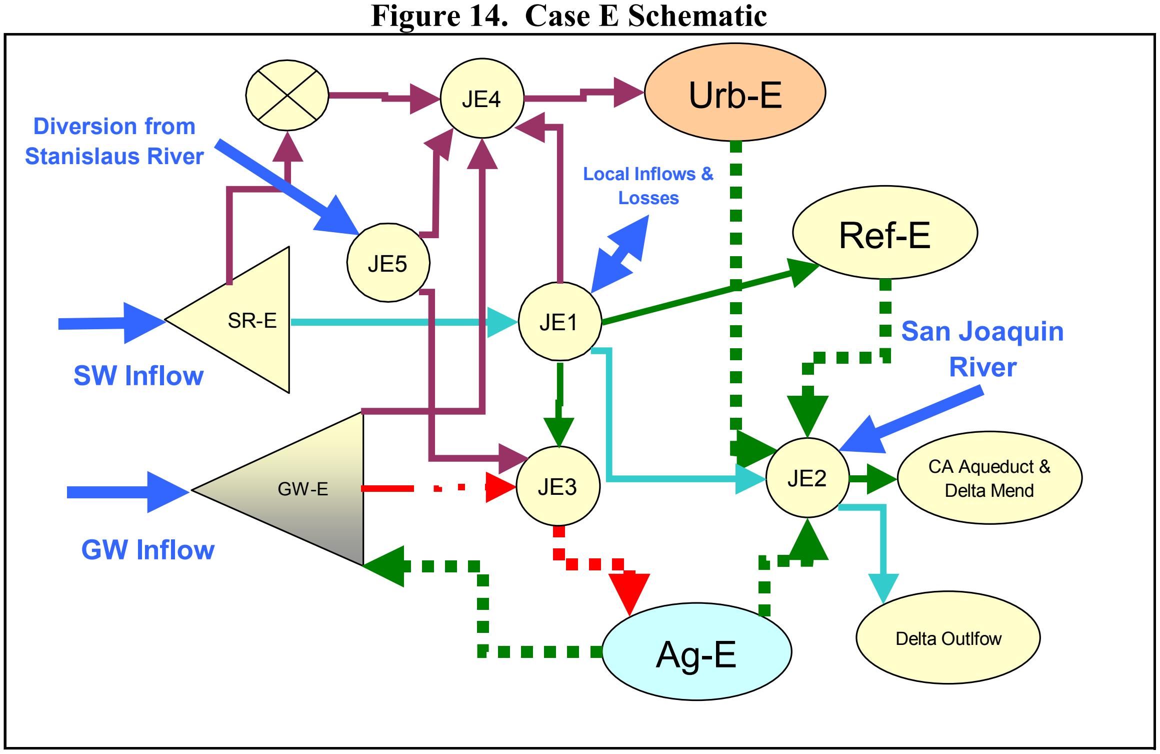


Figure 15. Case F Schematic