121: Intersite Comparisons of Rainfall-runoff Processes

JULIA JONES

Department of Geosciences, Oregon State University, Corvallis, OR, US

This article argues that despite the limitations of rainfall-runoff data, there are compelling reasons for hydrologists to conduct many more intersite comparisons of rainfall-runoff data. Inferences about hydrologic processes are drawn from an unnecessarily narrow subset of temporal scales, spatial scales, and geographic conditions given the range of data available. In intersite comparison of rainfall-runoff data, we face the same challenge as hydrologic modelers, namely: how can we discriminate among alternative mechanistic explanations for any given rainfall-runoff (or runoff-runoff) dataset? This article (1) provides a justification for additional intersite comparisons given the history and lessons learned from rainfall-runoff and related studies, (2) demonstrates how intersite comparison helps discriminate among alternative hydrologic mechanisms, and (3) outlines steps involved in conducting intersite comparative analysis with a particular focus on analysis of primary data. The article argues that the best approach to intersite hydrology is (re)analysis of original data, which allows us to (1) expand the sample size to include data not yet analyzed and replicate hypothesis tests; (2) ask new questions, even of old data, posed by the current context for hydrology; (3) make new comparisons among records not formerly juxtaposed; (4) use novel statistical approaches to reveal hitherto obscure features of the data, and (5) use ancillary data to refine hypothesis tests about hydrologic mechanisms. The article provides suggestions for the steps involved in intersite comparison including (!) identifying a question, (2) developing a study design, (3) selecting, accessing, and merging datasets, and (4) choosing statistics for comparison of rainfall and runoff data. The endeavor of intersite comparison of rainfall and runoff data is analogous and complementary to the parallel quest for methods of parameter identification and model structure selection in hydrologic modeling.

OBJECTIVE AND SCOPE OF THIS ARTICLE

Of the tools available to hydrologic science – process studies, modeling, and analysis of rainfall-runoff data – data analysis has taken a far back seat over the past few decades. Reasons for this neglect are varied, but many hydrologists are deeply disparaging of the potential for new insights from rainfall-runoff data. A common view is that it is impossible to gain insights about hydrologic mechanisms from rainfall-runoff data because they are inevitably limited to a black-box form of analysis.

This study takes the opposite view, namely, that a great deal can be learned from intersite comparison of rainfall-runoff data. Our inferences about hydrologic processes are drawn from an unnecessarily narrow subset of temporal scales, spatial scales, and geographic conditions, given the range of data available. A distressingly large number of publications simply repeat or challenge the results of a few primary analyses, without exploring unexamined datasets.

In analysis of rainfall-runoff data, we face the same problem as that faced by hydrologic modelers. Hydrologic modelers have great difficulty in identifying the correct parameters from among the very large set of parameters that can optimize any given rainfall and/or runoff dataset applied to a given model structure (Beven, 2000). Moreover, hydrologic modelers must find some means of discriminating among alternative model structures, other than optimizing models to fit rainfall and/or runoff datasets. In intersite comparison of rainfall-runoff data, we face the same challenge, namely: **how can we discriminate among alternative**

Encyclopedia of Hydrological Sciences. Edited by M G Anderson.

mechanistic explanations for any given rainfall-runoff (or runoff–runoff) dataset?

The creative analysis of multiple rainfall-runoff records can help test hypotheses about hydrological processes. This optimism is based on the following observations:

1. *A great many rainfall-runoff datasets have never been examined.* Even in datasets that have been analyzed, there exist time periods and temporal resolutions that have never been examined. Because many bitter controversies in interpretation of hydrologic processes hinge upon hypotheses tested using very small samples, replicating these hypothesis tests with enlarged samples of basins and time periods is an obvious first step toward constructing hydrologic theory from analysis of observed data.
2. *New questions can be asked of rainfall-runoff records*, even those that have been analyzed to answer old questions. Old questions that have dominated the analysis of rainfall-runoff data include (i) (forest harvest) treatment effects, (ii) tests of unit hydrograph, and (iii) flood and water yield forecasting. New questions include tests for (i) trends related to climate and/or vegetation and (ii) scaling of rainfall and runoff in space and time. Prediction of runoff in ungaged basins is an ongoing question.
3. *New comparisons may be made among records not previously compared.* The most common intersite comparison has been of runoff in small, treated-control basin pairs within a site. However, many opportunities exist for other kinds of comparisons. For example, (i) two or more control basins can be compared to reveal the influence of some factor that differs between them, (ii) a small basin may be compared to a larger basin that contains it to reveal the effect of spatial scale and routing, or (iii) two or more large basins may be compared to reveal the influences of factors that differ between them. It may even be possible to compare (iv) two basins in very different, distantly spaced sites, to infer the progression of some external forcing, such as climate change.
4. *New forms of statistical analyses may reveal new aspects of rainfall-runoff data.* Intersite comparisons of rainfall-runoff data have been limited to a few statistical methods, which provide limited insights into the patterns of rainfall and runoff. Linear regression, such as of runoff at a treated versus a control basin, has dominated analyses in the past. However, many other statistical techniques such as analysis of variance, autocorrelation methods, cross-correlations, and even spectral analysis or wavelets, can reveal hitherto unknown aspects of rainfall-runoff relationships among sites.
5. *Ancillary datasets are now available to reveal processes operating inside instrumented basins.* With improved computer technology and Internet access, rainfall and runoff data can be readily accessed and merged with rainfall-runoff data. Examples of ancillary data include historical climate data, digital elevation data, maps of geology, soils, and vegetation, and process studies and experiments. In combination with greater computer power and the ready availability of multiple statistical analysis tools, this information can be used to stratify and subdivide rainfall records to more precisely test hypotheses about the hydrologic mechanisms operating in instrumented basins of all sizes and types.

Intersite comparison of rainfall-runoff data involves **analysis of primary datasets, using new data, new questions, novel comparisons, novel statistical approaches, and ancillary datasets to discriminate among alternative mechanistic hypotheses about hydrologic processes**. Records from all gaged basins are relevant, irrespective of basin size, climate type, ecosystem or vegetation type, continent, time period, or length of record. Until recently, primary analyses of rainfall-runoff data have attempted to draw inferences on the basis of the behavior of streamflow, precipitation, or runoff ratios using a single basin at one site, or a single paired basin. *Intersite hydrology comparison studies enlarge this perspective to include multiple basins within a site or across sites, or multiple paired-basin experiments within a site, or across sites.*

The objective of this article is to describe intersite analyses of rainfall-runoff processes, and evaluate their potential and limitations for contributing to basic advances in hydrology. The article consists of

- justification for additional intersite comparisons given the history and lessons learned from rainfall-runoff and related studies
- how intersite comparison helps discriminate among alternative hydrologic mechanisms, and
- an outline of steps involved in conducting intersite comparative analysis with a particular focus on analysis of primary data.

JUSTIFICATION FOR INTERSITE COMPARISON STUDIES

Analysis of primary rainfall and runoff data is by far the most practical and promising approach to intersite comparisons aimed at elucidating fundamental hydrologic processes. Analysis of primary data is necessary because to date, analyses of rainfall-runoff data have not produced accepted generalizations or theory. On the contrary, rainfall-runoff studies, as well as reviews and metaanalyzes of these studies, have left a great many questions and controversies about mechanisms that control runoff.

Some authors have approached intersite comparison by conducting literature reviews or summarizing findings (e.g. Robinson *et al.*, 2003). A somewhat more formal, but still qualitative, approach involves tabulating streamflow studies. These approaches to intersite comparison are weak because they are qualitative, and rely upon the original authors' assessments of whether findings were significant, independent of sample size or the type of test used. Such approaches also may lead to incorrect synthesis, particularly when they involve many studies with small sample sizes (e.g. short hydrologic records).

Metanalysis overcomes some of the limitations of reviews or summary tables, but it is still limited to answering the questions posed by the original authors. Metanalysis is a quantitative approach to synthesizing results from multiple studies (Gurevich and Hedges, 2001). In metanalysis, studies that share common methodologies – usually an experimental treatment – are compared, so that collective results are not confounded by differences in approaches. Metanalysis techniques estimate the average effect of some treatment across a range of studies, determine whether that average effect is significantly different from zero, and examine whether differences in effects can be attributed to certain characteristics of studies. The "average effect of the treatment" is a measure of the difference between treated and control sample units (e.g. instrumented watersheds) and may be a simple difference or a log-transformed difference (Gurevich and Hedges, 2001). These differences may be weighted to account for differences in sample sizes.

Reviews, summaries, and metanalysis of hydrologic data have focused on predicting the effects of forestry treatments on water yield, flooding, and nutrient/sediment fluxes, although many other questions could potentially be examined. For example, Hibbert (1967) examined the relationship between annual water yield responses and forestry treatments, and concluded that water yield was inversely related to forest cover, although responses to forestry treatments were highly variable and unpredictable. Bosch and Hewlett (1982) expanded Hibbert (1967) study, collecting data on the maximum increase in water yield in the first five years following reduction of forest cover from 94 paired watershed experiments around the world. They concluded that streamflow increases were weakly positively related to the proportion of forest cover removed, and the strength of the relationship between percent forest removed and annual streamflow increase was greatest for conifers, followed by hardwood and scrub vegetation. Bosch and Hewlett (1982) also noted that annual streamflow response was positively related to mean annual precipitation, at least for the sample of watersheds with conifers. The inverse relationship between forest cover and water yield from these studies gave rise to a large set of studies examining the potential for water yield augmentation from forestry treatments. These studies (e.g. Harr, 1983) revealed that short-term water yield increases were not preserved as forests regenerated, and, furthermore, streamflow surpluses often occurred in wet seasons when downstream users did not need additional water. More recent reviews, such as Stednick (1996), also have evaluated the effects of forestry treatments on water yield. More recently, metaanalyzes have examined water yield and flooding consequences of reforestation. For example, based on a metanalysis of 28 small basin studies across Europe, Robinson *et al.* (2003) concluded that reforestation was unlikely to have much effect on floods or lowflows, except where conifer plantations were established on poorly drained soils.

Many intersite comparisons in the form of reviews, summaries, or metaanalyzes have addressed the relationships of rainfall-runoff processes to nutrient and sediment. Binkley and Brown (1993) extended the metanalysis approach to examine the effect of forestry treatments on stream temperature, dissolved oxygen, nitrate, and suspended sediment based on studies in the United States. They concluded that "best management practices" were able to mitigate undesirable changes in most of these properties, except for occasional large storm event-related sediment pulses. Martin *et al.* (1984) and Hornbeck *et al.* (1997) summarized and evaluated the implications of forestry treatments for water quality with a focus on the eastern United States. Binkley *et al.* (2004) summarized the nitrogen and phosphorus concentrations of streams draining forests in the United States, but found no factors that explained the variation among sites.

Reviews, summaries, and metanalysis are an inadequate approach to intersite hydrology comparison in many instances. Reviews are limited to repeating the conclusions drawn by the authors, which may be limited or biased, depending on the experimental design, sample size, and analysis. Frequently, experiments combined in a review or summary have small sample sizes, which lack the power to detect treatment effects. Thus, a summary or evaluation of a set of experiments may provide misleading conclusions biased toward no detected effect (see discussions in Hedges and Olkin, 1985). Although metanalysis is quantitative, and does not rely upon the conclusions of the original authors (Gurevich and Hedges 2001), it is still a weak form of synthesis because it is limited to estimates of effects determined by the original authors of the studies.

HOW INTERSITE COMPARISON HELPS DISCRIMINATE AMONG ALTERNATIVE HYDROLOGIC MECHANISMS

The essence of an intersite comparison is that it examines primary rainfall and/or runoff data to determine the difference, or change, between two or more datasets, at two or more time periods, in order to elucidate hydrologic processes. The best-understood example of this approach is paired-basin forest harvest treatment experiments, in which the ratio of the runoff at the treated and control basin is compared between the pretreatment and the treated period to infer how the treatment (removal or replacement of vegetation) has affected runoff. *In this article, intersite comparisons encompass all analyses in which rainfall and/or runoff is compared among any two or more sites or two or more time periods, with an aim to inferring hydrologic mechanisms.*

Intersite comparison, when it involves analysis of primary data, can be a useful tool for discriminating among alternative process explanations for observed rainfall-runoff relationships. Analysis of primary rainfall-runoff data can overcome many of the deficiencies of secondary analyses such as literature review and metaanalyzes. Analysis of original data allows us to

- expand the sample size to include data not yet analyzed and replicate hypothesis tests;
- ask new questions, even of old data, posed by the current context for hydrology;
- make new comparisons among records not formerly juxtaposed;
- use novel statistical approaches to reveal hitherto obscure features of the data; and
- use ancillary data to refine hypothesis tests about hydrologic mechanisms.

Replicating Hypothesis Tests and Expanding Sample Size

Intersite comparison studies can expand the sample size of records tested for certain hydrologic responses, and thereby replicate hypotheses tested in earlier studies. Our inferences about hydrologic processes are drawn from an unnecessarily narrow subset of temporal scales, spatial scales, and geographic conditions, given the range of data available (Figure 1).

A vast amount of rainfall and runoff data have been collected, but never analyzed (Figure 1). Rainfall and runoff data have been collected for centuries to aid in predictions of water yield, flooding, and other hydrological properties. Over the past century, and particularly the past 50 years, streamflow gaging technology has permitted the acquisition of continuous records with very fine temporal resolution (5 or 15 minutes). Sites with streamflow gaging have proliferated and number in the millions, with more than 1.5 million sites gaged historically by the United States Geological Survey alone. Continued developments in computer database technology and informatics have permitted data to be digitized and stored for ready access. Many rainfall and runoff records, especially those managed by government agencies, are now publicly available on the Internet.

The fact that rainfall and runoff records are held by a variety of agencies may partly explain why some records have never been analyzed. Streamflow and precipitation have been monitored systematically in basins in the United States and abroad since the early to mid twentieth century. In the United States, such monitoring has occurred under the auspices of individual states, the US Forest Service, the US Geological Survey, the USDA Agricultural Research Service, and state agencies. In the United Kingdom, monitoring is conducted under the auspices of the Center for Ecology and Hydrology. In Australia, monitoring is conducted by the water resources branch of the CSIRO, Forestry Tasmania, and state and municipal agencies. In Japan, streamflow monitoring is conducted by the Forestry and Forest Products Research Institute and Japanese universities. All countries have some form of streamflow monitoring, although record length, quality, and availability vary.

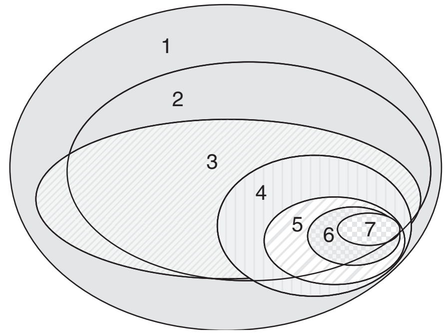

Figure 1 Most inferences from data about hydrologic processes are drawn from an unnecessarily narrow subset of temporal scales, spatial scales, and geographic conditions, given the range of data available. A distressingly large number of publications simply repeat (or challenge) the results of a few primary analyses without exploring unexamined datasets. Only a small fraction of the published studies in hydrology (1) utilize any measured streamflow data (2) (boxes not to scale). The majority of these studies are based on data from small basins (3) despite the enormous numbers of gaged large basins. Of the small, instrumented basins whose records are noted in any publication (3), only a fraction have been analyzed (4), so that many basin records remain unexamined. Few of these studies involve analysis of data at finer than annual time steps (5), and only a fraction of studies utilize records longer than a decade (6), so very few studies examine daily data over long periods comparable to most hydrologic modeling efforts (7). Apart from analyses of instantaneous peak flows, virtually no studies have compared streamflow data among sites at finer than daily time steps. Thus, there are at least four opportunities for intersite comparisons involving hypothesis testing and expanded data use: (a) using data to verify modeling and process studies, which have not involved measured data; (b) analyzing large basin records that have never been looked at; (c) analyzing small basin records that have never been looked at; and (d) extending the time periods, or increasing the temporal resolution (e.g. from annual to daily, daily to hourly) of analyses for basins whose records have been examined. A color version of this image is available at http://www.mrw.interscience.wiley.com/ehs

Also, intersite comparisons may be made more difficult by the fact that agency jurisdiction and consequent objectives and study designs in monitored basins vary with basin size. Small basins are more likely to be instrumented for research and forest or agricultural land use management, whereas large basins are more likely to be instrumented for water management and supply assessment objectives. In the United States, most basins monitored for research purposes by federal agencies (USFS of the USDA, NCRS of USDA, USGS research branch) are small, usually $<1\,km^2$, usually with a single land owner and one dominant vegetation type. In contrast, basins monitored for water supply and flood control purposes may drain areas up to millions of square kilometers and typically encompass multiple land uses and vegetation types.

Intersite comparisons could replicate a number of hypotheses that have been advanced about hydrologic responses, mostly based on experiments designed to test the effect of vegetation removal on runoff. These hypotheses are based on analyses conducted from records using a small number of basins, limited record length, basins of only one – usually small – size, and only one temporal resolution – usually annual runoff (Figure 1). Several key hypotheses merit replicated testing, including statements (from Hibbert, 1967; Bosch and Hewlett, 1982; Jones and Grant, 1996; Beschta *et al.*, 2000) that the hydrologic response to vegetation treatment is positively related to

- the amount of vegetation removed;
- the amount of precipitation;
- the evergreen-ness of the vegetation;

and negatively related to

- the amount of runoff;
- the size of the basin.

Analyses of primary rainfall and runoff data using more sites, longer records, and finer temporal resolutions of data could replicate tests of these hypotheses, and resolve contradictions among them. Reviews (e.g. Hibbert, 1967, Bosch and Hewlett, 1982) have included many small paired-basin records in the initial posttreatment period. However, the primary data from many of these small, paired basins have not been reanalyzed, although much of these data are now available through the Internet *(see* e.g. the Clim-DB, Hydro-DB website **`http://www.fsl.orst.edu/climhy/`**). Moreover, the primary records from other small basin sites that were not part of paired-basin experiments, recently established paired-basin experiments, and a great many large basins monitored for flood control and water yield have not been analyzed (Figure 1). In cases where streamflow and climate records are continually expanding, new trends may emerge from examination of longer records. The treated-control relationships in small paired-basin experiments summarized by Bosch and Hewlett (1982) could be reanalyzed using longer records to ascertain how vegetation succession alters water yield over time. Examination of finer resolution data may help discriminate multiple scales of pattern indicative of several interacting hydrologic mechanisms. For example, Jones and Post (2004) showed that basins with similar hydrologic response to forest harvest at the annual streamflow resolution had very different day-to-day patterns of hydrologic response. Continuous records of streamflow or precipitation that may have been analyzed at one temporal resolution (such as annual or monthly) can be resampled at finer resolutions (such as storm, daily, or hourly) to seek evidence for hydrologic processes that are not detectable at monthly or annual time scales.

Posing and Answering New Questions

Often, intersite hydrology involves asking new questions about old data. Typically, streamflow monitoring studies were established for reasons that are no longer principal motivating factors for research or management. For example, an original objective for many USGS instrumented basins included establishing flood frequency curves for designing downstream engineering structures such as dams or bridges. However, in many cases, the construction of engineering works has been completed, while upstream land use may have changed the flood frequency distributions. Also, many USFS sites established paired watershed experiments to determine how water quantity would be affected by forest harvest treatments. However, forest harvest and road building practices have changed and no longer resemble the treatments imposed in the experiments.

Changing contexts have created important new questions about rainfall-runoff processes. New areas of concern include considerations of climate change, scaling, and the ongoing challenge of prediction in ungaged basins. For example, we can ask how the rainfall-runoff relationship has changed over time, and interpret those findings to infer changes in climate and vegetation. Also, we can examine how the rainfall-runoff relationship varies with spatial scale to understand how perturbations propagate in streamflow networks. We can combine these two general questions to determine how the rainfall-runoff relationship varies as a function of basin scale and the temporal resolution of the data to help scale up our understanding of hydrologic processes to that of global climate models. We can also ask how the geography of basins influences the rainfall-runoff relationship based on intersite comparisons, and perhaps derive principles on which to base predictions of runoff in ungaged basins.

Making Novel Intersite Comparisons

A key component of intersite comparison of rainfall and runoff involves the choices of basins to compare. Intersite hydrology involves comparisons of rainfall and runoff

Table 1 Novel comparisons among basins represent new experimental designs for intersite comparison to test a variety of hydrologic mechanisms influencing rainfall-runoff relationships. Examples are based on runoff data, but could be extended to rainfall/runoff ratios in cases where both rainfall and runoff data exist for all the sites

Compare	To	Study question	Hydrologic interpretation
Runoff at treated basin (vegetation removal, roads)	Runoff at control basin	Was there a detectable effect of an experimental treatment?	Water balance altered by vegetation removal and/or roads routing water
Runoff at treated basin, by time period since treatment	Runoff at control basin, by time period since treatment	How did the treatment effect change over time?	Water balance affected by vegetation recovery or other covarying factors
Runoff at treated basin, by season	Runoff at control basin, by season	How did the treatment effect vary by season?	Water balance affected by seasonal variation in wetness, temperature, and/or other factors
Runoff at treated, by event size	Runoff at control, by event size	How did the treatment effect vary by the size of the runoff event?	Water balance affected by antecedent wetness, precipitation, snow, and/or other factors
Runoff at multiple treated basins, by time since most recent pretreatment disturbance	Runoff at multiple control basins, by time since most recent pretreatment disturbance	How did the treatment effect differ among basin pairs according to the time since most recent pretreatment disturbance?	Water balance affected by changes in vegetation structure and species composition over succession
Runoff at small basin	Runoff at large basin containing the small basin	How does runoff vary with basin scale?	Hydrograph affected by routing of water in network and/or differences in vegetation type and cover and/or roads small to large
Runoff at basin in climate type a	Runoff at basin in climate type b	How does runoff vary with climate?	Water balance affected by differences in timing of moisture, freezing
Runoff at basin in vegetation type a (or land use type a)	Runoff at basin in vegetation type b (or land use type b)	How does runoff vary with vegetation type (or land use)?	Water balance affected by differences in timing, amount of transpiration, evaporation, interception

among basins according to the differences in those basins, whether differences arise from "controlled" or inadvertent experiments, or geographical differences. It involves examining flow regimes and responses to treatment across divergent vegetation, hydrologic systems. Intersite comparison studies could make important contributions to inferring the importance of various hydrologic processes by examining novel pairings of basins that have not been compared before (Table 1). Obvious comparisons involve (i) multiple neighboring or nearby basins, and (ii) nested basins.

Multiple small, neighboring, or nearby basins are a logical choice for intersite comparisons. Small basins are instrumented at many United States Forest Service-sponsored experimental forests (e.g. Andrews, Caspar Creek, Coweeta, Fernow, and Hubbard Brook Experimental Forests) as well as United States Department of Agriculture Agricultural Research Service sites (e.g. Reynolds Creek, Walnut Gulch). When multiple basins are instrumented in a single site, they allow examination of environmental variation and treatment effects on streamflow. Some sites have implemented one or more "paired watershed experiments" in which a small control and one or more small treated (usually adjacent) basins are monitored over some pretreatment period, and then a treatment, such as vegetation removal, is imposed on one basin, and the monitoring of both the treated and control basin is continued for some posttreatment period. Bosch and Hewlett (1982) identified nearly 100 such treated/control basin pairs in several dozen locations with experiments involving modification of forest, shrub, and grassland vegetation.

Multiple large basins also could be compared using intersite analysis. Large basins generally are instrumented for flood and water yield forecasting. Large paired-basin experiments do not exist because of the difficulty of implementing a treatment over a large area and because of the lack of large basins in a "control" state. However, neighboring or nearby large basins with contrasting vegetation histories can be considered as parts of inadvertent experiments (*see* e.g. Jones and Grant, 1996; Thomas and Megahan, 1998). Nested basin comparisons are possible when small, instrumented basins occur inside large, instrumented basins (see e.g. the Andrews Forest, Hubbard Brook, Coweeta). Nested basins allow the testing of scaling relationships (Gupta and Waymire, 1998).

Table 2 Novel statistical approaches, multiple sites, and longer records permit asking novel questions in intersite comparisons of rainfall and runoff data

Questions	Number of Variables	Sites	Periods	Analysis tool	Illustrative citations
1. What is the mean value of rainfall, runoff?	One	One	One	Mean	Post *et al.* (1998)
2. What is the variation in rainfall, runoff?	One	One	One	Variance, standard deviation, coefficient of variation	Post and Jones (2001)
3. What is the shape of the distribution of rainfall, runoff?	One	One	One	Mean, variance, and higher moments; quantiles	Gupta and Waymire (1989)
4. How has the mean value of rainfall, runoff varied over time?	One	One	Two or more	Means	Andreassian *et al.* (2003)
5. How is rainfall, runoff related to itself over time?	One	One	Two or more	Time-series methods: autocorrelation, spectral analysis, wavelets	Kirchner *et al.* (2000) Lafrenière and Sharp (2003) Tague and Grant (2004)
6. How is runoff (rainfall) related across sites?	One	Two or more	One	Linear regression	Hibbert (1967), Bosch and Hewlett (1982), Andreassian (2004)
7. How does the shape of the rainfall or runoff distribution change across sites or time?	One	Two or more	One or more	Statistical self-similarity; quantiles	Gupta and Waymire (1993, 1998), Andreassian *et al.* (2003)
8. How does the rainfall or runoff relationship between two sites change over time?	One	Two	Two or more	Linear regression for each time period; anova by time period	Jones and Grant (1996), Thomas and Megahan (1998), Beschta *et al.* (2000), Jones (2000), Jones and Post (2004)
9. How is rainfall related to runoff?	Two	One	One	Rainfall-runoff ratio	Post and Jones (2001), Kokkonen *et al.* (2004)
10. How is rainfall related to runoff at a site over time?	Two	One	Two or more	Cross-correlation; cross-spectral, cross-wavelet	Post and Jones (2001), Lafrenière and Sharp (2003)

Novel Statistical Approaches Reveal Hitherto Unquantified Aspects of Data

Most of the published analyses of rainfall and runoff data are based on a narrow set of statistical tools. Linear regression has been widely used. However, other techniques, including time-series analysis, autocorrelation, and cross-correlation approaches have the potential to reveal aspects of rainfall and runoff data that could be compared among sites, and lend themselves to interpretations of physical or biological processes (Table 2).

New Ancillary Data to Refine Hypotheses

Counter to the assertion that rainfall-runoff data are a "dead end" for determining hydrologic processes, this article contends that intersite comparisons of rainfall and runoff data can shed light on the hydrologic processes operating within the basin. While such interpretations are not definitive, rainfall, runoff, and ancillary data can be used to narrow the range of possible interpretations of hydrologic mechanisms operating at a given site for a given time period. In so doing, researchers are engaged in a process analogous to that conducted by hydrologic modelers in selecting the appropriate model structure and identifying correct parameters.

Intersite comparisons involve drawing inferences about the water balance – revealing or inferring the processes inside the black box. The water balance is the basic theoretical tool for the interpretation of analyses of rainfall-runoff data (Figure 2). Any analysis of rainfall and runoff confronts the difficulty of discriminating which of the large number of terms in the water balance explain the hydrologic processes at a site under a given set of conditions (Figure 2a).

Intersite comparisons can make progress by utilizing ancillary data to narrow down the number of possible terms in the water balance that could be operating under certain conditions, thereby reducing the number of possible mechanistic interpretations (Jones and Post, 2004). If certain moisture, temperature, or eco-physiological conditions are

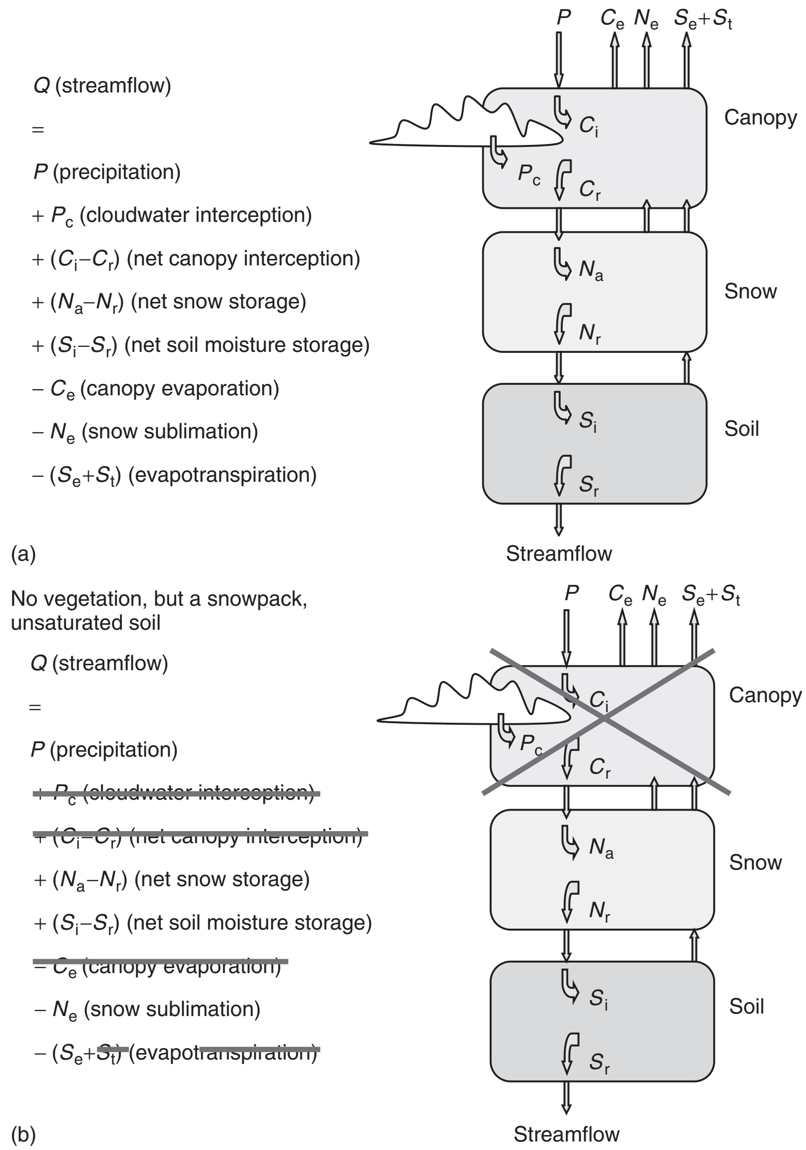

Figure 2 An approach to simplifying the number of possible hydrologic processes that can explain rainfall-runoff data involves partitioning data into subsets, each of which can be argued to involve a subset of terms in the water balance. Thus, the complete water balance for a given site (a) can be simplified into a set of water balances for conditions when (b) vegetation is absent, or (c) both vegetation and snowpack are absent, or (d) snowpack is absent and soils are near-saturated, or (e) vegetation is not actively transpiring, snowpack is absent, and soils are near-saturated. Analyses of rainfall-runoff relationships for each of these subsets of data can help bracket the range of values for the terms in the water balance. A color version of this image is available at http://www.mrw.interscience.wiley.com/ehs

met, key terms can be effectively removed from consideration in the water balance for a given place or time period (Figure 2b–2e). For example, if vegetation is absent, soils are unsaturated, and a snowpack is present, any observed change in streamflow relative to precipitation must involve snow and soil moisture storage only (Figure 2b). If, in addition, no snowpack is present, the water balance can be simplified to consider only the effects of the soil moisture storage (Figure 2c). Alternatively, when vegetation is present, but snow is absent, and if the soil is saturated, then the water balance is regulated by interception, evaporation, and transpiration processes in the vegetation canopy (Figure 2d). To further simplify the water balance, it may be possible to identify periods when vegetation is not actively transpiring, and the water balance can be reduced to the interception and evaporation terms (Figure 2e).

This kind of approach can guide intersite comparison studies, because data are collected on related climate variables in addition to precipitation and streamflow at most

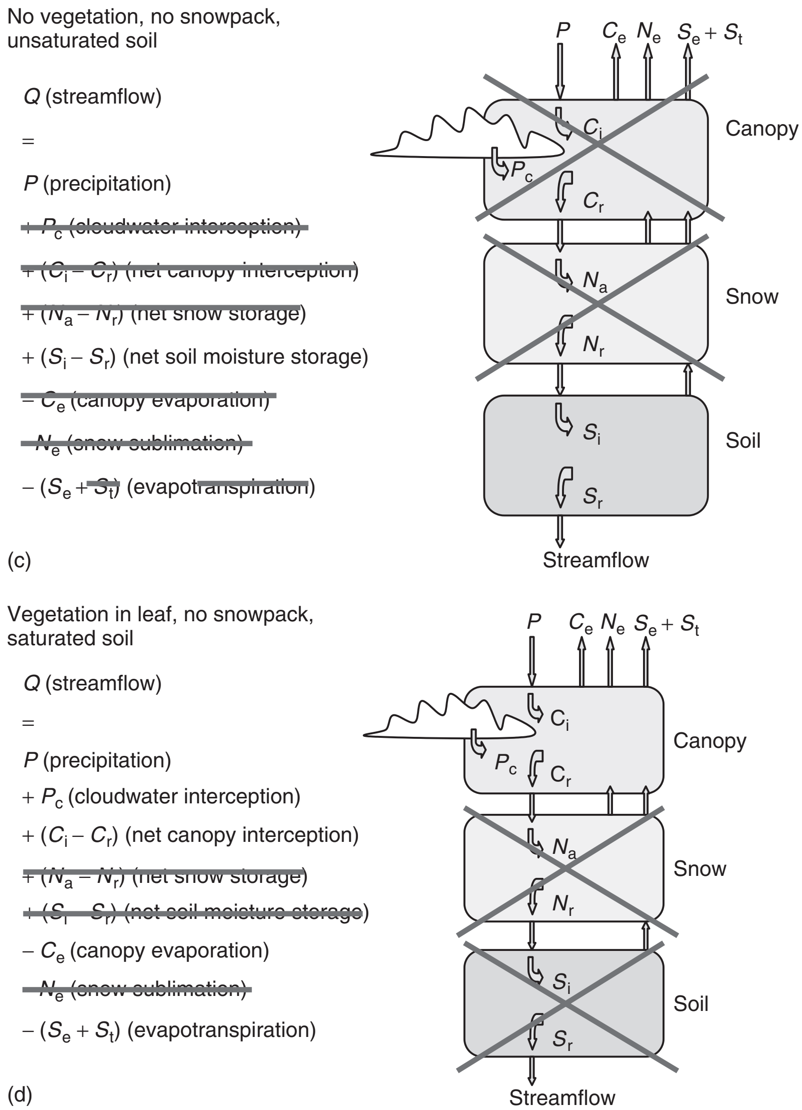

Figure 2 (*continued*)

instrumented sites (see the Reynolds Creek dataset publications for an excellent example, described in Slaughter *et al.*, 2001; Pierson *et al.*, 2001). Instrumented variables differ among sites, and, over time, within a site, depending upon the objectives for the site and whether the data have been or are being utilized for research and/or management.

AN OUTLINE OF STEPS INVOLVED IN CONDUCTING INTERSITE COMPARATIVE ANALYSIS

The steps involved in conducting an intersite comparison are

- identifying a question;
- developing a study design;
- selecting, accessing, and merging datasets; and
- choosing statistics for comparison of rainfall and runoff data.

Identifying a Question

Many questions are possible for intersite comparisons (Tables 1 and 2). The list below (and Tables 1 and 2) contains questions that hydrologists can answer using intersite comparison. Answers to these questions would represent important contributions to the current state of knowledge in hydrology.

1. *Are basins unique, or are there consistent types of behavior among basins?*

Intersite comparisons among multiple instrumented basins may reveal consistent groupings of basin behavior. If

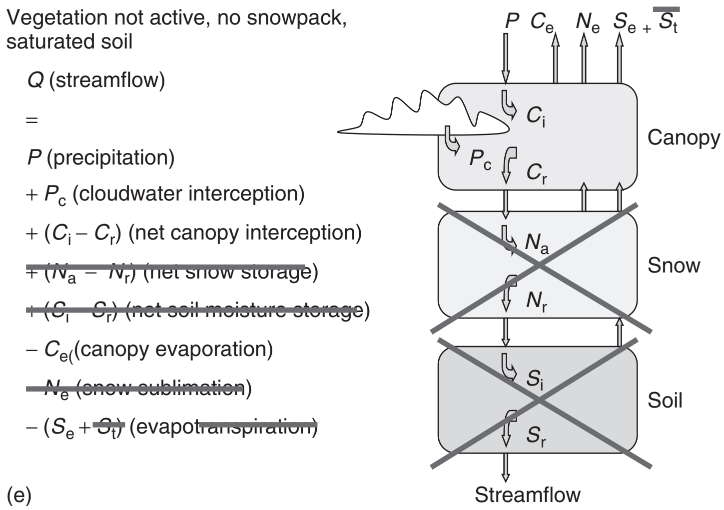

Figure 2 (*continued*)

consistent types of rainfall-runoff behavior can be identified for groups of instrumented basins, a classification system may be developed, which could facilitate the prediction of runoff from ungaged basins (see the IAHS PUB initiative).

2. *What factors (climate, vegetation, subsurface properties, land use history) explain similarities and differences in streamflow among basins?*

Intersite comparisons can involve *post hoc* superexperiments, in which rainfall-runoff responses associated with one factor (such as vegetation) can be examined while controlling for other factors (such as climate, land use history). Understanding, and separating, the roles of climate, vegetation, and other factors will help elucidate hydrologic mechanisms that account for streamflow in ungaged basins.

3. *Can streamflow be predicted in ungaged basins based on a priori knowledge?*

Records from multiple instrumented basins could be examined in such a way as to test models for predicting streamflow in ungaged basins.

4. *Are precipitation, streamflow, or rainfall-runoff relationships changing over time?*

Rainfall-runoff relationships are the basis for predictions of water yield and flooding and other practically relevant aspects of hydrology. Intersite comparison of long-term rainfall-runoff records can reveal changes in these relationships, and, potentially, the factors that cause the changes.

5. *How does the variability of streamflow and rainfall change with time scale – diurnal, storm, seasonal, or the time scales of vegetation succession and climate change?*

Existing records of streamflow and precipitation have high temporal resolution, potentially revealing processes occurring at multiple time scales. However, most published analyses of these records have dealt with annual or mean annual data. Intersite analyses of these records may reveal hydrologic processes that produce streamflow responses at diurnal or storm time scales, while other analyses may reveal processes operating at the time scales of vegetation succession and climate change.

6. *Are changes over time in streamflow occurring in response to vegetation change, climate change, or other drivers?*

Intersite comparisons can involve examination of changes in basins whose vegetation and climate history is well documented, potentially revealing responses to vegetation, climate change, natural and anthropogenic disturbances, and their influences on streamflow.

7. *Do basins respond differently to similar perturbations or treatments? If so, what accounts for differential responses?*

Intersite comparisons can investigate the causes of divergent streamflow responses to similar treatments (such as vegetation removal). For example, the presence of snow, soil depth and texture, and the seasonality of vegetation water use; all may produce different responses to the same treatment in different basins.

8. *Do basins respond predictably to perturbations or treatments?*

Very few of the existing long-term rainfall/runoff records are regularly used for testing hydrological models. Yet, the lengths of these records, the detailed associated knowledge about these basins, and the differences among them provide many opportunities for testing hydrologic model predictions.

9. *Are streamflow responses to perturbations or treatments linear or nonlinear? Are there thresholds of response?*

Intersite analysis of rainfall-runoff records offers the opportunity to seek nonlinear responses, or thresholds, in hydrologic system behavior.

Developing a Study Design

The second step in an intersite hydrology comparison is to establish an experimental design. The experimental design involves identifying a comparison that appropriately tests the study question (Table 1). It also involves selecting the number of basins, variables, and time periods for the analysis (Table 2). In an intersite comparative analysis of primary rainfall and runoff, data are included from a sample of basins selected because the hydrologic processes are expected to differ among these basins in a predictable fashion. Study basins may be selected to replicate the effects of a given treatment (deliberate or inadvertent), or to examine rainfall or runoff relationships across a gradient of conditions.

A number of decisions are made in designing any intersite comparison. These include the following:

1. Choosing the study basins.
2. Choosing a temporal resolution.
3. Stratifying data by time period, season/climatic conditions, event size, or other factors.

The study basins should be chosen so as to allow testing of the question of interest, usually by selecting basins that differ from one another according to the factor of interest (whether or not this difference was the result of a human-imposed treatment) (Table 1). The temporal resolution of the data (ranging from 15 minutes to annual or decadal) also should be chosen so as to permit testing of the question of interest. For example, questions involving the timing and routing of runoff (e.g. flooding) focus on the time scale of individual storms (peak discharges, storm hydrographs), whereas questions involving the amounts of water in various components of the water balance may focus on hourly, daily, or annual data. Because the influence of hydrologic mechanisms varies according to the conditions, it can be quite important to stratify data by time period (e.g. age of vegetation, or time since treatment), by season or climatic conditions (e.g. subzero vs warm temperatures, wet vs dry conditions), by event size (e.g. amount of precipitation or runoff), or other factors.

Selecting, Accessing, and Merging Datasets

Selecting, accessing, and merging datasets is the third step in an intersite comparison. Data availability poses some restrictions on the types of intersite hydrology comparisons that can be made (Figure 3). Streamflow, precipitation, temperature, snow, vegetation, geology, topography, and other records are all pertinent to intersite comparisons. Each form of data has its own advantages and weaknesses.

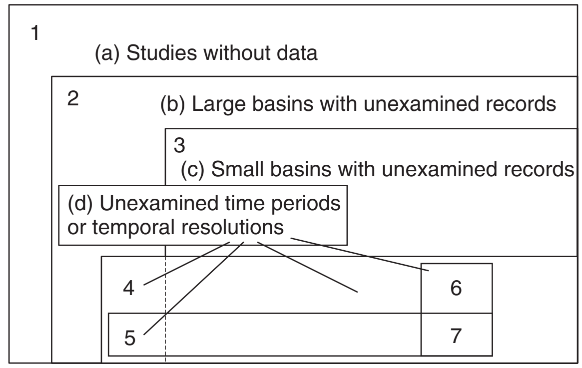

Figure 3 Data availability poses some restrictions on the types of intersite hydrology comparisons that can be made (ellipses not to scale). Many basins have streamflow records (1), but only some have been digitized (2). Ancillary data on precipitation and temperature, not all of which is digitized, are available only for some basins (3). Of the total number of gaged basins, only a few involve treated/control pairs, and not all of these records are digitized (4). Ancillary data on vegetation, snow, and other ecological factors is available for only a subset of basins, whether they are treated-control pairs or not (5). Online datasets that involve streamflow, precipitation, temperature, and ancillary data on vegetation, snow, and other ecological factors are only a small fraction of the total data, and these are limited to daily time resolutions (6). A very few sites have online data available at finer than daily time steps (7)

Streamflow measurements vary in accuracy and precision, and these differences among sites determine the minimum change that can be detected in an intersite comparison of runoff data. In small basins, streamflow is measured using flumes or weirs with controlled cross sections, so stage height observations are readily converted to discharge. The greatest uncertainty, however, is associated with the largest flows for which measurements are most difficult to obtain. In large basins, discharge is calculated on the basis of the measured stage height on a known cross section; changes in the cross section associated with flooding add additional uncertainty to discharge estimates from large basins. Streamflow records typically are continuous, so, if the original charts are saved, or the record is digitized, it is possible to capture streamflow patterns at very fine temporal resolutions, including small diurnal fluctuations (*see* e.g. **www.fsl.orst.edu.lter/data/hydrology**). The weir, flume, or cross section determines the ability of the streamflow record to capture small variations in discharge. For example, trapezoidal flumes typically cannot capture very small fluctuations as well as V-notch weirs. Thus, the minimum detectable change in a given streamflow record is smallest for records collected in small basins using V-notch weirs, and largest for records collected in large basins where cross sections change frequently.

Precipitation and air temperature records are also typically available from sites adjacent or near to streamflow gages, but it is challenging to match point measurements of precipitation with area measurements of streamflow. Capturing the spatial variability of precipitation (or air temperature) is problematic, particularly for large basins, and especially for particular time periods in the past. Most small basin streamflow studies involve one or more precipitation and air temperature gages. When multiple precipitation (air temperature) gages are present, it is possible to spatially interpolate precipitation or temperature over the basin (Daly *et al.*, 2002; Smith, 2002). Maps of interpolated precipitation can be found at **http://www.ocs.orst.edu/prism/**, and illustrative maps of temperature interpolations can be found at **http://www.ocs.orst.edu/pub/smithjw/hja/**).

Accurate representation of precipitation is more difficult in large compared to small basins, and this difficulty is compounded by the fact that precipitation gages in large basins are maintained by different agencies than those that manage streamflow data collection. Thus, for example, in the United States, the US Geological Survey manages streamflow data collection from large basins, but state climate services or the National Oceanic and Atmospheric Administration manage precipitation records.

In an intersite comparison study, it is generally possible to match fine-resolution temporal streamflow data with precipitation data at the small basin scale, because these records usually were collected as part of the same monitoring effort. However, at the large basin scale (above 1 km^2), it may not be possible to match historical streamflow with precipitation at temporal resolutions finer than monthly, particularly if spatial heterogeneity is important. Obtaining historical precipitation data for periods of record that match gaged periods in large basins can be a significant challenge. Nevertheless, historical climate records have been compiled at the national scale in some cases (*see* e.g. for the United States, **http://cdiac.esd.ornl.gov/epubs/ndp019/ndp019.html**). Temporal resolution of precipitation and air temperature data at the small basin scale is usually high, and may even match that of streamflow (e.g. 15-minute data). However, obtaining spatially interpolated historical – or even current – precipitation and air temperature data at temporal resolutions finer than monthly may be difficult for many large basin sites (*see* **http://www.ncgc.nrcs.usda.gov/branch/gdb/products/climate/data/index.html**).

Other data, such as vegetation cover and type, soils, geology, and land use history may be critical for interpreting intersite hydrology comparisons. However, these data are much less likely than streamflow or climate data to be available or tabulated in a format that is easily interpreted. Some good examples of vegetation, soils, and other spatial datasets linked to streamflow and precipitation can be seen at long-term ecological research (LTER) sites (**www.lternet.edu**), especially those that were originally experimental forests, in the United States. Even if they are available, vegetation, soils, and land use history data are very unlikely to have been collected in a consistent framework across multiple sites, increasing the difficulty of making use of such data in an intersite hydrology comparison. When conducting an intersite hydrology comparison, it is usually necessary to contact researchers or managers at individual sites for help in obtaining and interpreting this sort of data.

Intersite hydrology comparisons are limited by the availability of data (Figure 3). Efforts are underway to systematically collect and provide hydrologic data (see hydrodb website, **http://www.fsl.orst.edu/hydrodb/**). However, hydrologic data may still be in original analog formats, or not yet publicly available, at many sites of interest for intersite hydrology studies. Researchers will need to seek out data of interest from these individual sites, and they will also need to support and participate in efforts to make such data publicly available.

Intersite comparisons require that hydrologic data be collected using comparable measurement methods, units, and data records. Relative to other kinds of comparative studies, such as vegetation or stream chemistry, intersite hydrology and climate studies are made easier by the fact that streamflow data collection methods are largely standardized and common units (e.g. cfs, cms) are used. Both precipitation and streamflow data are typically collected at 15-minute or finer resolution, and are tabulated at daily, monthly, and annual resolution.

An intersite hydrology comparison also may require overlapping records, but certain intersite hydrology comparisons may not. Runoff ratios require overlapping precipitation and streamflow periods of record, and intersite comparisons of runoff ratios probably should be made controlling for climate by using a common period of record. Paired-basin experiments require the collection of simultaneous records at treated and control basins. Intersite comparisons looking for trends over time due to climate change also require simultaneous records. On the other hand, intersite hydrology comparisons of paired-basin experiments need not have records from the same time periods, since the questions of interest relate to the periods before and after imposition of a treatment.

The lengths of record available limit the scope of what can be determined from an intersite hydrology comparison (Figure 3). Short records (i.e. a few years) lend themselves to calculations of metrics, such as runoff ratios, which may be compared to other sites, or examined to compare hydrologic processes at finer than annual temporal resolutions such as seasonal, storm, or daily time scales. In contrast, long records allow examination of changes over time in metrics, such as runoff ratios, that may be related to longer-term processes, including vegetation succession, disturbance, or climate change.

The numbers of sites with records of hydrology, climate, and environmental variables also limit the scope of what can be learned from intersite hydrology comparisons (Figure 3). There is no complete listing of small monitored basins, or of small paired-basin experiments. Only a couple of dozen small basins ($<1\,km^2$) exist in the United States where streamflow and precipitation data have been consistently monitored for periods of several decades and where good records exist describing vegetation cover, soils, and landscape history (see **`http://www.fsl.orst.edu/hydrodb/`**). Bosch and Hewlett (1982) noted a total of 26 sites in 16 states in the United States where small paired-basin experiments had been conducted, but many of these records are not available in digital form. On the other hand, hundreds of large basins (10^1 to $10^6\,km^2$) have long-term streamflow records collected by the USGS. However, these sites may not have matching precipitation records, and they do not have compiled records of vegetation cover, soils, or landscape history.

Choosing Statistics for Comparison of Rainfall and Runoff Data

Lots of ink and breath have been expended in statistical controversies in hydrology. Some argue that discussion of statistical methods distracts from the essential questions. However, statistical approaches force researchers to clearly articulate the questions and hypotheses they are testing. Many statistical techniques are now available (Table 2). As in any statistical analysis, the analysts must keep in mind that if data are not independent, or if data are not normally distributed, certain statistical findings will be biased. Thus, tests of autocorrelation, and data transformations, are an important component of intersite comparisons.

The statistics used may be measures of single variables, or they may be measures involving two or more variables. Statistics that are useful for intersite comparisons include means, variances, general linear models (regression and analysis of variance), autocorrelation methods (including spectral analysis and wavelets), and cross-correlation methods (Table 2). Single-variable measures include means, variability (e.g. standard deviation), distributions or measures of distribution shape, or measures of autocorrelation. Bivariate measures include linear regression, analysis of variance, rainfall-runoff ratios, and cross-correlation.

Single-variable Analysis – Comparison of Means

Mean daily (or monthly, or annual) streamflow may be compared across a range of sites in an intersite hydrology study (Table 2). For example, Post *et al.* (1998) compared mean annual precipitation and mean annual runoff across sites. This kind of comparison of long-term averages can reveal differences in climate, or which types of vegetation use larger or smaller proportions of water input as precipitation. Post *et al.* (1998) showed that rainfall-runoff ratios were similar across four forested sites with strongly contrasting climates.

Single-variable Analysis – Comparison of Variance or Variation

A second single-variable approach to comparison involves measures of variability (Table 2). Standard deviations, standard errors, or coefficients of variation (standard deviation divided by the mean) of streamflow can be evaluated to compare variability among sites or time periods. For example, Post and Jones (2001) conducted an intersite comparison based on the coefficients of variation of daily, monthly, and annual streamflow at small, forested basins in the United States. They replicated the common finding of decreased variability with increasing temporal aggregation, noting that streamflow variability was much higher at daily timescales than at seasonal or annual time scales, at four sites with temperate deciduous, temperate conifer, and tropical forest. More interestingly, from the point of view of intersite comparison, Post and Jones (2001) showed that the greatest variability in rainfall and runoff varied among sites for any give time scale: variability at the daily time scale was greatest at eastern United States temperate forest sites in mesic climates, variability at the seasonal time scale was greatest at western United States conifer forest sites in xeric (Mediterranean) climates, and annual variability was greatest at tropical forest sites in hurricane-affected Puerto Rico.

Single-variable Analysis – Comparison of Higher Moments of Distributions or Quantiles

A third single-variable approach involves creating and comparing measures of the shapes of distributions of precipitation, streamflow or rainfall-runoff ratios (Table 2). For example, Andreassian *et al.* (2003) created measures of the proportions of the daily streamflow distributions that fell above or below certain benchmarks (e.g. flood flows, low flows, base flows) and compared how these proportions changed over time for a number of small, forested basins. This kind of comparison can be used to detect changes over time, or cross-site differences, in low flows, high flows, or other parts of the flow frequency distribution. Andreassian *et al.* (2003) showed that flow frequency distributions had remained relatively constant at a recently reforested basin in Ohio, USA (Coshocton) and a burnt shrubland basin in southern France (Real Collobrier), but low flows and baseflow had shifted downwards over time at an old-growth basin in Oregon, USA (Andrews WS 2). Gupta and

Waymire (1998) used the moments of the distributions of rainfall and runoff to test for self-similarity across a set of basins arranged in increasing size.

Single-variable Analysis – Comparison of Autocorrelation

A fourth single-variable approach involves measures of how streamflow or rainfall are related to themselves over time (Table 2). Measures of autocorrelation can be used to characterize and compare streamflow, precipitation, or rainfall/runoff ratios among sites at various time scales, which can then be compared across sites or time periods. For example, Post and Jones (2001) conducted an intersite comparison based on the autocorrelation coefficients for a range of lags using both daily and monthly streamflow at four forested sites. They were able to detect differences among sites in the seasonal and daily timescales rainfall-runoff relationships that were linked to differences in climate and vegetation. Tague and Grant (2004) applied autocorrelation analysis to streamflow records from two large basins in Oregon, showing how the seasonality of flows differed between two basins with contrasting geology.

Single-variable Analysis – Comparison of Spectra or Wavelets

More complex single-variable approaches to intersite comparison of rainfall and runoff data involve the use of spectral analysis or wavelets (Table 2). Kirchner *et al.* (2000) constructed log–log plots of the spectral power of rainfall, streamflow, and stream chloride concentrations in catchments in Plynlimon (Wales), and showed that the plot of spectral power of chloride concentration was fractal (had a positive slope). Lafrenière and Sharp (2003) used wavelet analysis of streamflow from glacially controlled lakes to infer and separate the influences of snowmelt from rainfall on discharge.

Bivariate Analysis – Comparison Using Linear Regression

Linear regression has been widely used to compare two sites over two or more time periods (Table 2). The slope of a regression line comparing e.g. runoff at site 1 versus site 2 (y/x) is a ratio that can be compared across sites or time periods. The earliest studies of paired-basin experiments used a regression of streamflow at the treated (y-axis) versus the control (x-axis) to characterize the treatment effect. Beschta *et al.* (2000) used linear regression to reexamine data from the small basins in Jones and Grant (1996). Jones and Grant (1996) used regression analysis to examine the ratio of change in peak discharge versus change in cumulative area harvested in large (60 to 600 km^2) basins, as a way to detect whether forest harvest was associated with changes in peak discharges. Thomas and Megahan (1998) used analysis of covariance to reexamine data from large basins in Jones and Grant (1996).

Bivariate Analysis – Comparison Using Analysis of Variance

Analysis of variance has been less widely used than linear regression, but it also can be used to compare two sites over two or more time periods (Table 2). The ratio of e.g. runoff at site 1 versus site 2 can be tested for significant changes by time period using ANOVA. A more useful ratio is the log-transformed ratio, i.e. $\ln(y/x)$, which is equivalent to the difference of the log-transformed values, i.e. $\ln(y) - \ln(x)$. Eberhardt and Thomas (1991) recommend that the log-transformed ratio be used in analysis of streamflow from paired-basin experiments. For example, Jones and Grant (1996) and Jones (2000) used the log-transformed ratio of streamflows at the treated versus the control basins as a dependent variable in an analysis of paired-basin treatment effects on peak discharges. Using this ratio, they were able to show how peak discharges changed in the treated versus the control in small paired-basin experiments over 5-year periods after forest harvest (Jones and Grant, 1996; Jones, 2000). Jones and Post (2004) used this ratio to calculate how daily flows changed in the treated versus the control basins in small paired-basin experiments over 5-year periods after forest harvest.

Bivariate Analysis – Comparison Using Rainfall-runoff Ratios

Rainfall-runoff ratios are a second form of ratio that has been widely used in intersite comparisons, but only at certain time scales (Table 2). This kind of comparison allows examination of the fraction of water released from the basin relative to precipitation inputs, over time or among sites. For example, Post and Jones (2001) used mean annual runoff ratios (streamflow divided by precipitation) as a basis for comparison across sites. Annual runoff ratios also can be used as dependent variables in more complex analyses involving ratios. For example, Post and Jones (2001) conducted an intersite comparison based on the slopes of annual rainfall-runoff regressions across sites for a group of small, forested basin sites. Using annual runoff ratios in long (up to 50 year) records, they were able to show that forested basins respond to interannual variation in precipitation by taking up more or less water, and water use was similarly responsive to interannual variations in precipitation for all four forest types (Post and Jones, 2001).

As these examples illustrate that ratios of streamflow in treated versus control basins are the most commonly used measures for assessing treatment effects in paired-basin experiments. However, ratios also have the potential to be used for comparisons among basins that have not been subjected to formal treatments, but whose behavior may be diverging as the result of inadvertent treatments (e.g. Table 1). Inadvertent treatments in such “out of control” basins may include forest mortality from disease or disturbances like windthrow or fire, climate change, or changes in species composition and leaf area associated with forest growth and succession. The potential for learning about rainfall-runoff relationships by applying standard treated/control comparison methods to pairs of control basins (which may actually be "out of control") deserves further exploration.

Runoff ratios have potential uses that have not been explored. For example, an intersite comparison could examine the change over time in the runoff ratio (e.g. Kokkonen *et al.*, 2004), and compare this measure across basins. Or, runoff ratios could be calculated for finer than annual time periods, such as seasons or even days. This latter approach resembles the calculations involved in water balance estimates.

Bivariate Analysis – Comparison Using Cross-correlation

The cross-correlation is a form of the rainfall-runoff ratio that takes into account the temporal lag between precipitation input and streamflow output in a basin, or the lag in rainfall or runoff between one basin and another (Table 2). Cross-correlations of rainfall versus runoff at a basin estimate the lags involved between input and output in a given basin, while cross-correlations of runoff (or precipitation) across sites can be used to determine the delays in transmission of a pulse of rainfall or runoff through a stream network. For example, Post and Jones (2001) used cross-correlation of daily and monthly precipitation versus streamflow as a measure for comparing across small, forested basins. Streamflow responses to precipitation lasted for only a day or two at deciduous forest sites with mesic climates and shallow soils, but they persisted for almost two weeks at conifer sites with a Mediterranean climate and deep soils. Bond *et al.* (2002) correlated hourly sapflow with streamflow at various lags to compare across two adjacent small basins with old-growth and young conifer forest (Andrews, Oregon, US). They were able to show that diurnal variations in streamflow were most pronounced and occurred within a few hours of maximum sapflow in the early summer, but, by late summer, the relationship had diminished and the lag between maximum sapflow and streamflow had lengthened to eight hours. Moore (2003) correlated hourly precipitation, soil moisture, sapflow, and streamflow at various lags to extend this comparison of two, small, adjacent forested basins. She showed that diurnal streamflow variations were related to soil moisture, sapflow, and vapor pressure. However, strong lagged relationships were apparent during only one of several periods in three successive summers. Tague and Grant (2004) used cross-correlation of streamflow in two, large, adjacent basins to demonstrate how contrasting geology produced consistent lags in runoff at the basin whose geology permitted longer storage and slower transmission of water.

SUMMARY

Despite the limitations of rainfall-runoff data, there are compelling reasons for hydrologists to conduct many more intersite comparisons of rainfall-runoff data. Studies of rainfall-runoff processes in hydrology have generated many important predictions. Our inferences about hydrologic processes are drawn from an unnecessarily narrow subset of temporal scales, spatial scales, and geographic conditions, given the range of data available. Too many publications simply repeat (or challenge) the hypotheses inferred from a few analyses of primary data, without exploring unexamined datasets. However, verification of these hypotheses depends on replicating them using rainfall and streamflow records from locations and time periods not included in the original studies. Many thousands of basin-years of data from a wide variety of ecosystems, climate types, and basin sizes have been collected, and much of these data are now available on the Internet. However, only a fraction of these data have been analyzed. Existing published analyses of these records do not address all of the questions of current interest to hydrologists and researchers from related disciplines, including climate science, geomorphology, and ecology. New questions, novel combinations of basins, new statistical approaches, and ancillary data can all be used to more carefully test important hypotheses regarding hydrologic mechanisms operating inside basins. The endeavor of intersite comparison is analogous and complementary to the parallel quest for methods of parameter identification and model structure selection in hydrologic modeling.

Acknowledgment

This work has been supported by grants to the H.J. Andrews Long-term Ecological Research program, and for Intersite Hydrology comparisons, from the Long-term Studies program of the National Science Foundation, and by a Bullard Fellowship at Harvard Forest. Younes Alila provided helpful comments on a draft of this manuscript.

REFERENCES

Andreassian V. (2004) Waters and forests: from historical controversy to scientific debate. *Journal of Hydrology*, **291**, 1–27.

Andreassian V., Parent E. and Michel C. (2003) A distribution-free test to detect gradual changes in watershed behavior. *Water Resources Research*, **39**(9), 1252–1262.

Beschta R.L., Pyles M.R., Skaugset A.E. and Surfleet C.G. (2000) Peakflow responses to forest practices in the western cascades of Oregon, USA. *Journal of Hydrology* **233**, 102–120.

Beven K.J. (2000) Uniqueness of place and process representations in hydrological modelling. *Hydrology and Earth System Sciences*, **4**(2), 203–213.

Binkley D. and Brown T.C. (1993) Forest practices as nonpoint sources of pollution in North America. *Water Resources Bulletin*, **29**(5), 729–741.

Binkley D., Ice G.G., Kaye J. and Williams C.A. (2004) Nitrogen and phosphorus concentrations in forest streams of the United States. *Journal of the American Water Resources Association*, **40**(5), 1277–1291.

Bond B.J., Jones J.A., Moore G., Phillips N., Post D. and McDonnell J.J. (2002) The zone of vegetation influence on baseflow revealed by diel patterns of streamflow and vegetation water use in a headwater basin. *Hydrological Processes*, **16**, 1671–1677.

Bosch J.M. and Hewlett J.D. (1982) A review of catchment experiments to determine the effect of vegetation changes on water yield and transpiration. *Journal of Hydrology*, **55**, 3–23.

Daly C., Gibson W.P., Taylor G.H., Johnson G.L. and Pasteris P. (2002) A knowledge-based approach to the statistical mapping of climate. *Climate Research*, **22**, 99–113.

Eberhardt L.L. and Thomas J.T. (1991) Designing environmental field studies. *Ecological Monographs*, **61**(1), 53–73.

Gupta V.K. and Waymire E. (1989) Statistical self-similarity in river networks parameterized by elevation. *Water Resources Research*, **25**(3), 463–476, doi:10.1029/88WR04115.

Gupta V.K. and Waymire E. (1993) A statistical analysis of mesoscale rainfall as a random cascade. *Journal of Applied Meteorology*, **32**, 251–267.

Gupta V.K. and Waymire E. (1998) Scale invariance and regionalization in floods. In *Scale Dependence and Scale Invariance in Hydrology*, Sposito G. (Ed.), Cambridge University Press: Cambridge, pp. 88–135.

Gurevich J. and Hedges L.V. (2001) Meta-analysis: combining the results of independent experiments. In *Design and Analysis of Ecological Experiments*, Scheiner S.M. and Gurevich J. (Eds.), Oxford University Press: New York, pp. 347–369.

Harr R.D. (1983) Potential for augmenting water yield through forest practices in western Washington and western Oregon. *Water Resources Bulletin*, **19**(3), 383–393.

Hedges L.V. and Olkin I. (1985) *Statistical Methods for Meta-Analysis*, Academic Press, New York.

Hibbert A.R. (1967) Forest treatment effects on water yield. In *International Symposium on Forest Hydrology*, Sopper W.E. and Lull H.W. (Eds.), Pergamon, Oxford.

Hornbeck J.W., Bailey S.W., Buso D.C. and Shanley J.B. (1997) Streamwater chemistry and nutrient budgets for forested watersheds in New England: variability and management implications. *Forest Ecology and Management*, **93**, 73–89.

Jones J.A. (2000) Hydrologic processes and peak discharge response to forest removal, regrowth, and roads in ten small experimental basins, western Cascades, Oregon. J.A. Jones. *Water Resources Research*, **36**(9), 2621–2642.

Jones J.A. and Grant G.E. (1996) Peak flow response to clearcutting and roads in small and large basins, western Cascades, Oregon. *Water Resources Research*, **32**, 959–974.

Jones J.A. and Post D.A. (2004) Seasonal and successional streamflow response to forest cutting and regrowth in the northwest and eastern United States. *Water Resources Research*, **40**, W05203, doi:10.1029/2003WR002952.

Kirchner J.W., Feng X. and Neal C. (2000) Fractal stream chemistry and its implications for contaminant transport in catchments. *Nature*, **403**, 524–527.

Kokkonen T., Koivusalo H., Karvonen T., Croke B. and Jakeman A. (2004) Exploring streamflow response to effective rainfall across event magnitude scale. *Hydrological Processes*, **18**(8), 1467–1486.

Lafrenière M. and Sharp M. (2003) Wavelet analysis of inter-annual variability in the runoff regimes of glacial and nival stream catchments, Bow Lake, Alberta. *Hydrological Processes*, **17**, 1093–1118.

Martin C.W., Noel D.S. and Federer C.A. (1984) Effects of forest clearcutting in New England on stream chemistry. *Journal of Environmental Quality* **13**(2), 204–210.

Moore G.W. (2003) *Drivers of Variability in Transpiration and Implications for Stream Flow in Forests of Western Oregon*, PhD. dissertation, Forest Science, Oregon State University.

Pierson F.B., Slaughter C.W. and Cram Z.K. (2001) Long-term stream discharge and suspended-sediment database, Reynolds Creek Experimental Watershed, Idaho, United States. *Water Resources Research*, **37**(11), 2857–2861, doi:10.1029/2001WR000420.

Post D.A., Grant G.E. and Jones J.A. (1998) Ecological hydrology: expanding opportunities in hydrologic sciences. *EOS*, **79**(43), 517, 526.

Post D.A. and Jones J.A. (2001) Hydrologic regimes at four long-term ecological research sites in New Hampshire, North Carolina, Oregon, and Puerto Rico. *Advances in Water Resources*, **24**(9–10), 1195–1210.

Robinson M., Cognard-Plancq A.-L., Cosandey C., David J., Durand P., Fuhrer H.-W., Hall R., Hendriques M.O., Marc V., McCarthy R., *et al.* (2003) Studies of the impacts of forests on peak flows and baseflows: a European perspective. *Forest Ecology and Management*, **186**, 85–97.

Slaughter C.W., Marks D., Flerchinger G.N., Van Vactor S.S. and Burgess M. (2001) Thirty-five years of research data collection at the Reynolds Creek Experimental Watershed, Idaho, United States. *Water Resources Research*, **37**(11), 2819–2823.

Smith J.W. (2002) *Mapping the Thermal Climate of the H.J. Andrews Experimental Forest, Oregon*, MS thesis, Department of Geosciences, Oregon State University.

Stednick J.D. (1996) Monitoring the effects of timber harvest on annual water yield. *Journal of Hydrology*, **176**(1–4), 79–95.

Tague C. and Grant G.E. (2004) A geological framework for interpreting the low-flow regimes of cascade streams, Willamette River Basin, Oregon. *Water Resources Research*, **40**, W04303, doi:10.1029/2003WR002629.

Thomas R.B. and Megahan W.F. (1998) Peak flow responses to clear-cutting and roads in small and large basins, western Cascades, Oregon: A second opinion. *Water Resources Research* **34**(12), 3393–3403.