How uncertainty analysis of streamflow data can reduce costs and promote robust decisions in water management applications

McMillan, Hilary ; Seibert, Jan ; Petersen-Overleir, Asgeir ; Lang, Michel ; White, Paul ; Snelder, Ton ; Rutherford, Kit ; Krueger, Tobias ; Mason, Robert ; Kiang, Julie

DOI: https://doi.org/10.1002/2016WR020328

Originally published at:
McMillan, Hilary; Seibert, Jan; Petersen-Overleir, Asgeir; Lang, Michel; White, Paul; Snelder, Ton; Rutherford, Kit; Krueger, Tobias; Mason, Robert; Kiang, Julie (2017). How uncertainty analysis of streamflow data can reduce costs and promote robust decisions in water management applications. Water Resources Research, 53(7):5220-5228.
DOI: https://doi.org/10.1002/2016WR020328

Water Resources Research

COMMENTARY
10.1002/2016WR020328

Special Section:
Engagement, Communication, and Decision-Making Under Uncertainty

Key Points:
- Ignorance of uncertainties in streamflow data can lead to suboptimal decisions and economic costs in water management applications
- We present three case studies where uncertainty analysis reduced costs and promoted robust conclusions
- We recommend development of practical tools that generate multiple streamflow time series realizations rather than simple error bounds

Correspondence to:
H. McMillan,
hmcmillan@sdsu.edu

Citation:
McMillan, H., J. Seibert , A. Petersen-Overleir, M. Lang, P. White, T. Snelder, K. Rutherford, T. Krueger, R. Mason, and J. Kiang (2017), How uncertainty analysis of streamflow data can reduce costs and promote robust decisions in water management applications, *Water Resour. Res.*, *53*, 5220–5228, doi:10.1002/2016WR020328.

Received 12 DEC 2016
Accepted 7 JUN 2017
Accepted article online 14 JUN 2017
Published online 20 JUL 2017

How uncertainty analysis of streamflow data can reduce costs and promote robust decisions in water management applications

Hilary McMillan[1], Jan Seibert[2], Asgeir Petersen-Overleir[3], Michel Lang[4], Paul White[5], Ton Snelder[6], Kit Rutherford[7], Tobias Krueger[8], Robert Mason[9], and Julie Kiang[9]

[1]Department of Geography, San Diego State University, San Diego, California, USA, [2]Department of Geography, University of Zurich, Zurich, Switzerland, [3]Market Operations Hydrology, Statkraft Energi AS, Oslo, Norway, [4]Irstea, UR HHLY, Hydrology-Hydraulics, Villeurbanne, Lyon, France, [5]Institute of Geological & Nuclear Sciences, Taupo, New Zealand, [6]LWP Ltd., Christchurch, New Zealand, [7]National Institute of Water and Atmospheric Research, Napier, New Zealand, [8]IRI THESys, Humboldt-Universität zu Berlin, Berlin, Germany, [9]U.S. Geological Survey, Office of Surface Water, United States Geological Survey, Reston, Virginia, USA

Abstract Streamflow data are used for important environmental and economic decisions, such as specifying and regulating minimum flows, managing water supplies, and planning for flood hazards. Despite significant uncertainty in most flow data, the flow series for these applications are often communicated and used without uncertainty information. In this commentary, we argue that proper analysis of uncertainty in river flow data can reduce costs and promote robust conclusions in water management applications. We substantiate our argument by providing case studies from Norway and New Zealand where streamflow uncertainty analysis has uncovered economic costs in the hydropower industry, improved public acceptance of a controversial water management policy, and tested the accuracy of water quality trends. We discuss the need for practical uncertainty assessment tools that generate multiple flow series realizations rather than simple error bounds. Although examples of such tools are in development, considerable barriers for uncertainty analysis and communication still exist for practitioners, and future research must aim to provide easier access and usability of uncertainty estimates. We conclude that flow uncertainty analysis is critical for good water management decisions.

Plain Language Summary In this commentary, we show how analyzing uncertainty in river flow data can reduce costs and promote robust conclusions in water management applications. River flow data can contain large uncertainties but are often communicated and used without uncertainty information. We give case studies from Norway and New Zealand where flow uncertainty analysis has uncovered economic costs in the hydropower industry, improved public acceptance of a controversial water management policy, and tested the accuracy of water quality trends. We conclude that flow uncertainty analysis is critical for good water management decisions.

1. Introduction

Streamflow data are used for a wide variety of water management applications. Flow magnitude data are used to design levees, dams, and other hydraulic structures, and to delineate floodplains. Analysis of the natural flow regime is used to specify environmental flow requirements and manage water quality, and real-time flow data determine authorizations of water withdrawal or pollutant releases and serve as the basis of river forecasts and flood warnings. Flow data guide commercial management of reservoirs and dams, and thus affect economic returns.

When flow data are imprecise or inaccurate, unnecessary costs or poor ecological or societal outcomes may follow. For example, overestimation of a design flood would lead to higher bridge building costs, whereas underestimation would lead to higher failure risks. When extracting stream water for irrigation, the balance between economic benefits and ecological flow requirements might be wrongly estimated if streamflow data are uncertain. Such costs can be managed by treating flow as an uncertain variable by specifying a measurement model, using this together with appropriate tools for decision-making under uncertainty, and investing in data collection to reduce uncertainty where economically justified (e.g., by collecting additional high flow measurements or directly measuring water velocity as well as stage [*Levesque and Oberg*, 2012]).

Streamflow data used by commercial and governmental organizations are often treated as "error-free" information. However, these data are the products of models (most often the stage-discharge rating curve) that may be imprecise and/or biased and may thus produce inaccurate flow information even under best-practice stage and flow measurement protocols [*Petersen-Øverleir et al*., 2009; *Coxon et al*., 2015]. Flow data are typically provided without accompanying uncertainty estimates, or with only very general guidance such as $+/-$ uncertainty bounds or flagging of suspect data. Data users, therefore, have little opportunity to account for flow uncertainty, apart from pragmatic approaches such as ignoring flagged data, or using a wide safety margin. Uncertainty analysis may also be perceived as arduous and expensive, with results that could add doubt to reported conclusions. If the imprecision and inaccuracy of the rating curve model were better quantified, users would be able to better use the flow data in their predictions and decision making.

We argue in this commentary that including uncertainty analysis as a standard part of applied water management applications can save money and increase the robustness and perceived trustworthiness of decisions. We discuss the causes and magnitude of uncertainty in flow data. Through a series of case studies, we demonstrate the economic consequences of inaccurate flow data, and the improvements in economic, environmental, and societal outcomes when uncertainty analyses are included. These case studies describe previous experiences of the authors, and are in part drawn from examples presented at a Workshop on Discharge Uncertainty Analysis, held at TU Vienna, Austria, April 2016. We conclude by discussing current challenges and recommendations for operational, practical flow uncertainty analysis.

2. Uncertainty Sources and Magnitudes in Flow Data

Most flow data are derived from a measured time series of river stage, and a rating curve model that relates stage to flow. The resulting flow data are, therefore, estimates that contain uncertainties relating to: (1) the measured stage series, (2) the measurements of stage and flow used to derive the rating curve, (3) interpolation and extrapolation in the rating curve model, and (4) unrecognized changes in the channel cross section due to scour/fill, vegetation and ice, backwater and hysteresis effects, which cause change in the deterministic rating curve. The commonly used term "flow data uncertainty" refers to these uncertainty sources in the derived flow estimate. This uncertainty may also be understood as part of a "measurement model" that relates measured stage to an estimate of the true flow, and should include both random and systematic components of uncertainty. We must acknowledge that our measurement model will always be incomplete, and sometimes we do not have much evidence upon which to base our measurement model.

Streamflow uncertainty analysis methods attempt to quantify the combined magnitude of the four uncertainty sources outlined above. Uncertainties in measurements of stage and discharge relate to instrument precision, fluctuations in water surface height, and estimation of average cross-section discharge from a series of discrete measurements. Uncertainties in the rating curve combine known uncertain approximations (interpolation/extrapolation) with unknown systematic errors due to cross-section change. While the latter could theoretically be minimized by frequent measurements and reassessments of the rating curve model, in practice, these changes can be more frequent (and unpredictable) than limited measurement budgets can cover. Therefore, uncertainty analysis techniques identify the magnitude of likely systematic errors based on changes between previous measurements, and treat these systematic errors as an uncertainty source in the measurement model.

McMillan et al. [2012] provide a thorough overview of uncertainty sources and their typical magnitudes in flow data, including for alternative flow measurement techniques such as Acoustic Doppler instruments. Typical confidence bounds for flow uncertainties when using the rating curve method were found to be $\pm$50–100% for low flows, $\pm$10–20% for medium or high (in-bank) flows, and $\pm$40% for out of bank flows. However, the bounds for any particular rating are highly dependent on gauge construction, measurement regime, and channel characteristics [*Westerberg et al*., 2016]. If flow estimates for ungauged catchments are derived from a hydrologic model, further uncertainties in rainfall inputs, model structure, and parameters have to be considered.

3. Case Studies

In the following section, we describe three water management case studies where analyses of flow uncertainties were critical to the accuracy and robustness of the results.

3.1. Economic and Environmental Costs of Flow Uncertainties

Case study 1: Rating curves in the Norwegian hydropower industry. *In this case study, we demonstrate high economic costs associated with errors in flow data used to maintain minimum environmental flows*.

The Norwegian Water Resources and Energy Directorate legislates that hydropower companies must maintain hydrometric networks to monitor minimum environmental flow releases in regulated rivers, and to monitor nearby natural rivers to assess the comparative effects of dams and abstraction. These minimum flows mitigate the impacts of flow-regulation by preserving biological diversity and landscape quality, accommodating the needs of other water users, diluting pollutants, and maintaining groundwater levels. Waterfall flows are regulated to protect the majestic landscape valued by Norwegians: Vøringsfossen (Eidfjord, Hordaland) and Mardalsfossen (Nesset, Møre, and Romsdal) described below are protected waterfalls managed by the hydropower company Statkraft. Governing authorities can impose penalty fees for insufficient flows resulting from reservoir management operations, but environmental consequences are difficult to quantify using economic measures. Releasing too much water has a quantifiable economic cost for the power producer.

The minimum flow required at Vøringsfossen is 11 m^3/s during the summer holiday period 1 June–15 September. Most of this flow is released from a large reservoir upstream, especially during dry spells. A gauging station just upstream of the waterfall measures the discharge and controls water releases. A typical rating curve error of 1 m^3/s, combined with plausible power prices, could result in an economic loss of €5000/d. In 2010, an error of this magnitude was suspected, caused by erosion of the hydraulic control. A new rating curve was quickly measured, but, assuming the error had persisted throughout the season, the estimated economic losses were approximately €500,000. The new rating curve used eight measurements in the range 7.4 to 14.2 m^3/s. Rating measurements were limited to this interval because the main purpose of the gauging station is to ensure that the discharge is 11 m^3/s or higher, and therefore, the accuracy of extrapolated flows is not of great concern to the hydropower company or the regulator.

Using the Bayesian rating curve analysis of *Reitan and Petersen-Øverleir* [2009] with default priors, the 95% credible interval at 11 m^3/s is about 5%. Given the potential costs for the power producer of a 5% rating error, investment in further measurements to improve the rating curve precision may have positive benefit compared to cost. For comparison, the costs of developing a new rating curve are approximately €10,000, consisting of one day's work for a hydrometric consultant team to perform the discharge measurements (fee typically around €3000) and the cost of stepped water releases over the desired flow interval (average additional release of 3 m^3/s for 10 h; cost approximately €7000).

The framework governing water releases at the second waterfall, Mardalsfossen is complicated, but an error of 0.3 m^3/s is plausible based on the rating curve quality. In the case of underestimation, this could result in total costs for the power producer of up to €100,000 per summer season. This cost was calculated based on a recent system average power price of 30 €/MWh in conjunction with a calculation of the energy equivalent of water volume in the upstream reservoir. In the case of overestimation, a fine for insufficient summer releases could occur such as that of 1986 when Statkraft was fined €1.5M (€3M equivalent today). The Mardalsfossen rating curve is already high quality, with an estimated 95% credible interval of 3% for the regulated minimum flows of 2, 2.5, and 3 m^3/s; but again, it might be economically beneficial to improve the rating curve further given the high economic costs and penalty fees involved.

3.2. Robust Conclusions in Water Management Applications

Case study 2: Determination of a catchment boundary for Lake Rotorua, New Zealand. *In this case study, we demonstrate how transparent uncertainty estimates improved public acceptance of a controversial water management policy*.

Lake Rotorua is an 80 km^2 freshwater lake in the North Island of New Zealand. The local government body intends to control land use in the lake catchment to reduce nitrogen and phosphorus inflows and improve water quality. The management strategy will have significant economic impact on land owners within the catchment boundary. For example, if rules restricting nitrogen losses from land lead to a land parcel being used for dryland sheep/beef farming rather than irrigated dairy farming, the profit per hectare could fall from US$2900 to US$200 (estimates from New Zealand South Island) [*Saunders and Saunders*, 2012]. Determining an accurate catchment boundary is therefore important to the strategy's success. While the surface catchment boundary can be accurately mapped, the lake also has groundwater inflows originating from outside the surface catchment.

Analysis of 13 years of recorded flow at the lake outlet and at four streams flowing into the lake was used to quantify the size of the "missing" groundwater catchment area [*Rutherford and Palliser*, 2014a]. Annual runoff, as a fraction of rainfall, was calculated for each flow record, along with its uncertainty (requiring estimates of the uncertainty in flow, rain, and catchment area). By comparing the runoff fraction in the streams without groundwater influence to the catchment outlet runoff fraction, the additional outlet runoff was calculated and used to estimate the "missing" area. It was important to test whether the additional runoff could be attributed to uncertainty in the measured flows. Flow uncertainty was calculated using two different methods, recognizing that deviations of gaugings from the rating curve can represent systematic error (i.e., the rating curve was incorrect or had changed with time), and/or random error in the measurements of stage and flow used to derive the rating curve. For a worst case scenario, all error was assumed systematic, and so average error was calculated as the mean difference between gaugings and rating curve. For a best case scenario, errors were assumed to be a combination of systematic and random components. An empirical estimate of rating uncertainty and its autocorrelation in time was used to estimate a total uncertainty distribution.

Even with worst-case uncertainty, there was only a 33% chance of the observed flow differences occurring if there was no groundwater inflow. Therefore, in a second step, a catchment water balance model, together

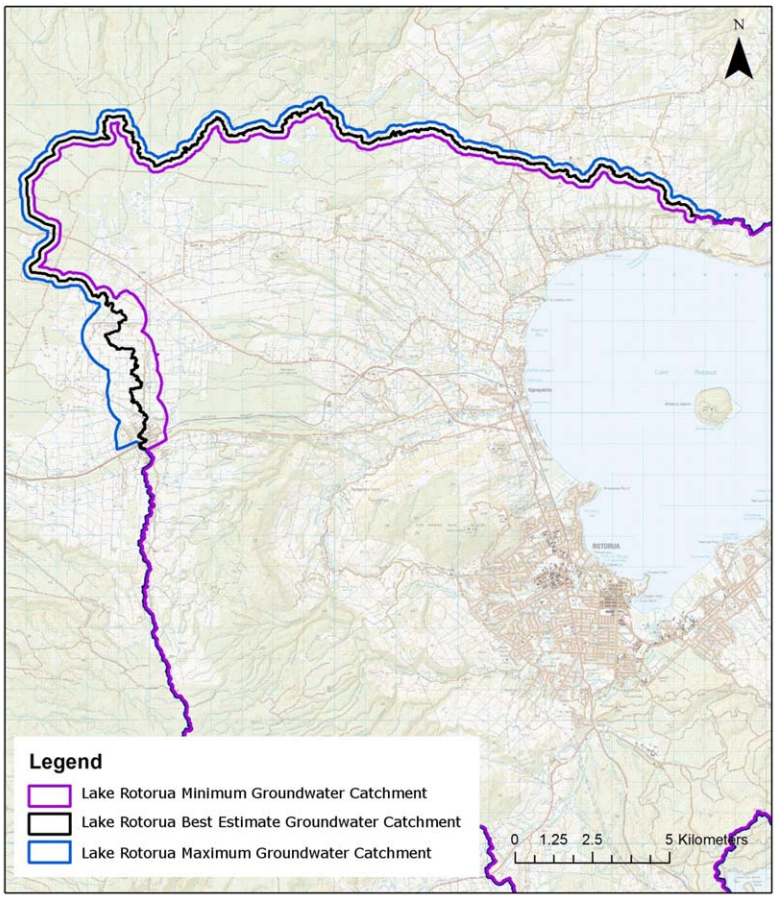

Figure 1. Map of the minimum, best estimate, and maximum Lake Rotorua groundwater catchment boundaries in the Mamaku Plateau area. Figure reproduced from *White et al*. [2014] with the permission of Bay of Plenty Regional Council.

with local hydrogeological knowledge, was used to estimate the location of the missing area [*White et al*., 2014; *Rutherford and Palliser*, 2014b]. Discharge uncertainty estimates in the outlet and streams were used to provide inner and outer uncertainty limits for the location of the groundwater catchment boundary; see Figure 1 for an example. Maps of these uncertainty limits were provided to the public and to the governing body, demonstrating the robustness of the groundwater boundary estimate, which was essential given the high economic impact of imposed land-use change. Public hearings on the land management rule changes are in progress, and have not included any submissions challenging the calculated catchment boundary positions.

Case study 3: Water quality trends in ungauged catchments, New Zealand. *In this case study, we illustrate how information on errors in flow values was used to test the reliability of water quality trends that used flow estimates from a hydrologic model in the absence of gauged flow data.*

Water quality is of concern in many New Zealand rivers and lakes, due to increased levels of nitrogen, phosphorus, pathogenic micro-organisms, and other agricultural emissions. The New Zealand government therefore commissioned a national-scale water quality analysis to quantify current state and future pressures on water quality. Water quality measurements from around the country were compiled and analyzed to determine state and trends in water quality variables at measurement sites.

Water quality is commonly strongly dependent on flow, and therefore a flow-adjustment procedure is preferable prior to trend analysis [*Hirsch et al*., 1982]. Flow adjustment reduces the possibility that, if samples at the beginning or end of the time period are biased toward high or low flows, a trend can be induced that reflects the flow rather than a real change in the concentration. The flow adjustment procedure fits a relationship between flow and the water quality variable, such that the trend analysis is computed only on the residuals. However, in the New Zealand case, 547 out of 785 water quality measurement sites had no flow data available. In this case study, we investigated whether these sites could be included in the national analysis by substituting flow estimates from a hydrologic model. The alternative to using the modeled flows for flow adjustment would be to use the "raw" (i.e., nonflow adjusted trends), which risks incorrectly detecting trends as explained above.

Trend classification uncertainty due to errors in flow was analyzed by comparing the two sets of trends (i.e., trends calculated using the observed versus modeled flows). This allowed us to explore how error in flow values (here due to substituting measured flow with flow estimates from a hydrologic model) affects the reliability of subsequent analyses. Although this is not a full uncertainty analysis (e.g., quantifying the distribution of total uncertainty in modeled flow), it demonstrates how information on errors in flow values can be used to make inferences about the reliability of the final classification of trend direction.

The New Zealand national hydrologic model [*McMillan et al*., 2016] was used to estimate mean daily flow for sites that lacked flow measurements. However, the national hydrologic model was originally designed for large-scale estimates of water stores and fluxes. It was therefore important to establish whether single-day estimates of flows at water quality measurement sites were sufficiently accurate for flow-adjustment use. This uncertainty analysis was unusual in that error in flow magnitude (which can be large) is only an intermediate variable, and our final aim is to detect errors in the correct classification of water quality trend into three categories (improving/insufficient data/degrading). We therefore used the set of sites/dates where flow data were available to quantify errors in the model prediction of flow, compared to measured flow. The contribution of these flow errors to errors in trend classification was calculated by repeating the full trend analysis using (a) water quality data adjusted using measured flow data, (b) water quality data adjusted using model flow data, and (c) water quality data not adjusted for flow.

Trends were assessed in three steps. (1) Flow-adjustment: a second-order polynomial was fitted to the log of the measured values of the water quality variable and the log of the associated flow using a generalized additive model (GAM). (2) Trend analysis: the Seasonal Kendall Slope Estimator [*Hirsch et al*., 1982] was used to determine whether the trend direction could be confidently inferred from the flow adjusted data and to estimate the trend magnitude. This analysis was carried out for six water quality variables: clarity, ammonium, nitrate, total nitrogen, dissolved reactive phosphorus, and total phosphorus.

Results of the uncertainty analysis showed that for the 20 year time period ending at the end of 2013, when model flow data were used instead of measured flow data, the trend direction was correctly classified for between 82% and 95% of trends (mean 90%), depending on which water quality variable was tested. When raw trends were used, the direction was classified correctly for a lower mean rate of 86% of trends. The error rate in trend classification resulting from model flow uncertainty was considered sufficiently small that it was reasonable to proceed with using model flow estimates at ungauged water quality measurement sites. The uncertainty analysis results were presented to the New Zealand government alongside the analysis of water quality trends [*Larned et al*., 2015, 2016], to allay concerns that hydrological model predictions may not be sufficiently accurate.

4. Discussion

4.1. Wide Influences of Flow Uncertainty

We presented three case studies here, but there are numerous other water management applications that are strongly influenced by flow uncertainty. In some cases, the lack of uncertainty assessment has led to erroneous conclusions. For example, *Lang et al*. [2010] discuss a trend analysis of French floods and droughts, at 195 gauging stations with long records and low human influences [*Lang et al*., 2006]. An initial analysis showed a significant trend in the annual maximum flood series for 18% of the sites, using a likelihood ratio test based on the generalized extreme value distribution. However, a more thorough analysis showed that 49% of the significant trends were explained by errors in the flow data, such as changes caused by revisions in the rating curve extrapolation. Analysis of the remaining 124 gauges without such problems found fewer changes, and these were assessed as nonsignificant by *Renard et al*. [2008].

A further analysis of flood magnitudes in southern France by *Lang et al*. [2010] asked whether initial flood estimates using data since 1980 could be constrained by including historical records. Results showed that adding historical information without considering flow uncertainty led to biased and misleadingly small flood discharge credible intervals. When uncertainties were included, adding manual records since 1892 reduced uncertainty, but adding historical flood estimates since 1741 increased uncertainty because rating curve errors in old data were significantly higher.

Even in water-scarce regions such as the Southwest United States where the value of accurate streamflow information is widely accepted, flow-record uncertainty, and the potential inefficiency that accompanies it is under appreciated. For example, because Colorado River water is over allocated amongst the States of California, Colorado, Nevada, and Utah, the US Supreme Court, in its 1963 Decree, required the monitoring of all diversions and returns of water from and to the Colorado River but the Court said nothing about the required accuracy of the streamflow records. That oversight is a significant consideration: streamflow records for the Colorado River at Lees Ferry, Arizona, are rated excellent—generally thought to be within 5% of true flow 95% of the time. Even so, the cumulative annual uncertainty exceeds the entire volume of Colorado River water appropriated to Utah (typically about 3% of the total annual flow at Lees Ferry).

When water has high economic value, early inclusion of uncertainty analysis allows decision-making to properly account for uncertainty in the flow data. For a review of the role of uncertainty analysis in risk-type decision-making see *Hall and Solomatine* [2008]. Theories of decision-making can include both quantitative economic and sociological perspectives [*Johnson and Busemeyer*, 2010], and can account for multicriteria current and future risks [*Marinoni et al*., 2011]. In Norway, in addition to the minimum waterfall flows discussed in our Case Study 1, flow data are used to characterize flow regimes and hence design hydropower systems. In these systems, inaccurate rating curves can cause large costs due to long-lasting nonoptimal energy management. Flow data are further used to calibrate hydrologic models, which provide the forecasts of reservoir inflow needed by energy management models. Unrecognized data uncertainty can lead to water use for energy production during nonoptimal periods. General economic figures on possible losses due to forecast imprecision are difficult to quantify, but potential losses of millions of Euros are a realistic estimate in Norway.

4.2. Challenges and Recommendations in Communicating Flow Uncertainty

Flow uncertainty analysis is critical for comparing the value of additional hydrologic data for water management applications against the time and cost involved in collecting these data. But uncertainty analysis may not be taken up by users unless it can be applied to the most frequently used data types (generally, real-

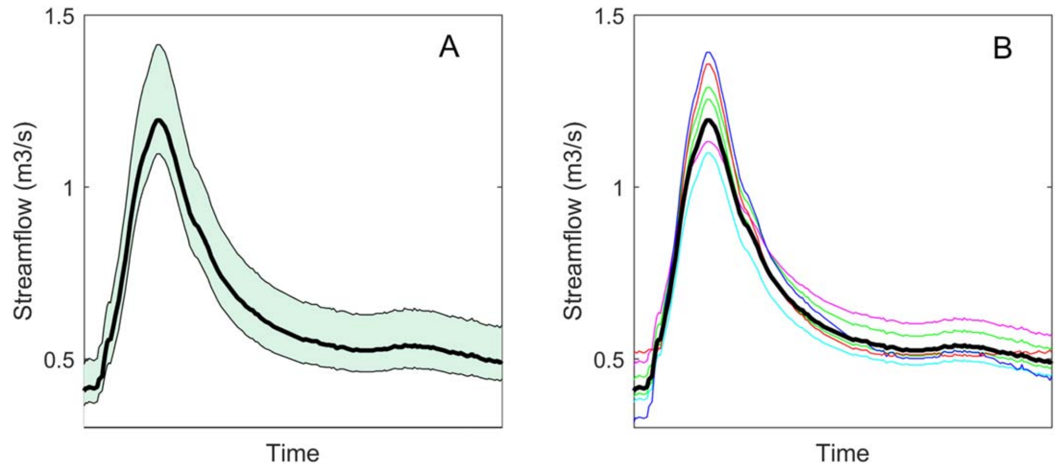

Figure 2. Two methods of presenting flow uncertainty: (a) as bands around a best estimate discharge and (b) in the form of random samples from the distribution of possible flow series.

time information) and is sensitive to differences in data collection costs and accuracies between sites. The U.S. Geological Survey assessed the cost-effectiveness of its national streamgage network from 1982 to 1986 and identified optimization strategies to minimize overall uncertainty of the network annual flow estimates given a stated fiscal budget [*Thomas and Wahl*, 1993]. However, because the optimization focused on annual flow estimates and not real-time data, and because the analysis was insensitive to flow variation at specific sites, the results lacked practical impact for users motivated by specific flow regimes, were not well received, and were never fully implemented.

The efforts required for a full uncertainty analysis can be large and overwhelming for many users. Such analyses require access to raw gauging data and stage records, which varies by country and recording authority, and is usually not as straightforward as obtaining processed flow series. Readily available flow uncertainty estimates are therefore valuable. Typically, if available at all, these are presented as bands around a best estimate discharge, e.g., Bayesian credible intervals around the median (Figure 2a). This is problematic because if the users are only given uncertainty information in the form of bands, they cannot infer information on error correlation in time (typical for rating curve models that contain unknown systematic errors as well as random errors). They therefore cannot specify whether flow data exhibit consistently high or low bias, or a combination such as overestimation of low flows and underestimation of high flows. All this extra information is required if flow uncertainty analysis is to be followed by further uncertainty analysis of, e.g., trends, hazard estimates, or economic losses, typical of the case studies in this paper.

Such further analyses usually require uncertainty estimates in the form of random samples from the distribution of possible flow series, drawn from different realizations of the rating curve [*Westerberg and McMillan*, 2015]. These "spaghetti plots" (Figure 2b) of multiple possible flow series retain the full error structure of the stage-flow rating model, and enable analyses to be repeated on each flow series, and the results aggregated to give an uncertainty distribution of the variable of interest. As an example from dam safety, the standard approach is to route a design flood toward a dam with some initial water level to check whether the dam crest is overtopped. An alternative approach is to simulate several thousand possible flood events and initial water levels from estimated probability distributions, computing the final dam water level each time, and therefore directly assessing the probability distribution of levels [*Arnaud et al*., 2015]. This overcomes the arbitrary choice of an initial level by properly combining the uncertainty of both initial condition and flood event. Similarly, in hydrologic modeling, one needs full rating curve error distributions for streamflow in order to obtain unbiased model calibrations [*McMillan et al*., 2010; *Engeland et al*., 2016].

We therefore recommend promotion of practical, operational tools that enable straightforward generation of multiple possible "spaghetti" flow series that are consistent with the measured stage series, and take into account uncertainty in the rating curve model of the stage-flow relationship. Many such flow uncertainty analysis methods exist [e.g., *Coxon et al*., 2015; *Le Coz et al*., 2014; *Mason et al*., 2016; *Morlot et al*., 2014; *Reitan and Petersen-Overleir*, 2008; *Sikorska et al*., 2013; *Westerberg and McMillan*, 2015] and an experimental comparison project is ongoing to compare differences in assumptions and results. We encourage flow data providers to consider integrating a software tool at the data download stage to calculate discharge realizations based on water level series and rating curve shape or parameter distributions stored for each gauging station. Such a method would avoid large data sets associated with direct storage of multiple flow series. The tool could also aggregate standard statistics such as monthly mean and extreme values (and their uncertainty bounds). We know of two similar tools made available online by hydrologic researchers: BaRatin from IRSTEA, https://forge.irstea.fr/projects/baratinage_v2/news; and HydraSub from the University of Oslo, https://folk.uio.no/trondr/hydrasub/ratingcurve.html. Our proposal requires commitment from the hydrologic community to continue work alongside hydrometrists to develop generalizable and easily accessible methods for rating curve uncertainty estimation across large numbers of rating sites.

5. Conclusion

The significant costs associated with the streamflow uncertainties presented in this paper show that estimation of uncertainty as part of a water management application can reduce project costs and lead to more robust and publicly acceptable decisions. In addition to explicit presentation of flow uncertainty bounds, uncertainty analysis can be extended to derived variables such as flood volume, and can be used to link the hydrologic/hydraulic uncertainties with an economic model (e.g., minimization of the expected cost, considering flow uncertainty). Rationalization and optimization of hydrometric networks can also benefit from cost-benefit analysis accounting from flow uncertainty. Our results show that it is often economically beneficial to invest in further streamflow gaugings to reduce the potential costs of rating curve error. However, even with excellent data sets it is not possible to eliminate flow uncertainty. We recognize that the complexities of uncertainty analysis (especially under data-poor conditions) and difficulties in communicating uncertainty results can present barriers for practitioners. We therefore call for a commitment from the hydrologic community to develop practical tools to support more widespread analysis of flow uncertainty.

Acknowledgments
This commentary is based on discussions at a Workshop on Discharge Uncertainty Analysis, held at TU Vienna, Austria, April 2016, and draws from discussions at the U.S. Geological Survey John Wesley Powell Center for Analysis and Synthesis. The work benefited from the framework of the Panta Rhei Research Initiative of the International Association of Hydrological Sciences (IAHS), working groups "Hydrologic Services and Hazards in Multiple Ungauged Basins" and "Epistemic Uncertainty." Data on Lake Rotorua catchment boundaries are owned by Bay of Plenty Regional Council (https://www.boprc.govt.nz/) and we thank them for permission to reproduce Figure 1. Data on New Zealand water quality and trends are available from Land Air Water Aotearoa (LAWA; https://www.lawa.org.nz/). We thank Grey Nearing, Hoshin Gupta, and one anonymous reviewer for their constructive comments on the manuscript. Grey in particular provoked us to connect our data uncertainty language better with the alternative language of measurement models.

References

Arnaud, A., E. Paquet, Y. Aubert, P. Royet, J. A. Fine, and M. Lang (2015), Assessment of the distribution of water level within a reservoir using a stochastic rainfall-runoff generator. Case study on a fictive reservoir, in *25th ICOLD Meeting, Norway, June 2015* [in French], Stavanger.

Coxon, G., J. Freer, I. K. Westerberg, T. Wagener, R. Woods, and P. J. Smith (2015), A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations, *Water Resour. Res.*, *51*, 5531–5546, doi:10.1002/2014WR016532.

Engeland, K., et al. (2016), Effects of uncertainties in hydrological modelling. A case study of a mountainous catchment in Southern Norway, *J. Hydrol.*, *536*, 147–160.

Hall, J., and D. Solomatine (2008), A framework for uncertainty analysis in flood risk management decisions, *Int. J. River Basin Manage.*, *6*(2), 85–98.

Hirsch, R. M., J. R. Slack, and R. A. Smith (1982), Techniques of trend analysis for monthly water quality data, *Water Resour. Res.*, *18*, 107–121.

Johnson, J. G., and J. R. Busemeyer (2010), Decision making under risk and uncertainty. Wiley Interdisciplinary Reviews, *Cogn. Sci.*, *1*(5), 736–749.

Lang, M., B. Renard, E. Sauquet, P. Bois, A. Dupeyrat, C. Laurent, O. Mestre, H. Niel, L. Neppel, and J. Gailhard (2006), A national study on trends and variations of French floods and droughts, IAHS Publ. 308, pp. 514–519, Int. Assoc. of Sci. Hydrol., Gentbrugge, Belgium.

Lang, M., K. Pobanz, B. Renard, E. Renouf, and E. Sauquet (2010), Extrapolation of rating curves by hydraulic modelling, with application to flood frequency analysis, *Hydrol. Sci. J.*, *55*(6), 883–898 [Special Issue, "Court of Miracles of Hydrology," Engref Paris, 18–20 June 2008].

Larned, S., T. Snelder, M. Unwin, G. McBride, P. Verburg, and H. McMillan (2015), Analysis of water quality in New Zealand lakes and rivers: Data sources, datasets, assumptions, limitations, methods and results, *Prepared for Ministry for the Environment. NIWA Tech. Rep. CHC2015–033*, Wellington, NZ.

Larned, S. T., T. Snelder, M. Unwin, and G. B. McBride (2016), Water quality in New Zealand Rivers: Current state and trends, *N. Z. J. Mar. Freshwater Res.*, *50*(3), 389–417, doi:10.1080/00288330.2016.1150309.

Le Coz, J., B. Renard, L. Bonnifait, F. Branger, and R. Le Boursicaud (2014), Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: A Bayesian approach, *J. Hydrol.*, *509*, 573–587, doi:10.1016/j.jhydrol.2013.11.016.

Levesque, V. A., and K. A. Oberg (2012), Computing discharge using the index velocity method, *U.S. Geol. Surv. Tech. Methods 3-A23*, 148 pp. [Available at http://pubs.usgs.gov/tm/3a23/.]

Marinoni, O., P. Adkins, and S. Hajkowicz (2011), Water planning in a changing climate: Joint application of cost utility analysis and modern portfolio theory, *Environ. Modell. Software*, *26*(1), 18–29.

Mason, R. R., J. E. Kiang, and T. A. Cohn (2016), Rating curve uncertainty: A comparison of estimation methods, in *River Flow 2016*, edited by G. Constantinescu et al., pp. 251–252, Taylor and Francis, London.

McMillan, H., J. Freer, F. Pappenberger, T. Krueger, and M. Clark (2010), Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions, *Hydrol. Processes*, *24*(10), 1270–1284.

McMillan, H., T. Krueger, and J. Freer (2012), Benchmarking observational uncertainties for hydrology: Rainfall, river discharge and water quality, *Hydrol. Processes*, *26*(26), 4078–4111.

McMillan, H. K., D. J. Booker, and C. Cattoën (2016), Validation of a national hydrological model, *J. Hydrol.*, *541*, 800–815.

Morlot, T., C. Perret, A.-C. Favre, and J. Jalbert (2014), Dynamic rating curve assessment for hydrometric stations and computation of the associated uncertainties: Quality and station management indicators, *J. Hydrol.*, *517*, 173–186.

Petersen-Øverleir, A., A. Soot and T. Reitan (2009), Bayesian rating curve inference as a streamflow data quality assessment tool, *Water Res. Manag.*, *23*(9), 1835–1842.

Reitan, T., and A. Petersen-Øverleir (2009), Bayesian methods for estimating multi-segment discharge rating curves, *Stochastic Environ. Res. Risk Assess.*, *23*(5), 627–642.

Renard, B., et al. (2008), Regional methods for trend detection: Assessing field significance and regional consistency, *Water Resour. Res.*, *44*, W08419, doi:10.1029/2007WR006268.

Rutherford, K., and C. Palliser (2014a), Lake Rotorua catchment boundaries—Phase 2 water budget, *NIWA Tech. Rep. Prepared for GNS and Bay of Plenty Reg. Counc., Rep. HAM2014–045*, 35 pp., Whakatāne, NZ.

Rutherford, K., and C. Palliser (2014b), Lake Rotorua catchment boundaries—Phase 3 water budget calculations, *NIWA Tech. Rep. Prepared for GNS and Bay of Plenty Reg. Counc., Rep. HAM2014–048*, 27 pp., Whakatāne, NZ.

Saunders, C., and J. Saunders (2012), The economic value of potential irrigation in Canterbury, 42 pp., Comm. Rep. for Lincoln Univ., Canterbury, New Zealand. [Available at http://www.cdc.org.nz/wp-content/uploads/2014/12/aeru-value-of-irrigation-report.pdf.]

Sikorska, A. E., A. Scheidegger, K. Banasik, and J. Rieckermann (2013), Considering rating curve uncertainty in water level predictions, *Hydrol. Earth Syst. Sci.*, *17*, 4415–4427.

Thomas, W. O., Jr., and K. L. Wahl (1993), Summary of the nationwide analysis of the cost effectiveness of the U.S. Geological Survey streamgaging program (1983–88), *U.S. Geol. Surv. Water Resour. Invest. Rep. 93–4168*, 27 pp.

Westerberg, I. K., and H. K. McMillan (2015), Uncertainty in hydrological signatures, *Hydrol. Earth Syst. Sci.*, *19*(9), 3951–3968.

Westerberg, I. K., T. Wagener, G. Coxon, H. K. McMillan, A. Castellarin, A. Montanari, and J. Freer (2016), Uncertainty in hydrological signatures for gauged and ungauged catchments, *Water Resour. Res.*, *52*, 1847–1865, doi:10.1002/2015WR017635.

White, P. A., C. Tschritter, A. Lovett, and M. Cusi (2014), Lake Rotorua catchment boundary relevant to Bay of Plenty Regional Council's water and land management policies, *GNS Sci. Consult. Rep. 2014/111*, 99 pp., Taupo, NZ.