



A retrospective on hydrological catchment modelling based on half a century with the HBV model

Jan Seibert^{1,2} and Sten Bergström^{a,☆}

¹Department of Geography, University of Zurich, 8057 Zurich, Switzerland

²Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden

^aformerly at: Hydrological Research, Swedish Meteorological and Hydrological Institute, 60176 Norrköping, Sweden

☆retired

Correspondence: Jan Seibert (jan.seibert@geo.uzh.ch)

Received: 29 October 2021 – Discussion started: 19 November 2021

Revised: 29 January 2022 – Accepted: 3 February 2022 – Published: 14 March 2022

Abstract. Hydrological catchment models are important tools that are commonly used as the basis for water resource management planning. In the 1960s and 1970s, the development of several relatively simple models to simulate catchment runoff started, and a number of so-called conceptual (or bucket-type) models were suggested. In these models, the complex and heterogeneous hydrological processes in a catchment are represented by a limited number of storage elements and the fluxes between them. While computer limitations were a major motivation for such relatively simple models in the early days, some of these models are still used frequently despite the vast increase in computational opportunities. The HBV (Hydrologiska Byråns Vattenbalansavdelning) model, which was first applied about 50 years ago in Sweden, is a typical example of a conceptual catchment model and has gained large popularity since its inception. During several model intercomparisons, the HBV model performed well despite (or because of) its relatively simple model structure. Here, the history of model development, from thoughtful considerations of different model structures to modelling studies using hundreds of catchments and cloud computing facilities, is described. Furthermore, the wide range of model applications is discussed. The aim is to provide an understanding of the background of model development and a basis for addressing the balance between model complexity and data availability that will also face hydrologists in the coming decades.

1 Introduction

The fundamental questions to hydrologists have remained similar over the past decades. How much water is available today and will be in the future? How much discharge is there on average, at minimum and maximum flow? How extreme can it get? How does discharge vary between catchments? How do vegetation and climate influence discharge? What is the role of subsurface flow? How will river flow change over the days and months to come?

Despite these long-standing questions, many aspects of hydrological catchment modelling have changed over the past 50 years, as can be seen from the modified advice to hypothetical young PhD students:

- Around 1970, students were told that “We need to simulate runoff – go and construct a hydrological catchment model.”
- Around 1995, the advice given was “Use a physically based model, not something as simplistic as the HBV model, which requires calibration to work well. Simple models will not be used for much longer. As soon as computers become more powerful, everyone will use more advanced models.”. Admittedly, at that time, some researchers were already arguing that simple models such as the HBV model were ideal for simulating catchment runoff, whereas more complex models should mainly be used for applications in which other variables were also of interest (Beven, 1989; Refsgaard and Knudsen, 1996).

- Around 2000, students were told that “Model predictions are necessarily uncertain. That’s the way it is, so you better study how we can estimate that uncertainty.”
- Around 2020, this advice changed to “HBV for your PhD work – good choice, but why are you only using 100 catchments? How will you ever get any generally valid results with so few?”

In other words, doing a PhD in hydrological catchment modelling has remained challenging over the years, but the type of challenge has shifted from handling or sorting punched cards to making sense of gigabytes of data. Imagine Sten, Jan and a student, who is working on her (or his) PhD thesis, discussing hydrological modelling (Fig. 1). Such a dialogue between PhD students from these different times talking about the HBV model (Fig. 2) might proceed as follows.

Student: “You used punched cards? How did that work? Was there an app for that?”

Sten: “In the very beginning, we even used paper tapes with the model code and data punched on it. If we were lucky, we could count on one or two model runs per day. By the 1990s, computers had become really powerful; since then, we have run HBV on our desktop computers.”

Jan: “Yes, I could even do Monte Carlo runs with a few million model runs, but I had to borrow all the department’s computers over the Christmas break.” (Seibert, 1999)

Student: “Only a few million model runs? This is what cloud computing does for me during the coffee break. By the way, you were calling HBV conceptual, bucket-type, physically based, lumped and (semi-)distributed; now, what is it?”

Jan: “We have not always been consistent in model classification over the years. The term conceptual model is widely used in hydrology for these models, representing the basic concepts of water storages and fluxes. However, I now prefer using the terms bucket-type and semi-distributed. The first refers to the limited number of storage elements, and the latter refers to the use of elevation (and sometimes vegetation) zones as catchment area fractions.”

Sten: “HBV is all of this. A simple model can be as physically based as a more complex one. After all, in HBV we obey the water balance, which is a basic physical law. There have been endless discussions on which types of models are more physically based (for an interesting exchange, see Beven, 1990, and the subsequent discussion, later reflected on by, for instance, Beven, 2001, and Refsgaard et al., 2016). But don’t forget that model complexity must match the input data available, including all areal variabilities in a basin. For example, if we do not describe the albedo or atmospheric inversions reasonably well, we cannot expect a physically based snowmelt model to work either because the energy balance at the surface of the snow will be compromised. Honestly, I care more that a model provides useful simulations than that it intends to describe all processes in full detail. Performance counts! So, let’s stick with conceptual and semi-distributed for now.”

Student: “I read some of your papers and understand that you put some thought into model development, but what on Earth were you thinking when deciding on the name HBV model?”

Sten: “Well, once we had the model, we needed a name for its first scientific publication. The name of the unit at SMHI (Swedish Meteorological and Hydrological Institute) where the model was developed (Hydrologiska Byråns Vattenbalansavdelning) seemed like a good choice at the time. Had we only known how long we would stick to this name ...”

Student: “And HBV light?”

Jan: “My version of HBV was designed for easier use in teaching, so we started calling it HBV light and also used this name in an early publication. Each time I get an email such as “I liked your light model, can you send me the real one”, I question the name.”

Student: “And the performance measure R^2 , which is so often confused with r^2 ?”

Sten and Jan: “That was not us; blame Nash and Sutcliffe and their classic publication on hydrological modelling (Nash and Sutcliffe, 1970). We will come back to this later.”

Student: “It seems that a lot of researchers are using HBV these days, weren’t there other conceptual models developed in the 1970s–1990s?”

Sten: “Yes, there were several. We will come back to them.

Student: “And today HBV is one of the most popular of these models; it seems I chose the best model for my PhD studies. What made HBV so successful?”

Jan: “In short, the three Ps”

Student: “What do you mean?”

Jan: “Parsimony – one thing I especially like about the HBV model is that the different model routines are all simple and use a limited number of parameters that need adjustment when tuning the model. Thus, the frustration when trying to calibrate an overparameterized model can be avoided.”

Sten: “Performance – the HBV model was able to do what was needed and performed quite well in several model inter-comparisons” (Breuer et al., 2009; WMO, 1986)

Sten: “... and not to forget, persistence – it certainly helped that HBV was widely applied at SMHI and also became the main tool for operational hydrological forecasting in some other countries. In Sweden it also became the standard tool for the estimation of design floods (Bergström, 1992). Soon after its introduction, many researchers started and continued using the model for all kind of applications.”

Student: “Interesting. Would you say there is anything one cannot do with HBV?”

Jan: “Models like HBV focus on the simulation of catchment runoff. If you are interested in the spatial variation of internal water fluxes and states, you might have to use a fully distributed model after all. Furthermore, quantifying impacts of climate change and land-use/land-cover changes on hydrology as well as simulations for ungauged catchments are difficult ...”

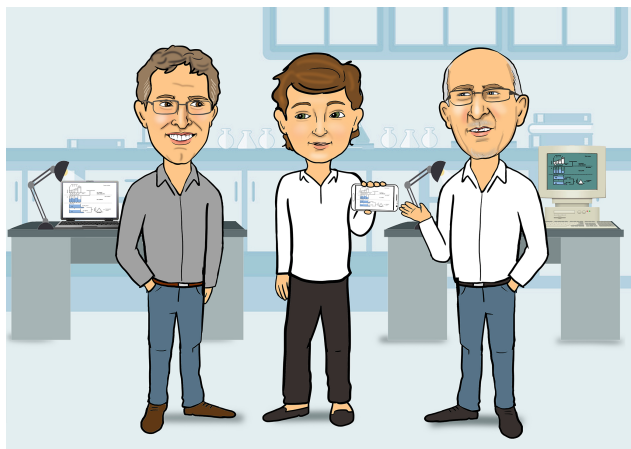


Figure 1. Discussion of hydrological catchment modelling between Jan, a PhD student and Sten. The smartphone that the PhD student is holding in her (or his) hand is far more powerful than the computers available when Sten started developing the HBV model in the early 1970s (drawing by Steph's Sketches).

Sten: "... but we do it anyway. More seriously, these types of applications were, for a long time, considered impossible with a conceptual type of model like HBV. Still, later in this text, we will describe how HBV can contribute here nevertheless."

Student: "It seems that HBV has become a widely used, powerful tool. How did this all happen? What were the ideas driving its development, and who were the people behind this development? What did the journey of HBV look like, and how will it continue?"

Sten and Jan: "Good questions! We will try to address these in this paper, in which we describe the history of model development, from thoughtful considerations of the formulation of different model structures to modelling studies using hundreds of catchments and cloud computing facilities. We show that the HBV model performed surprisingly well during several model intercomparisons, and we discuss the wide range of model applications."

2 History of model development

2.1 Early thoughts on catchment modelling

The International Hydrological Decade (IHD), a research programme initiated by the United Nations Educational, Scientific and Cultural Organization (UNESCO) that spanned from 1965 to 1974, became instrumental in advancing hydrology as a science, with new financial resources and intensified national and international cooperation. Computers had arrived; however, in the beginning, they were only used by a specialized few. The first hydrological computer-based models, such as the Streamflow Synthesis and Reservoir Regulation (SSARR) model (Brooks et al., 1975) and the Stanford

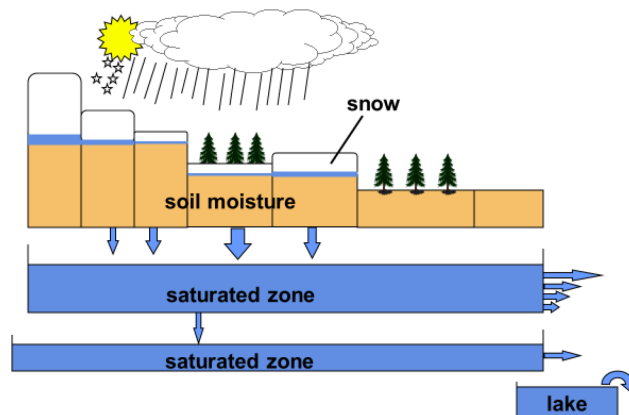


Figure 2. Structure of the HBV model (drawn by Sten Bergström).

Watershed Model (Crawford and Linsley, 1966) had already been developed in the 1960s.

Internationally, there was a lively debate on the principles of hydrological modelling. The initial enthusiasm about the calculation power of the new computers was met by arguments against the development of model systems that were too complex (Dawdy and O'Donnell, 1965). With too many details in a model, it was argued that the number of model parameters introduced so many degrees of freedom that it was difficult to handle and control the model.

In 1970, Eamonn Nash and co-authors published a set of papers that strongly advocated for best practices in model development (Mandeville et al., 1970; Nash and Sutcliffe, 1970; O'Connell et al., 1970). In the first paper, a strategy for model development was suggested under the heading "Progressive modification". At the end of their eminent 1970 paper (p. 290), Nash and Sutcliffe wrote the following:

If one accepts that it is desirable to have a simple rather than a complex model, and this is certainly true if it is hoped to obtain stable values of the optimised parameters, then it would seem that a systematic procedure would be as follows:

1. Assume a simple model, but one which can be elaborated further.
2. Optimise the parameters and study their stability.
3. Measure the efficiency R^2 .
4. Modify the model– if possible by the introduction of a new part – repeat (2) and (3), measure R^2 and decide on acceptance or rejection of the modification.
5. Choose the next modification. A comparative plotting of computed and observed discharge hydrographs may indicate what modification is desirable.

6. Because all models cannot be arranged in increasing order of complexity it may be necessary to compare two or more models of similar complexity. This may be done by comparing R^2 .

Thus, from the beginning, there were two fundamentally different approaches in hydrological modelling. On the one hand, start simple and add what is needed; on the other hand, attempt to represent all relevant hydrological processes from the beginning, i.e. put in all we know at once. The development of the HBV model followed the first approach. It is important to note that while some modellers claimed (and partly still do) that more complexity would provide a better representation of the “true” processes in a catchment, there have also been contributions from the early days in which it was argued that more complex models are only needed if variables other than catchment runoff are of interest.

The paper by Nash and Sutcliffe (1970) is one of the most cited in hydrological modelling, although this is mainly because of the introduction of the so-called model efficiency or explained variance. The authors called this measure R^2 , but other names, such as Nash–Sutcliffe efficiency (NSE), are preferable to avoid confusion with the coefficient of determination. The NSE has drawbacks and has been shown to overemphasize errors in the timing and magnitude of flood peaks, whereas less emphasis is placed on volume errors and low flows. Several alternative performance measures have been proposed over the years, but the NSE is still probably the most commonly used measure. Only recently has the Kling–Gupta efficiency (KGE) measure (Gupta et al., 2009) gained popularity. The advantage of this measure, as well as its modification the non-parametric efficiency (NPE; Pool et al., 2018), is that it allows different components of the total model error (correlation, bias and variability) to be distinguished.

Model calibration soon became an issue. As model calibration (where model parameters are successively adjusted until the best model performance is obtained) became a standard procedure, there was the risk that hydrological modelling would degenerate into simple curve fitting. The way to overcome this was to judge the model by its performance over a period that had not been used to calibrate its parameters. This split-sample testing became standard for most applications, including the international model intercomparisons carried out by the World Meteorological Organization in 1975 and 1986 (see below).

In the early days of model development, the most trusted way of judging the performance of a hydrological model was to simply look at computed and observed hydrographs and to judge model performance by this so-called “visual inspection”. Objective criteria of goodness of fit were used as support, but, for practical reasons, they soon became the standard when the number of applications grew. The R^2 value, as suggested by Nash and Sutcliffe in 1970, was the most

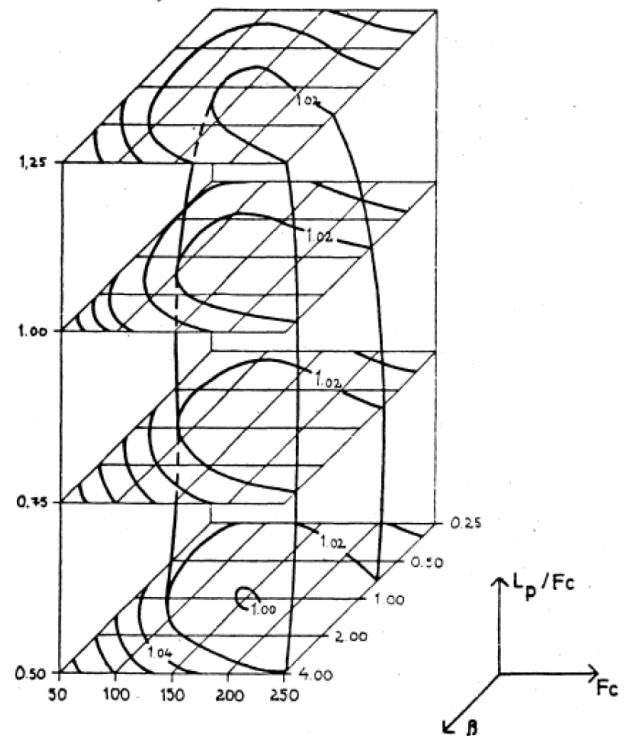


Figure 3. Mapping of the error function around the optimum values of the three parameters in the soil moisture routine of the HBV model when applied to the Filefjell research basin in Norway. Hand-made drawing based on 100 model runs from Bergström (1976).

popular criteria, supplemented with a measure of the volume error. The calibration process meant that model parameters were adjusted until a maximum value of R^2 was found. This tuning was not a trivial problem, as the limited capacity of the early computers did not allow for a great number of simulations. Therefore, various techniques to search for the optimum of the error function were used, such as the algorithm by Rosenbrock (1960). However, the number of model runs was still a limiting factor.

To ensure that the introduced model parameters really improved the model, the error function was sometimes mapped, and its topography around optimum parameter values was analysed. A topography without a distinct peak at the optimum parameter values reveals parameters with low significance or severe interaction between them (Fig. 3).

In the early days of work on the HBV model, we spent hours and hours simply looking at the computed and observed hydrographs and trying to understand what was really going on when trying to tune the model. Sometimes, severe problems in the observations were also detected this way. During manual calibration, 10–30 model runs were typically performed (note that one model run took about half a day in the 1970s). Today, automatic calibration of hydrological models is commonly used. Visual inspection, however,

should not be ruled out in the development phase. It gives the hydrologist an insight into the dynamics and sensitivity of the model to disturbances in its parameters, indicates important interactions, and helps with identifying feedback mechanisms between model components. It is time well spent.

2.2 Catchment model developments

Early models from the 1970s included the Canadian UBC model (Quick and Pipes, 1976); the Danish NAM model (Nielsen and Hansen, 1973); the Japanese TANK model (Sugawara, 1979); the Swiss–American SRM model (Rango and Martinec, 1979); the US National Weather Service River Forecast System (NWSRFS, Anderson, 1973), which is based on the Sacramento Catchment Model (Burnash, 1995); the GR4J model (Michel, 1983; Perrin et al., 2003); and the Swedish HBV model (Bergström, 1976). Later models included the British TOPMODEL (Beven and Kirkby, 1979), the Chinese Xinanjiang model (Zhao, 1992), the Danish MIKE-SHE model (Refsgaard et al., 2010), the Italian ARNO model (Todini, 1996) and the American VIC model (Liang et al., 1994). An overview of some of the best-known hydrological models at the end of the 20th century can be found in the work by Singh (1995). Peel and McMahon (2020) provide a recent review on the development of runoff models.

While there has been an exchange of ideas between some model developers, there have also been independent developments in parallel. For example, the philosophy behind the development of the French GR4J model by Claude Michel (Michel, 1983; Perrin et al., 2001) has large similarities to the HBV philosophy; however, despite developing the models around the same time, there was no exchange between the two model developers.

Early in the hydrological modelling debate, the term “conceptual model” was contrasted with “physically based models”, and “lumped models” were contrasted with “distributed models”. These are rather illogical classifications that have persisted over the years. It was understood that a conceptual model in this sense lies somewhere between a black-box approach and a fully physical representation of the hydrological system. Still, as almost all models are based on the water balance in some way, all models could be said to be physically based. Using such a broad interpretation of the term physically based implies that the classification becomes meaningless. There is an obvious difference between bucket-type models, such as the HBV model, and more sophisticated spatially distributed models such as the SHE model or the modelling system presented by Kollet et al. (2018), but we argue that the latter type of model is not automatically more physically based in the sense of being a (more) correct representation of the real catchment processes. Beven (1989) pointed out that using effective parameter values at the grid scale in physically based models implies that these models are necessarily conceptual approximations; even if the Richards equa-

tion applies, it does not average linearly. Model classification can also be seen as a more academic question, whereas operational model users are more interested in the performance and feasibility of the model.

In the beginning, the simpler models regarded the catchment as one homogeneous unit, so-called lumped models, but conceptual models are currently usually applied with some form of spatial variation. In mountainous catchments, for instance, the use of elevation zones is common, and these models are also used with vegetation zones or sub-catchments. In other words, spatial variations are considered when describing fluxes and storages. While these cannot be mapped to exact locations in the catchment, the model represents spatial variability by distribution functions; thus, the term semi-distributed is used.

There are also gridded implementations of the HBV model. It started as a way of producing water balance maps and made it possible to better communicate and cooperate with climate modellers, who were used to gridded formats. The first gridded HBV model was used to produce hydrological maps for the National Atlas of Sweden (Raab and Vedin, 1995). Since then, simulations with gridded model versions have become more common. However, the grid cells are typically rather large (1 km² or more) to ensure that lateral groundwater flow between the cells is negligible (i.e. small in comparison with streamflow) and that the groundwater storages in adjacent cells can be treated as independent of one another. It should also be noted that such a distributed version of the HBV model does not necessarily lead to better runoff simulations (Das et al., 2008). On the other hand, distributed versions of the HBV model have been used to calibrate parameter values simultaneously against data from many different catchments. This approach has been used, for instance, to generate runoff maps for Norway (Beldring et al., 2003) and Georgia (Beldring et al., 2017).

Using field observations of internal model variables such as snow pack, soil moisture deficit or groundwater dynamics to evaluate the performance of a catchment model is challenging. There is the commensurability issue (Beven, 2018) when comparing observed point values to internal areal model variables. Direct comparisons are also tricky when using aggregated observation data or spatial observations (e.g. remote sensing of snow cover). Nevertheless, several attempts to compare the internal variables of the HBV model against field observations have been carried out (Andersson, 1988; Bergström and Lindström, 2015; Bergström and Sandberg, 1983; Brandt and Bergström, 1994; Seibert, 2000). These studies suggest, at least qualitatively, that the dynamics of the internal variables of the model agree reasonably well with observations and, thus, help enhance the confidence in the model.

2.3 The story of HBV in Sweden

In Sweden, intense large-scale hydropower development commenced at the beginning of the 20th century and was due to be completed in the 1970s. By this time, the hydropower system had become the backbone of the electricity supply system, but further development met strong environmental concerns and opposition. Thus, it was time to focus on the operation of the system rather than further exploitation of the rivers. Hence, the industry began to look for a reliable and practical operational forecasting tool for inflow to reservoirs and power plants.

The Swedish Meteorological and Hydrological Institute (SMHI) had a long tradition of working with hydropower development. It also had the advantage of hosting meteorology and hydrology under the same roof, where the meteorological service could provide computer resources and meteorological data for the hydrologists. The small hydrological research unit at SMHI, HBV (Hydrologiska Byråns Vattenbalansavdelning), was also a very active national part in the scientific cooperation within the International Hydrological Decade (IHD). Thus, the scene was set for computerized hydrological modelling at SMHI.

Eamonn Nash spent some time at SMHI in Stockholm in 1972, and his ideas came to have a strong impact on the hydrological modelling work at the institute. With operational use in mind, it was important that the complexity of the model could be justified by its performance and that data demands could be met, even in remote areas with limited data coverage. In particular, the number of model parameters, which are subject to calibration, should be kept as low as possible to avoid parameter interactions and subroutines representing insignificant process details. Today, we call this “overparameterization”. It was a surprise how soon the point was reached where increased complexity did not improve model performance.

The first successful run by an early version of a hydrological model at SMHI was carried out in the 12.7 km² Lilla Tivsjön research basin in the spring of 1972 (Fig. 4). The results were surprisingly good with this model embryo with only one runoff component. Later applications showed that three components are generally needed for good agreement between modelled and observed runoff. Still, the model performed surprisingly well in spite of a very simple structure and only few parameters to calibrate. The model was then named the HBV model (from the abbreviation of the research unit where it was developed), and the first report (in Swedish) was published the same year (Bergström, 1972). Later that year, modelling results were presented to Nordic colleagues at the Nordic Hydrological Conference in Sandefjord, Norway. In 1973, the first publication in an international scientific journal appeared (Bergström and Forsman, 1973). It is interesting to note that simultaneously, and in the very same edition of *Nordic Hydrology*, the Danish NAM model was published by Nielsen and Hansen (1973). It is similar to the

HBV model in its simplistic set-up, but subroutines are formulated differently.

In the pioneering days, when computerized hydrological modelling had just begun, working conditions differed a lot from today. Even if you had access to a computer, you could not count on more than a few test runs per day. At SMHI, for example, you programmed yourself, and you also had to punch the model code and the input and validation data yourself on paper tapes or cards. Geographical information systems (GISs) were unheard of, so extracting the hypsometric curve and similar mapping work had to be done manually from a paper copy of a topographic map. When you were ready for a test run, you simply walked into the noisy computer centre and kindly asked the staff to run your work in between the meteorological forecasts. Visual comparison of the graphs of the computed and the observed hydrograph was the most important way of judging the model performance. The R^2 value was used as a support criterion of performance only.

The first tests with the HBV model were made for the snow-free period only. The first subroutine for snow accumulation and melt was added to the model in 1975 (Bergström, 1975). With the introduction of an area–elevation zoning of the snow routine for mountainous terrain in 1976, the model was ready for operational tests (Bergström, 1976; Bergström and Jönsson, 1976). In 1976, the first operational forecasts were carried out for the hydropower industry. Both short-range (a few days) and long-range forecasts (several months) were carried out after the model had been updated to good starting conditions. The short-range forecasts were based on meteorological forecasts, whereas the long-range forecasts used a set of historical climate records for the forecasted period to simulate a range of possible future river flows (Bergström, 1978).

Parallel to the work for the power industry, interest in flood protection grew, in particular after severe flooding in the spring of 1977. Modelling started for some flood-prone rivers in central and southern Sweden in the late 1970s.

The HBV model was proven to perform well based on how it managed to simulate river runoff in an increasing number of basins; however, its merits were initially more appreciated by engineers and power producers than by academics. Maybe it was considered more of an engineering tool than a scientific achievement. The necessary calibration process was sometimes looked upon as mere curve fitting of a black-box model. It was evident that the hydropower industry and the national hydrological service, SMHI, had a more practical view of hydrological modelling than scientists from universities. It took some time before the physical considerations and interpretations in the HBV model were appreciated by academics in Sweden. Nevertheless, work on the HBV model resulted in a PhD thesis that was defended in 1976 (Bergström, 1976) and is still the standard reference for the HBV model. Since then, various aspects of the HBV model have been addressed in numerous PhD theses around the world. The con-

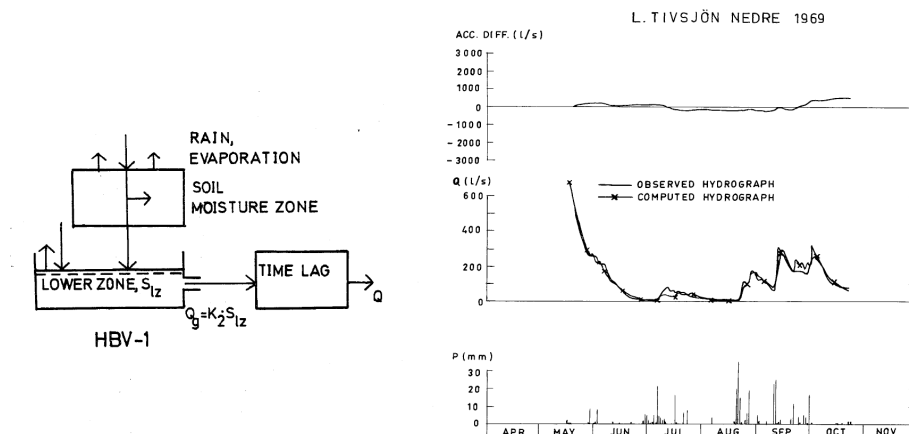


Figure 4. A sketch of the very first version of the HBV model and the very first successful simulation of runoff based on data from the Lilla Tivsjön basin (from Bergström, 1976).

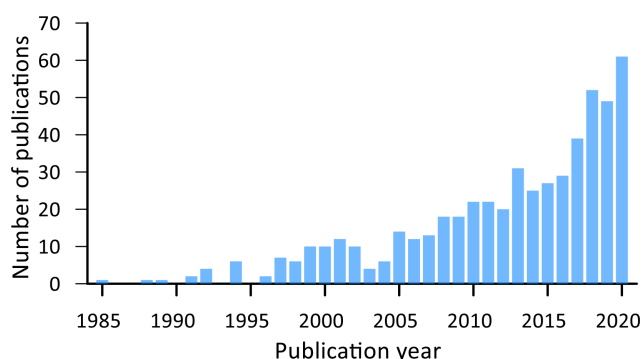


Figure 5. Number of publications found in the “Web of Science” using the search query “HBV AND (runoff OR hydrology)” (performed on 2 October 2021).

tinued popularity of the HBV model can also be seen from the steadily increasing number of publications (Fig. 5).

The hydropower industry quickly realized the potential of a model like HBV and strongly supported its further development. Thus, hydropower operation became the main driving force in the early development of the model. The day-to-day dialogue with hydropower companies, where the model performance was investigated in great detail, was instrumental. The industrial interest also helped to secure financial support for the model development and data collection.

It was soon realized that a production system with a practical and user-friendly interface is absolutely necessary for a hydrological model to become accepted as an operational tool. This system must include routines for model calibration, updating of the model as new data arrives, and an interface for runoff simulation and forecasting. The development of the production system would prove to be a lot more costly than developing the model itself. The operational requirements demanded the development of a user-friendly mod-

elling system that, in turn, contributed to the widespread use of the HBV model.

At SMHI, the Integrated Hydrological Modelling System (IHMS) was developed and ready for use in the mid-1990s (e.g. Lindell et al., 1996). The model could now be run by almost anyone. At the same time, the commercial importance of HBV modelling grew. The source code of the HBV model was made available to the academic community for research and education upon request, although with certain commercial restrictions. However, the advanced functionalities in the IHMS system, required from an operational point of view, were not needed by the academic research community, and new model codes were written by several researchers.

The operational modelling system (IHMS) also included a routine for automatic model calibration. This was a necessity when the number of applications grew. The criterion of model performance, R^2 , suggested by Nash and Sutcliffe in 1970, was supplemented by a penalty for modelled volume errors in the automatic calibration routine in the IHMS system (Lindström, 1997).

In the 1980s, several flood situations alerted the Swedish hydropower industry to the fact that the capacity of the spillways did not meet reasonable safety requirements and that the risk of dam failures was unacceptably high. After 5 years of intense studies, new national guidelines for estimating design floods for the spillways were formulated and adopted in 1990. The emphasis was on peak flows, but flood volumes were also of interest in the complex system of dams and reservoirs in the developed rivers. The HBV model played a key role in this process. The model was now developed for entire river systems, which meant that river regulation was also considered. By combining observed flood-generating factors such as heavy rainfall, extreme snowmelt situations and wet antecedent conditions, the model was used to simulate floods with magnitudes that far exceeded anything that had ever been recorded. These synthetic design floods were

often more than twice as big as the highest observed floods, which implied an extreme extrapolation far outside of the normal range during model calibration. This would never have been accepted in the model's infancy, but confidence in the performance had grown as the model has been successfully applied in an increasing number of basins and under a wide range of climatological conditions.

In 1990, new Swedish guidelines for spillway design floods were finally adopted, and they have been applied ever since with only small modifications (Bergström, 1992; Flödeskommittén, 1990; Lindström and Harlin, 1992; Norstedt et al., 1992). Today, these guidelines are the standard procedure not only for the hydropower industry but for all areal planning and developments in areas exposed to floods in the country. The safety of all major dams has been reassessed, and many of them have been modified to meet the new standards. In 2015 and 2021, the Swedish guidelines for design floods for dams were supplemented with a strategy for adaptation to climate change (Svenska_Kraftnät, Svensk_Energi and SveMin, 2022).

The focus on extreme floods influenced the re-evaluation of the structure of the HBV model, which was carried out by SMHI in the 1990s. The new model version, HBV-96, got a modified response function which performed better than the original function. This also meant that the number of free model parameters used for calibration of the model could be reduced by one, thereby reducing the risk of overparameterization, i.e. the inability to determine one single best parameter set (Lindström et al., 1997). It is interesting to note that the title of this paper is *Development and test of the distributed HBV-96 model*. To the authors' knowledge, there have been no objections from the scientific community to using the word "distributed" in this context.

3 The three Ps: parsimony, performance and persistence

The international success of the HBV model is, in our opinion, much related to the three Ps (parsimony, performance and persistence), as mentioned in Sect. 1. The parsimony made the model code easy to understand (and reprogram), the HBV model often performed well (especially in comparison studies), and it certainly helped that the model was intensely used by SMHI and the hydrological services of Norway and Finland and became a standard tool for the Nordic hydropower industry (in contrast to an individual researcher or research group using it, as was the case for other models). In this section, we further elaborate on these points.

3.1 Parsimony in model development

A key characteristic of the HBV model is its parsimony. In initial tests, parameters that were found to have little importance were eliminated. Thus, compared with most other

models, the HBV model has a rather low number of model parameters. The HBV model started with a simple lumped approach. The drainage basin was assumed to be a homogeneous unit and has gradually developed into a geographically distributed model where sub-basins are the prime hydrological unit and lakes can be modelled explicitly

In general, three distinct physical components can be identified in the structure of the HBV model: snow accumulation and melt, soil moisture accounting, and runoff response, including groundwater dynamics. The capability to simulate the latter was evaluated with satisfactory results in the early days of the HBV model (Bergström and Sandberg, 1983). As inflow and outflow of lakes has a strong influence on river flow dynamics, major lakes can be modelled explicitly. For parameters and equations used in the HBV model, we refer to previous descriptions such as the one in Seibert and Vis (2012).

3.1.1 Snow accumulation and melt

The simple traditional degree-day approach is the most commonly used approach for snowmelt modelling in the HBV model. The water retention capacity of snow is used to capture the delay between melt and the time when the water leaves the snowpack. Topography is usually accounted for by area–elevation zones and temperature and precipitation lapse rates. Land characteristics such as forest and open land can be separated, and a statistically distributed approach can be adopted to account for the redistribution of snow in open terrain. Glaciers are sometimes modelled separately.

Of great importance for the development of the snow routine of the HBV model was the World Meteorological Organization (WMO) international intercomparison of models of snowmelt runoff (WMO, 1986) in which the results from the HBV model were compared to other international models. One conclusion from that project is specifically worth citing: "On the basis of available information, it was not possible to rank tested models or classes of models in order of performance. The complexity of the structure of the models could not be related to the quality of the simulation results." (WMO, 1986, p. 35).

Since the publication of the results of the WMO intercomparison project, numerous attempts have been made to improve the simple temperature-index technique by introducing more physically based approaches, where the energy balance of the snowpack is accounted for at least to some extent. The results have been mixed or disappointing. A recent comprehensive study of snow model complexity versus performance by Girons Lopez et al. (2020), carried out for 54 basins in Switzerland and the Czech Republic, largely confirmed these results. Several model modifications were tested, which were argued to increase the realism of the model. However, only a few modifications resulted in improved simulations, and these improvements were generally small.

3.1.2 Soil moisture accounting

The calculation of the soil moisture deficit and its impact on runoff generation was the key to the immediate success of the HBV model. At first, it seemed obvious that any hydrological model should have a component for surface runoff, even if the soils were far from saturated. However, this was not supported by observations in the research basins in Scandinavia. Instead, a simple curve was introduced to relate runoff generation to soil moisture deficit using the modelled moisture state and one free parameter, the so-called BETA parameter. This meant that the soil acted as a buffer for runoff generation. When the catchment was dry most of the rainfall or snowmelt was used to fill up the moisture deficit in the soil; when it was wet, most additional water recharged the response function and eventually contributed to river runoff. This approach, first developed in 1972, has become the signature of all variants of the HBV model ever since. It has proven very versatile in applications to catchments over a great span of climatological and physiographical conditions worldwide.

This does not exclude the fact that surface runoff on unsaturated soils may occur locally due to very intense rainfall. However, local surface runoff of this kind is normally difficult to detect in the integrated river flow at the outlet of a drainage basin of a large catchment.

The key modelled variable in the soil moisture accounting of the HBV model is the computed soil moisture storage (SM), which is allowed to vary between zero and the parameter value FC (maximum soil water storage, not to be confused with what might be measured in soil samples as “field capacity”). The SM/FC relation varies from zero to one and is an index of the wetness of the soil in the basin, which determines how much of the water becomes recharge, and eventually runoff, and how much remains in the soil. The soil moisture storage is depleted by evapotranspiration, which has its potential value when the soil is wet and gradually declines as the soil dries out. For potential evapotranspiration, long-term monthly average values (“the twelve values”) are usually used, which are based on direct measurements or, for instance, the Penman–Monteith equation. Modifying these values based on temperature anomalies has led to improved runoff simulations (Lindström and Bergström, 1992).

The soil moisture accounting of the HBV model can be seen as an attempt to consider both the total soil moisture in the basin and its areal sub-basin or sub-grid variability. The soil moisture accounting approach probably also contributes to the generally good performance of the HBV model, as this implies that the model is robust with regard to the temporal distribution of precipitation and (small) timing errors. Similarly, model simulations might be relatively robust to random errors in precipitation amounts.

While the simple structure of the model allowed recoding, it can also be noted that some technical details were implemented differently. Most importantly, the soil routine of the HBV model is highly non-linear. Thus, in the original code,

the differential equation was solved by adding one millimetre at a time, as described by Bergström and Forsman (1973) and Bergström (1976). This detail has not been implemented in the same way in all model versions.

3.1.3 Response function

The role of the response function of a hydrological model is basically to give the simulated hydrograph a shape in time that agrees with observations. Traditionally, this has often been obtained using a set of reservoirs or buckets, where different outlets are used to represent quicker or slower runoff components.

When developing the HBV model, it was found that good agreement between the shape of the computed and observed hydrographs could generally be reached with two buckets and three runoff components represented by three holes and percolation from the upper to the lower bucket. This was supplemented with a transformation function to account for further damping of river runoff in the basin. This approach is still the most commonly used response function of the HBV model internationally. However, as mentioned above, when the HBV model was re-evaluated by SMHI in 1996, a new routine was introduced in which the two drainage holes of the upper reservoir were replaced by a continuous function. This approach proved to perform better, particularly for peak flow simulations. Another advantage was that the number of parameters in the response function could be reduced from five to four (Lindström et al., 1997). This new routine is standard at SMHI, whereas the original five-parameter version is still frequently used in other model versions.

The response function was subsequently modified again by SMHI. The aim of this modification was to better cope with the variable dynamics of winter and summer peaks. The idea was to account for the variation in groundwater levels and, thus, the extension of recharge and discharge areas between wet and dry conditions in the basin. It was an attempt to let the hydrologically active part of the basin gradually become smaller when the basin was drying out. Numerically, this effect was obtained by linking the drainage curve of the upper reservoir to the calculated soil moisture state as described by Bergström and Lindström (2015). Thus, it was possible to model a more peaked flood response for summers than for winters, without introducing further model parameters. This is a modification that is not that well known but has been shown to have potential to improve the simulation. It is an option in the HBV model package from SMHI.

3.2 Comparative tests of model performances

In 1969, the WMO initiated the “Intercomparison of conceptual models used in operational hydrological forecasting” project, with 10 participating rainfall–runoff models (WMO, 1975). It was followed by the “Intercomparison of models of snowmelt runoff” (WMO, 1986) and “Real-time intercom-

parison of hydrological models" (WMO, 1988) projects. This set of projects meant a lot for the practical use of hydrological models and set the standards for model intercomparisons. As an example, the validation of the performance of the individual models was based on an independent period of data not used for calibration of the models and was carried out by an independent team. The conclusions in the final report were agreed upon during a final conference between the participating modellers.

Thanks to the WMO projects intercomparing hydrological models, modellers from around the world had the opportunity to discuss practice and experience in hydrological modelling and to compare results. The projects confirmed that more complex models do not necessarily outperform the more simple ones with respect to catchment runoff simulations. In snow modelling, the simple degree-day method, which is basically based on a near-linear relationship between air temperatures and melt, was hard to outperform using more sophisticated models, which accounted for the complete energy balance of the snowpack. It was simply impossible for a model to describe the basin-wide areal distribution of the terms in the energy balance with sufficient accuracy.

About 20 years later, another model intercomparison workshop took place in Giessen, Germany (Breuer et al., 2009). Several groups were invited to apply their respective models to the Dill catchment (a 693 km² catchment north of Frankfurt) to simulate land-use change effects. In the end, 10 models were applied, ranging from fully distributed models (such as SHE with hundreds of parameters) to simple models (such as HBV). The initial step was a model intercomparison, and it turned out that the HBV model performed best (in terms of NSE) for both the calibration and validation periods (Breuer et al., 2009). This was an unexpected result, as one should, in theory, be able to achieve better model fits with more complex models and more parameters. It is important to note that only a few of the parameters were optimized during calibration for the more complex models. Furthermore, in such model intercomparisons, not only the model but also the experience of the respective modellers might influence the results. Discussions at the workshop also revealed that the fact that data were received only a few weeks before the workshop might have contributed to this result. With the HBV model, there was no problem setting up the model in time, whereas groups with more complex models had run into several problems. Still, the results demonstrated that, in practice, simple models could outperform more complex models.

One argument in favour of more complex models has always been the need for calibration for conceptual models. The argument is that for more complex, fully distributed models, it is possible, at least in theory, to determine model parameter values based on field observations without any model calibration. The view of the possibility to use models without any calibration has changed over the years, and most modellers would currently probably accept the need for at least some calibration, or tuning, even for more complex

models (Hrachowitz and Clark, 2017; Jens Christian Refsgaard, personal communication, 2021). A full discussion on whether model parameters in more complex models are observable or not is beyond the scope of this paper. Still, it is important to note that conceptual models can also be employed in an uncalibrated mode by using ensembles of randomly chosen parameter values. Recently the performance of the HBV model and the SHE model has been compared for more than 300 catchments in the UK (Seibert et al., 2018b). The parameter values for SHE had been derived based on detailed information of catchment characteristics, whereas an ensemble of random parameter sets was used for HBV. The performance of the two models with respect to simulated discharge was, on average, similar.

3.3 International development and use of the HBV model

There has been a great amount of persistency in the development and application of the HBV model over the years, both in Sweden and internationally. The first international adoption of the HBV model was carried out by colleagues from Norway, with the first HBV model results from Norway appearing in 1976. Comparative tests of hydrological models were also carried out in Norway (Saelthun, 1978). The exchange of ideas between Sweden and other Nordic countries intensified. In Norway, with its rugged mountainous terrain, Killingtveit and Aam (1978) contributed with a distributed approach to snow accumulation, which also influenced the development of the Swedish HBV model. The HBV model became a standard tool for the hydrologists of both the Norwegian and Finnish hydrological services (Killingtveit and Saelthun, 1995; Saelthun, 1996; Vehvilainen, 1986). During the 1980s, the international expansion of the HBV concept accelerated (Bergström, 1992). The simple structure and ease of understanding and handling of the HBV model were appreciated, and good results were reported from an increasing number of countries and modelling groups. The global spread was enhanced by developing an increasing number of new variants of the HBV model advanced by the international scientific community.

New versions of the HBV model appeared in many countries, some with modified descriptions of the physical processes, but the basic structure remained the same: Braun and Renner (1992) developed a Swiss variant named HBV-ETH, which was frequently used in alpine areas (Hagg et al., 2007); a German version, named HBV-D, was presented by Krysanova et al. (1999); Stahl et al. (2008) presented the Canadian HBV-CE version (Jost et al., 2012), which is currently commonly used (in particular) in mountainous basins with glaciers (Northern Climate Exchange, 2014); and another version of HBV has been developed in Austria at the Vienna University of Technology (Merz and Blöschl, 2004).

Another model version is HBV light, which was launched in the late 1990s by one of the authors of the present pa-

per (Seibert and Vis, 2012). The motivation for developing this version was to provide a user-friendly model version, especially for use in teaching. The user-friendliness was achieved by simplifying the model to its core functionality (e.g. leaving out functionalities related to hydropower operation) and providing a graphical user interface (GUI) (remember that GUIs were rather new at that time). In the meantime, HBV light has been further developed and includes features like automatic calibration, optional model variants, sub-catchments and a dynamic glacier routine (Seibert et al., 2018a). Furthermore, it is possible to use different time steps and to trace water particles through the model (Weiler et al., 2018).

The HBV concept has also been included in distributed model frameworks. A good example of this are model developments in Norway where a gridded version of the HBV model was combined with other process descriptions such as evapotranspiration modelling (Huang et al., 2019), stream-flow routing (Li et al., 2014), and glacier ice melting and retreat (Li et al., 2015).

Today the number of countries where the HBV model has been applied, to the authors' knowledge, is on the order of 100. The applications often started with hydrological forecasting in mind, but, nowadays, studies of climate change impacts on water resources and simulations in ungauged catchments are increasingly common.

4 Challenging applications

The development of the HBV model started in rather small research basins and then moved to larger basins, which were of more interest to the power industry. With the growing interest in climate change and its possible impact on water resources came attempts to model continental scales, like the drainage basin of the Baltic Sea. At first, this seemed to be a tremendous task, but the flexible structure of the HBV model allowed such applications. It seems that many of the small-scale processes average out at larger scales; thus, the simulated runoff often looks even better for very large basins than for very small ones.

The modelling of the drainage basin of the Baltic Sea was a truly distributed continental-scale model application covering some 1.6×10^6 km² (excluding the area of the Baltic Sea itself) and with territories from 14 nations. The meteorological database consisted of daily observations of air temperatures and precipitation from some 800 synoptic stations that were interpolated to a $1^\circ \times 1^\circ$ grid. The hydrological data were monthly values of river runoff from major rivers in the area, covering 86 % of the catchment. Runoff from the remaining 14 % was found by interpolation using neighbouring stations. The HBV model was applied to 26 sub-basins. Still, model calibration focused on five main subregional drainage basins (Bergström and Graham, 1998; Graham, 2004). This

model application delivered input to the international BALTEX (Baltic Sea Experiment) research programme.

Another important change in HBV applications is related to the increased computational power and availability of data. With increasing computational power, it has become possible (and expected) to quantify uncertainties. This can be done using Monte Carlo approaches, repeated model calibration trials and ensembles of suitable parameter sets. With the increasing availability of hydrological data sets large-sample hydrological modelling has become possible. There are, for instance, data sets with hundreds of catchments for the US, the UK, Chile, Brazil and Australia, which include all data necessary to run the HBV model (e.g. Addor et al., 2017; Alvarez-Garreton et al., 2018; Chagas et al., 2020).

While the use of the HBV model is straightforward for many hydrological questions, answering some of the questions posed in Sect. 1 remains challenging. These questions are mainly related to three types of model applications – land-cover change impacts, ungauged basins and climate change impacts – as discussed in the following.

4.1 Land-cover change impacts

The quantification of land-cover changes on hydrology is an important field of hydrological modelling. Here, it is important to distinguish between change prediction and change detection. It is difficult to use conceptual models for change prediction because model parameter values are not directly linked to certain types of land cover (Seibert and van Meerveld, 2016). For change detection, however, the HBV model has been proven to be useful (Brandt et al., 1988; Seibert et al., 2010; Seibert and McDonnell, 2010) by using the model as a control run (i.e. to simulate the discharge that would have been observed if there had been no change in the catchment). Another approach is the simultaneous calibration of the model to numerous catchments (Hundecha and Bárdossy, 2004). Such a regional calibration is possible for simple models, such as the HBV model, and allows a space-for-time approach to quantify land-use effects on catchment runoff using the HBV model.

4.2 Ungauged catchments

The most demanding type of application of a hydrological model is probably to use it without any calibration at all. This is referred to as prediction in ungauged basins. In the beginning, severe doubts were expressed when a conceptual hydrological model was used for this purpose, due to its need for calibration. Today, however, this is a standard type of application. The need for information about water resources is often so urgent that ungauged model results are accepted. The key to improving the reliability of the results is to generalize and use standard sets of model parameters from a vast number of applications. While there might be some relationships between catchment properties and indi-

vidual parameters (Seibert, 1999), spatial proximity is often more suitable than the use of catchment properties for determining model parameters for ungauged catchments (Merz and Blöschl, 2005; Oudin et al., 2008). In general, transferring entire parameter sets from donor catchments to ungauged catchments results in better simulations than the use of regression equations (Oudin et al., 2008). The model performance will, of course, be lower than if the model is calibrated, but the results are often acceptable, as they are better than no information at all. Thus, even though the HBV model was primarily developed for runoff simulation after calibration, it is currently routinely used in ungauged basins as well (Bergström, 2006). Using ensembles of random parameter sets can result in surprisingly good model results. An indication of the value of using ensembles for HBV model simulations is described by Seibert and Beven (2009), who found that the ensemble mean generally performed better than the best single parameter set (see their Fig. 2). This finding was recently confirmed in studies using large numbers of catchments (Pool and Seibert, 2021; Seibert et al., 2018b). Even better results can be obtained by ensembles of parameter sets that have been calibrated for other catchments (Seibert et al., 2018b). However, the question of how to identify optimal donor catchments is still not fully solved (Pool et al., 2021).

4.3 Climate change impacts

The late 1980s saw the start of a new era in natural sciences. The Intergovernmental Panel on Climate Change (IPCC) scientific assessments were instrumental when climate change became the new focus for many large-scale scientific programmes. Climate change (and its impact on water resources) was identified as a major concern and also became a main focus in hydrological modelling. A number of extreme hydrological situations in various parts of the world emphasized this further. This meant a new role for the HBV model. A joint Nordic project on climate change impacts on runoff and hydropower in the Nordic countries was carried out (Saelthun et al., 1998). It was to be followed by similar projects where, eventually, the Baltic states also participated (Beldring et al., 2006; Nordic Council of Ministers, 2011). In Sweden, the Swedish Regional Climate Modelling Programme, SWECLIM (Rummukainen et al., 2004), and the establishment of its modelling centre, the Rossby Centre, were a breakthrough. New sets of regional climate scenarios were developed. These were made available for hydrological scientists, who could also benefit from the large international network of climate researchers. The interdisciplinary scientific dialogue between meteorologists, hydrologists and oceanographers intensified under several international research programmes. Assessments of the impact of climate change on water resources using global climate models, dynamical downscaling and hydrological modelling with the HBV model became standard practice (Bergström et al., 2001). In an early application, Beldring et al. (2002) used a

gridded HBV model both for hydrological mapping in Norway and for simulations of the impact of climate change on the country's water resources (Beldring et al., 2006).

Today, the HBV model is routinely used to model the impacts of climate change on water resources all over the world such as the Himalayas (Akhtar et al., 2008), the Ouerthe (Driessen et al., 2010), the Rhine (Görgen et al., 2010), Finland (Veijalainen et al., 2010), the Nile (Booij et al., 2011), Sweden (Teutschbein and Seibert, 2012) or Canada (Stahl et al., 2008). One critical question asked when using a hydrological model for climate change studies was whether its parameter values, derived for one climate, could be used for the future climate.

An important question in this respect is the transferability of model applications between different climatic conditions. Differential split-sample tests have indicated that one has to be aware of additional uncertainties when parameter values that have been calibrated based on data from, for instance, cold and dry conditions are transferred to warm and wet conditions (Dakhlaoui et al., 2017). Another challenge is that climate change might affect catchments in several ways. Over longer periods, for instance, the treeline may change. This vegetation change will influence catchment functioning and should be reflected by changed model parameter values. However, in most climate impact studies, the hydrological model parameters are assumed to remain unchanged.

The use of dynamical downscaling of climate scenarios as input to a hydrological model for studies of climate change on water resources, means taking advantage of a chain of scientific efforts and the engagement of thousands of scientists. However, it also means that the overview of the whole modelling process is lost. No one has the full detailed knowledge of all components in the chain of results from global climate models via regional climate models to hydrological models. One important reflection about inconsistency in the modelling process is that each one of the three modelling concepts (global climate model, regional climate model and hydrological catchment model) has its own interpretation of the hydrological cycle. We leave the important task of the harmonization of the modelling of the water cycle in these three categories of models to the coming generation of modellers.

4.4 Model calibration and uncertainties

The three applications discussed above – land-cover change, ungauged catchments and climate change – cover important topics that are directly relevant in practice. However, HBV is also used for research on topics such as model calibration, uncertainty quantification and the value of data (e.g. van Meerveld et al., 2017; Seibert, 1997). Here, the HBV model provides an easy-to-use and computationally efficient model, which is a suitable representation of commonly used models. While many of these studies are more curiosity-driven and address basic issues, addressing these questions also pro-

vides the basis for good model application procedures in more practical applications.

A promising development is model applications using multi-criteria evaluation, i.e. the evaluation of model performances on more variables than just catchment runoff (Bergström et al., 2002). New sources of data (including spatial patterns derived from remote sensing) might present new opportunities (Stisen et al., 2018). Still, there is a challenging balance between more available data and more model complexity to make these data directly usable for model evaluation. While this applies to any type of catchment model, fully distributed, complex models might provide more opportunities to compare field observations directly with model simulation. In contrast, in bucket-type (or conceptual) models, such observations need to be aggregated or otherwise pre-processed to become applicable (Seibert, 2000; Seibert and McDonnell, 2002).

5 Concluding dialogue

So, let us come back to our hypothetical discussion with the PhD Student.

Student: “Cool! This was an interesting lesson in the history of hydrological model development. What has been most surprising with HBV for you?”

Sten: “That it is still in use after all these years. That it has spread worldwide and that its performance is still competitive.”

Jan: “That I am still using it after all these years. Remember that I started by showing how uncertain the HBV model was.”

Student: “So, what happened?”

Jan: “A model can be uncertain and still be useful.”

Sten: “Yes, a model does not have to be perfect. It needs to work under real-life circumstances, like with limited data coverage, and its performance must match requirements. I would also like to recommend that any young modeller spend some time in the field, in the basin subject to modelling, just to get a feeling for Mother Nature’s wonderful complexity.”

Student: “Is there any way we can improve the HBV model?”

Sten: “Sure, there have been improvements over the past decades, and with new data becoming available, there might be new options.”

Jan: “However, it is difficult to improve runoff simulations beyond the performance of relatively simple models such as HBV. Input data limitations are still one of the greatest obstacles.”

Student: “These days, data for hundreds of catchments are made available. Would this not allow us to decide on the best model variant easily?”

Jan: “In theory, yes. The large-sample hydrological modelling allows models to be evaluated in many more catchments, which leads to more robust results. On the other hand,

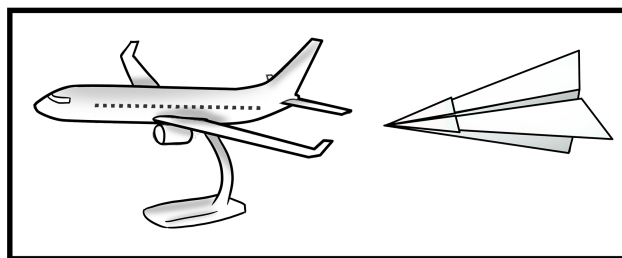


Figure 6. Which model will fly? (drawing by Steph’s Sketches).

performance differences are often small, and different model variants perform best in different types of catchments. Therefore, in practice, it is difficult to demonstrate that certain model variants are better than others.”

Sten: “I am glad to see that a reasoning hydrological modeller is still difficult to beat.”

Student: “What about data other than streamflow? Won’t they allow us to develop better models?”

Jan: “Again, in theory, yes. One could argue that a model (and parameterization) that represents not only streamflow but also internal variables, such as snow accumulation or groundwater storage, is realistically a better model and might provide more reliable simulations, especially when extrapolated beyond the conditions during calibration.”

Sten: “But, on the other hand, for most additional variables, model modifications are required to make these variables comparable with model simulations. This might result in increased parameter uncertainty and model complexity.”

Jan: “This is an interesting balance, and finding good ways to incorporate additional types of observations certainly requires hydrological creativity. This is especially true for novel types of observations such as remotely sensed data or citizen science observations.”

Sten: “So, in some ways, not much has changed in the past 50 years. Sound hydrological reasoning is still needed to decide on the most useful model structures and model testing.”

Student: “Ok. But to be on the safe side, we could represent all details of a catchment in our model. With all of the increasing computational power, this should be possible. Wouldn’t a fully realistic model not be great?”

Jan: “I’m not so sure. We are still limited by the available data to run our models. Even in places where there are a lot of data, there is often a mismatch in the scale of observations and simulations.”

Sten: “... and take a look at these two models of an airplane: one fully detailed scale model and a simple paper airplane (Fig. 6). What would you say? Which model will fly?”

Code and data availability. The data discussed in this paper are available from the corresponding publications referenced in the text and listed in the reference list. Several versions of the HBV model are freely available, including HBV light (<https://www.geo.uzh.ch/en/units/h2k/Services/HBV-Model.html>, University of Zurich, 2020).

Author contributions. JS and SB discussed the outline of the manuscript while enjoying delicious cinnamon rolls and wrote the text together.

Competing interests. At least one of the (co-)authors is a member of the editorial board of *Hydrology and Earth System Sciences*. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Special issue statement. This article is part of the special issue "History of hydrology" (HESS/HGSS inter-journal SI). It is not associated with a conference.

Acknowledgements. We thank the many "HBV colleagues" for stimulating studies and discussions over the years. We greatly appreciate the constructive comments from Keith Beven, Jens Christian Refsgaard, Axel Bronstert and András Bárdossy. Tracy Ewen, Sandra Pool and Marc Vis are acknowledged for providing valuable comments on an earlier version of this contribution.

Review statement. This paper was edited by Keith Beven and reviewed by Jens Christian Refsgaard, András Bárdossy and Axel Bronstert.

References

- Addor, N., Newman, A. J., Mizukami, N., and Clark, M. P.: The CAMELS data set: catchment attributes and meteorology for large-sample studies, *Hydrol. Earth Syst. Sci.*, 21, 5293–5313, <https://doi.org/10.5194/hess-21-5293-2017>, 2017.
- Akhtar, M., Ahmad, N., and Booi, M. J.: The impact of climate change on the water resources of Hindukush-Karakorum-Himalaya region under different glacier coverage scenarios, *J. Hydrol.*, 355, 148–163, <https://doi.org/10.1016/j.jhydrol.2008.03.015>, 2008.
- Alvarez-Garreton, C., Mendoza, P. A., Boisier, J. P., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J., and Ayala, A.: The CAMELS-CL dataset: catchment attributes and meteorology for large sample studies – Chile dataset, *Hydrol. Earth Syst. Sci.*, 22, 5817–5846, <https://doi.org/10.5194/hess-22-5817-2018>, 2018.
- Anderson, E. A.: National Weather Service River Forecast System – Snow Accumulation and Ablation Model, NOAA Technical Memorandum NWS HYDRO, 17, 1–87, 1973.
- Andersson, L.: Hydrological analysis of basin behaviour from soil moisture data, *Nord. Hydrol.*, 19, 1–18, <https://doi.org/10.2166/nh.1988.0001>, 1988.
- Beldring, S., Roald, L. A., and Voksø, A.: Avrenningskart for Norge. Årsmiddelværdier for avrenning 1961–1990 (Map Annual runoff in Norway for the period 1961–1990, in Norwegian), Norwegian Water and Energy Directorate, 2, 49 pp., 2002.
- Beldring, S., Engeland, K., Roald, L. A., Sælthun, N. R., and Voksø, A.: Estimation of parameters in a distributed precipitation-runoff model for Norway, *Hydrol. Earth Syst. Sci.*, 7, 304–316, <https://doi.org/10.5194/hess-7-304-2003>, 2003.
- Beldring, S., Andréasson, J., Bergström, S., Jónsdóttir, J. F., Rogozova, S., Rosberg, J., Suomalainen, M., Tonning, T., Vehviläinen, B., and Veijalainen, N.: Hydrological climate change maps of the Nordic region, in: European Conference on Impacts of Climate Change on Renewable Energy Sources, 5–9 June 2006, Iceland, 2006.
- Beldring, S., Kordzakhia, M., and Kristensen, S. E.: Runoff map of Georgia – Hydrological modelling of water balance, Norwegian Water Resources and Energy Directorate, 61 pp., ISBN 978-82-410-1579-3, 2017.
- Bergström, S.: Utveckling och tillämpning av en digital avrinningsmodell, Sveriges Meteorologiska och Hydrologiska Institut, Notiser och preliminära rapporter, Serie Hydrologi, 22, 1972.
- Bergström, S.: The development of a snow routine for the HBV-2 model, *Nord. Hydrol.*, 6, 73–92, 1975.
- Bergström, S.: Development and application of a conceptual runoff model for Scandinavian catchments, SMHI, Norrköping, Sweden, RHO 7, 134 pp., 1976.
- Bergström, S.: Operational Hydrological Forecasting by Conceptual Models in Sweden, in: Logistics and Benefits of Using Mathematical Models of Hydrologic and Water Resource Systems, edited by: Askew, A. J., Greco, F., and Kindler, J., 61–74 pp., 1978.
- Bergström, S.: The HBV model – its structure and applications, SMHI Reports Hydrology, Norrköping, Sweden, RH 4, 1992.
- Bergström, S.: Experience from applications of the HBV hydrological model from the perspective of prediction in ungauged basins, in: Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment–MOPEX, Proceedings of the Paris (2004) and Foz de Iguacu (2005) workshops, IAHS-AISH Publication, 307, 97–107, 2006.
- Bergström, S. and Forsman, A.: Development of a conceptual deterministic rainfall-runoff model, *Nord. Hydrol.*, 4, 147–170, 1973.
- Bergström, S. and Graham, L. P.: On the scale problem in hydrological modelling, *J. Hydrol.*, 211, 253–265, [https://doi.org/10.1016/S0022-1694\(98\)00248-0](https://doi.org/10.1016/S0022-1694(98)00248-0), 1998.
- Bergström, S. and Jönsson, S.: Tillämpning av HBV-2 modellen på regleringsmagasin i Ångermanälven (Application of the HBV-2 modell to hydropower reservoirs in River Ångermanälven, in Swedish), Swedish Meteorological and Hydrological Institute, HB Rapport 18, Stockholm, Sweden, 1976.

- Bergström, S. and Lindström, G.: Interpretation of runoff processes in hydrological modelling-experience from the HBV approach, *Hydrol. Process.*, 29, 3535–3545, <https://doi.org/10.1002/hyp.10510>, 2015.
- Bergström, S. and Sandberg, G.: Simulation of groundwater response by conceptual models – three case studies, *Nord. Hydrol.*, 14, 71–84, 1983.
- Bergström, S., Carlsson, B., Gardelin, M., Lindström, G., Petterson, A., and Rummukainen, M.: Climate change impacts on runoff in Sweden – Assessments by global climate models, dynamical downscaling and hydrological modelling, *Clim. Res.*, 16, 101–112, 2001.
- Bergström, S., Lindström, G., and Pettersson, A.: Multi-variable parameter estimation to increase confidence in hydrological modelling, *Hydrol. Process.*, 16, 413–421, <https://doi.org/10.1002/hyp.332>, 2002.
- Beven, K.: Changing ideas in hydrology – the case of physically-based models, *J. Hydrol.*, 105, 157–172, 1989.
- Beven, K.: How far can we go in distributed hydrological modelling?, *Hydrol. Earth Syst. Sci.*, 5, 1–12, <https://doi.org/10.5194/hess-5-1-2001>, 2001.
- Beven, K. J.: A discussion of distributed hydrological modelling, in: *Distributed hydrological modelling*, Springer, Dordrecht, The Netherlands, 255–277, 1990.
- Beven, K. J.: On hypothesis testing in hydrology: Why falsification of models is still a really good idea, *WIREs Water*, 5, 1–8, <https://doi.org/10.1002/wat2.1278>, 2018.
- Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, *Hydrol. Sci. B.*, 24, 43–69, <https://doi.org/10.1080/02626667909491834>, 1979.
- Booij, M. J., Tollenaar, D., van Beek, E., and Kwadijk, J. C. J.: Simulating impacts of climate change on river discharges in the Nile basin, *Phys. Chem. Earth*, 36, 696–709, <https://doi.org/10.1016/j.pce.2011.07.042>, 2011.
- Brandt, M. and Bergström, S.: Integration of field data into operational snowmelt-runoff models, *Nord. Hydrol.*, 25, 101–112, <https://doi.org/10.2166/nh.1994.0022>, 1994.
- Brandt, M., Bergström, S., and Gardelin, M.: Modelling the effects of clearcutting on runoff – examples from central Sweden, *Ambio*, 17, 307–313, 1988.
- Braun, L. N. and Renner, C. B.: Application of a conceptual runoff model in different physiographic regions of Switzerland, *Hydrolog. Sci. J.*, 37, 217–231, <https://doi.org/10.1080/02626669209492583>, 1992.
- Breuer, L., Huisman, J. A., Willems, P., Bormann, H., Bronstert, A., Croke, B. F. W., Frede, H.-G., Gräff, T., Hubrechts, L., Jakeman, A. J., Kite, G., Lanini, J., Leavesley, G., Lettenmaier, D. P., Lindström, G., Seibert, J., Sivapalan, M., and Viney, N. R.: Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: Model intercomparison with current land use, *Adv. Water Resour.*, 32, 129–146, <https://doi.org/10.1016/j.advwatres.2008.10.003>, 2009.
- Brooks, K. N., Davis, E. M., Kuehl, D. W., and the United States Army Corps of Engineers, North Pacific Division: Program Description and User Manual for SSARR Model: Streamflow Synthesis and Reservoir Regulation, U.S. Army Engineer Division, North Pacific, USA, 188 pp., 1975.
- Burnash, R. J. C.: The NWS River Forecast System-catchment modeling, in: *Computer Models of Watershed Hydrology*, edited by: Singh, V., Water Resources Publication, Colorado, USA, 311–366, 1995.
- Chagas, V. B. P., Chaffe, P. L. B., Addor, N., Fan, F. M., Fleischmann, A. S., Paiva, R. C. D., and Siqueira, V. A.: CAMELS-BR: hydrometeorological time series and landscape attributes for 897 catchments in Brazil, *Earth Syst. Sci. Data*, 12, 2075–2096, <https://doi.org/10.5194/essd-12-2075-2020>, 2020.
- Crawford, H. H. and Linsley, R. K.: *Digital Simulation in Hydrology: Stanford Watershed Model IV*, Stanford University, Dept. of Civil Engineering. Technical report, no. 39. 210 pp., 1966.
- Dakhlaoui, H., Ruelland, D., Tramblay, Y., and Bargaoui, Z.: Evaluating the robustness of conceptual rainfall-runoff models under climate variability in northern Tunisia, *J. Hydrol.*, 550, 201–217, <https://doi.org/10.1016/j.jhydrol.2017.04.032>, 2017.
- Das, T., Bárdossy, A., Zehe, E., and He, Y.: Comparison of conceptual model performance using different representations of spatial variability, *J. Hydrol.*, 356, 106–118, <https://doi.org/10.1016/j.jhydrol.2008.04.008>, 2008.
- Dawdy, D. R. and O'Donnell, T.: Mathematical models of catchment behavior, *J. Hydr. Eng. Div.-ASCE*, 91, 123–137, <https://doi.org/10.1061/JYCEAJ.0001271>, 1965.
- Driessen, T. L. A., Hurkmans, R. T. W. L., Terink, W., Hazenberg, P., Torfs, P. J. J. F., and Uijlenhoet, R.: The hydrological response of the Ourthe catchment to climate change as modelled by the HBV model, *Hydrol. Earth Syst. Sci.*, 14, 651–665, <https://doi.org/10.5194/hess-14-651-2010>, 2010.
- Flödeskommittén: Riktlinjer för bestämning av dimensionerande flöden för dammanläggningar. Slutrapport från Flödeskommittén [Final report from the Swedish committee on spillway design], Statens Vattenfallsverk, Svenska Kraftverksföreningen och Sveriges Meteorologiska och Hydrologiska Institut, <https://www.svk.se/siteassets/3.sakerhet-och-hallbarhet/dammsakerhet/rapporter-och-yttranden/riktlinjer-for-bestamning-av-dimensionerande-floden-for-dammanlaggningar-2022.pdf> (last access: 7 March 2022), 1990 (in Swedish, updated 2022).
- Girons Lopez, M., Vis, M. J. P., Jenicek, M., Griessinger, N., and Seibert, J.: Assessing the degree of detail of temperature-based snow routines for runoff modelling in mountainous areas in central Europe, *Hydrol. Earth Syst. Sci.*, 24, 4441–4461, <https://doi.org/10.5194/hess-24-4441-2020>, 2020.
- Görgen, K., Beersma, J., Brahmer, G., Buiteveld, H., Carambia, M., Krahe, P., Nilson, E., Lammersen, R., Perrin, C., Görgen, K., Beersma, J., Brahmer, G., Buiteveld, H., and Carambia, M.: Assessment of climate change impacts on discharge in the Rhine River Basin: Results of the RheinBlick2050 project, International Commission for the Hydrology of the Rhine Basin (CHR), 228 pp., 2010.
- Graham, L. P.: Climate change effects on river flow to the Baltic Sea, *Ambio*, 33, 235–241, <https://doi.org/10.1579/0044-7447-33.4.235>, 2004.
- Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *J. Hydrol.*, 377, 80–91, <https://doi.org/10.1016/j.jhydrol.2009.08.003>, 2009.
- Hagg, W., Braun, L. N., Kuhn, M., and Nesgaard, T. I.: Modelling of hydrological response to climate change in glacierized Central Asian catchments, *J. Hydrol.*, 332, 40–53, <https://doi.org/10.1016/j.jhydrol.2006.06.021>, 2007.

- Hrachowitz, M. and Clark, M. P.: HESS Opinions: The complementary merits of competing modelling philosophies in hydrology, *Hydrol. Earth Syst. Sci.*, 21, 3953–3973, <https://doi.org/10.5194/hess-21-3953-2017>, 2017.
- Huang, S., Eisner, S., Magnusson, J. O., Lussana, C., Yang, X., and Beldring, S.: Improvements of the spatially distributed hydrological modelling using the HBV model at 1 km resolution for Norway, *J. Hydrol.*, 577, 123585, <https://doi.org/10.1016/j.jhydrol.2019.03.051>, 2019.
- Hundecha, Y. and Bárdossy, A.: Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model, *J. Hydrol.*, 292, 281–295, <https://doi.org/10.1016/j.jhydrol.2004.01.002>, 2004.
- Jost, G., Moore, R. D., Menounos, B., and Wheate, R.: Quantifying the contribution of glacier runoff to streamflow in the upper Columbia River Basin, Canada, *Hydrol. Earth Syst. Sci.*, 16, 849–860, <https://doi.org/10.5194/hess-16-849-2012>, 2012.
- Killingtveit, Å. and Aam, S.: En fordelt modell for snoakkumulering og -avsmelting (A distributed model for snow accumulation and melt, in Norwegian), Norwegian Institute of Technology, Division of Hydraulic Engineering, Trondheim, Norway, 1978.
- Killingtveit, Å. and Saelthun, N. R.: Hydrological models, in: *Hydrology. Hydropower Development Volume No. 7*, Norwegian Institute of Technology, Division of Hydraulic Engineering, Trondheim, Norway, 1995.
- Kollet, S., Gasper, F., Brdar, S., Goergen, K., Hendricks-Franssen, H. J., Keune, J., Kurtz, W., Küll, V., Pappenberger, F., Poll, S., Trömel, S., Shrestha, P., Simmer, C., and Sulis, M.: Introduction of an experimental terrestrial forecasting/monitoring system at regional to continental scales based on the terrestrial systems modeling platform (v1.1.0), *Water (Switzerland)*, 10, 1697, <https://doi.org/10.3390/w10111697>, 2018.
- Krysanova, V., Bronstert, A., and Müller-Wohlfeil, D.: Modelling river discharge for large drainage basins: From lumped to distributed approach, *Hydrolog. Sci. J.*, 44, 313–331, <https://doi.org/10.1080/02626669909492224>, 1999.
- Li, H., Beldring, S., and Xu, C. Y.: Implementation and testing of routing algorithms in the distributed Hydrologiska Byråns Vattenbalansavdelning model for mountainous catchments, *Hydrol. Res.*, 45, 322–333, <https://doi.org/10.2166/nh.2013.009>, 2014.
- Li, H., Beldring, S., Xu, C. Y., Huss, M., Melvold, K., and Jain, S. K.: Integrating a glacier retreat model into a hydrological model – Case studies of three glacierised catchments in Norway and Himalayan region, *J. Hydrol.*, 527, 656–667, <https://doi.org/10.1016/j.jhydrol.2015.05.017>, 2015.
- Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *J. Geophys. Res.*, 99, 14415–14428, <https://doi.org/10.1029/94jd00483>, 1994.
- Lindell, S., Carlsson, B., Sanner, H., Reihan, A., and Vedom, R.: Application of the integrated hydrological modelling system IHMS-HBV to pilot basin in Estonia, *SMHI Hydrology*, 67, 66 pp., 1996.
- Lindström, G.: A simple automatic calibration routine for the HBV model, *Nord. Hydrol.*, 28, 153–168, 1997.
- Lindström, G. and Bergström, S.: Improving the HBV and PULSE models by use of temperature anomalies, *Vannet i Nord.*, No. 1, 16–23, 1992.
- Lindström, G. and Harlin, J.: Spillway design floods, in: Sweden: II. Applications and sensitivity analysis, *Hydrolog. Sci. J.*, 37, 521–539, <https://doi.org/10.1080/02626669209492616>, 1992.
- Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, *J. Hydrol.*, 201, 272–288, 1997.
- Mandeville, A. N., O’Connell, P. E., Sutcliffe, J. V., and Nash, J. E.: River flow forecasting through conceptual models part III – The Ray catchment at Grendon Underwood, *J. Hydrol.*, 11, 109–128, [https://doi.org/10.1016/0022-1694\(70\)90098-3](https://doi.org/10.1016/0022-1694(70)90098-3), 1970.
- Merz, R. and Blöschl, G.: Regionalisation of catchment model parameters, *J. Hydrol.*, 287, 95–123, <https://doi.org/10.1016/j.jhydrol.2003.09.028>, 2004.
- Merz, R. and Blöschl, G.: Flood frequency regionalisation – spatial proximity vs. catchment attributes, *J. Hydrol.*, 302, 283–306, <https://doi.org/10.1016/j.jhydrol.2004.07.018>, 2005.
- Michel, C.: Que peut-on faire en hydrologie avec modèle conceptuel à un seul paramètre?, *La Houille Blanche*, 69, 39–44, <https://doi.org/10.1051/lhb/1983004>, 1983.
- Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models Part I – A discussion of principles, *J. Hydrol.*, 10, 282–290, [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6), 1970.
- Nielsen, S. A. and Hansen, E.: Numerical simulation of the rainfall-runoff process on a daily basis, *Nord. Hydrol.*, 4, 171–190, 1973.
- Nordic Council of Ministers: Climate Change and Energy Systems: Impacts, Risks and Adaptation in the Nordic and Baltic countries, Nordic Council of Ministers, Copenhagen, Denmark, Report TemaNord 2011:502, 2011.
- Norstedt, U., Brandesten, C.-O., Bergström, S., Harlin, J., and Lindström, G.: Re-evaluation of hydrological dam safety in Sweden, *International Water Power and Dam Construction IWPCDM*, 44, 43–49, 1992.
- Northern Climate ExChange: Projected future changes in glaciers and their contribution to discharge of the Yukon River at Whitehorse, Yukon Research Centre, Yukon College, Whitehorse, YT, Canada, 53 pp., 2014.
- O’Connell, P. E., Nash, J. E., and Farrell, J. P.: River flow forecasting through conceptual models part II – The Brosna catchment at Ferbane, *J. Hydrol.*, 11, 317–329, [https://doi.org/10.1016/0022-1694\(70\)90221-0](https://doi.org/10.1016/0022-1694(70)90221-0), 1970.
- Oudin, L., Andréassian, V., Perrin, C., Michel, C., and Le Moine, N.: Spatial proximity, physical similarity, regression and ungauged catchments: A comparison of regionalization approaches based on 913 French catchments, *Water Resour. Res.*, 44, 1–15, <https://doi.org/10.1029/2007WR006240>, 2008.
- Peel, M. C. and McMahon, T. A.: Historical development of rainfall-runoff modeling, *WIREs Water*, 7, 1–15, <https://doi.org/10.1002/wat2.1471>, 2020.
- Perrin, C., Michel, C., and Andréassian, V.: Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, *J. Hydrol.*, 242, 275–301, 2001.
- Perrin, C., Michel, C., and Andréassian, V.: Improvement of a parsimonious model for streamflow simulation, *J. Hydrol.*, 279, 275–289, [https://doi.org/10.1016/S0022-1694\(03\)00225-7](https://doi.org/10.1016/S0022-1694(03)00225-7), 2003.
- Pool, S. and Seibert, J.: Gauging ungauged catchments – Active learning for the timing of point discharge observations in combination with continuous water level measurements, *J. Hydrol.*,

- 598, 126448, <https://doi.org/10.1016/j.jhydrol.2021.126448>, 2021.
- Pool, S., Vis, M., and Seibert, J.: Evaluating model performance: towards a non-parametric variant of the Kling-Gupta efficiency Kling-Gupta efficiency, *Hydrolog. Sci. J.*, 63, 1941–1953, <https://doi.org/10.1080/02626667.2018.1552002>, 2018.
- Pool, S., Vis, M., and Seibert, J.: Regionalization for Ungauged Catchments – Lessons Learned From a Comparative Large-Sample Study, *Water Resour. Res.*, 57, 1–25, <https://doi.org/10.1029/2021wr030437>, 2021.
- Quick, M. C. and Pipes, A.: A combined snowmelt and rainfall runoff model, *Can. J. Civil Eng.*, 3, 449–460, 1976.
- Raab, B. and Vedin, H. E.: *Climate, Lakes and Rivers*. National Atlas of Sweden, Almqvist and Wiksell International, Stockholm, Sweden, 1995.
- Rango, A. and Martinec, J.: Application of a snowmelt-runoff model using Landsat data, *Nord. Hydrol.*, 10, 225–238, <https://doi.org/10.2166/nh.1979.0006>, 1979.
- Refsgaard, J. C. and Knudsen, J.: Operational Validation and Intercomparison of Different Types of Hydrological Models, *Water Resour. Res.*, 32, 2189–2202, <https://doi.org/10.1029/96WR00896>, 1996.
- Refsgaard, J. C., Storm, B., and Clausen, T.: Système Hydrologique Européen (SHE): review and perspectives after 30 years development in distributed physically-based hydrological modelling, *Hydrol. Res.*, 41, 355–377, <https://doi.org/10.2166/nh.2010.009>, 2010.
- Refsgaard, J. C., Højberg, A. L., He, X., Hansen, A. L., Rasmussen, S. H., and Stisen, S.: Where are the limits of model predictive capabilities?, *Hydrol. Process.*, 30, 4956–4965, <https://doi.org/10.1002/hyp.11029>, 2016.
- Rosenbrock, H. H.: An automatic method for finding the greatest or least value of a function, *Comput. J.*, 7, 175–184, 1960.
- Rummukainen, M., Bergström, S., Persson, G., Rodhe, J., and Tjernström, M.: The Swedish regional climate modelling programme, SWECLIM: A review, *Ambio*, 33, 176–182, <https://doi.org/10.1579/0044-7447-33.4.176>, 2004.
- Saelthun, N. R.: Comparative tests of hydrological models for operational use on Norwegian catchments. An inter-institutional project, in: Contributions to Nordic Hydrological Conference, 31 July –3 August 1978, Helsinki, Norwegian Water Resources and Electricity Board, Hydrological Division, Oslo, 5.1–5.11, 1978.
- Saelthun, N. R.: The “Nordic” HBV Model. Description and documentation of the model version developed for the project Climate Change and Energy Production, NVE Publication, Nor. Water Resour. Energy Adm. Oslo, 7, 26 pp., 1996.
- Saelthun, N. R., Aittoniemi, P., Bergström, S., Einarsson, K., Jóhannesson, T., Lindström, G., Ohlsson, P.-E., Thomsen, T., Vehviläinen, B., and Aamodt, K. O.: Climate change impacts on runoff and hydropower production in the Nordic countries, Nordic Council of Ministers, Copenhagen, Denmark, Report TemaNord 1988:552, 1998.
- Seibert, J.: Estimation of Parameter Uncertainty in the HBV Model, *Nord. Hydrol.*, 28, 247–262, 1997.
- Seibert, J.: Regionalisation of parameters for a conceptual rainfall-runoff model, *Agric. For. Meteorol.*, 98–99, 279–293, [https://doi.org/10.1016/S0168-1923\(99\)00105-7](https://doi.org/10.1016/S0168-1923(99)00105-7), 1999.
- Seibert, J.: Multi-criteria calibration of a conceptual runoff model using a genetic algorithm, *Hydrol. Earth Syst. Sci.*, 4, 215–224, <https://doi.org/10.5194/hess-4-215-2000>, 2000.
- Seibert, J. and Beven, K. J.: Gauging the ungauged basin: how many discharge measurements are needed?, *Hydrol. Earth Syst. Sci.*, 13, 883–892, <https://doi.org/10.5194/hess-13-883-2009>, 2009.
- Seibert, J. and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, *Water Resour. Res.*, 38, 23-1–23-14, <https://doi.org/10.1029/2001WR000978>, 2002.
- Seibert, J. and McDonnell, J. J.: Land-cover impacts on streamflow: a change-detection modelling approach that incorporates parameter uncertainty, *Hydrolog. Sci. J.*, 55, 316–332, <https://doi.org/10.1080/02626661003683264>, 2010.
- Seibert, J. and van Meerveld, H. J.: Hydrological change modeling: Challenges and opportunities, *Hydrol. Process.*, 30, 4966–4971, <https://doi.org/10.1002/hyp.10999>, 2016.
- Seibert, J. and Vis, M. J. P.: Teaching hydrological modeling with a user-friendly catchment-runoff-model software package, *Hydrol. Earth Syst. Sci.*, 16, 3315–3325, <https://doi.org/10.5194/hess-16-3315-2012>, 2012.
- Seibert, J., McDonnell, J. J., and Woodsmith, R. D.: Effects of wildfire on catchment runoff response: a modelling approach to detect changes in snow-dominated forested catchments, *Hydrol. Res.*, 41, 378–390, <https://doi.org/10.2166/nh.2010.036>, 2010.
- Seibert, J., Vis, M. J. P., Kohn, I., Weiler, M., and Stahl, K.: Technical note: Representing glacier geometry changes in a semi-distributed hydrological model, *Hydrol. Earth Syst. Sci.*, 22, 2211–2224, <https://doi.org/10.5194/hess-22-2211-2018>, 2018a.
- Seibert, J., Vis, M. J. P., Lewis, E., and van Meerveld, H. J.: Upper and lower benchmarks in hydrological modelling, *Hydrol. Process.*, 32, 1120–1125, <https://doi.org/10.1002/hyp.11476>, 2018b.
- Singh, V. P.: *Computer Models of Watershed Hydrology*, Water Resources Publications, Colorado, USA, 1995.
- Stahl, K., Moore, R. D., Shea, J. M., Hutchinson, D., and Cannon, A. J.: Coupled modelling of glacier and streamflow response to future climate scenarios, *Water Resour. Res.*, 44, W02422, <https://doi.org/10.1029/2007WR005956>, 2008.
- Stisen, S., Koch, J., Sonnenborg, T. O., Refsgaard, J. C., Bircher, S., Ringgaard, R., and Jensen, K. H.: Moving beyond run-off calibration–Multivariable optimization of a surface–subsurface–atmosphere model, *Hydrol. Process.*, 32, 2654–2668, <https://doi.org/10.1002/hyp.13177>, 2018.
- Sugawara, M.: Automatic calibration of the tank model, *Hydrol. Sci. B.*, 24, 375–388, <https://doi.org/10.1080/02626667909491876>, 1979.
- Svenska_Kraftnät, Svensk_Energi and SveMin: Riktlinjer för bestämning av dimensionerande flöden för dammanläggningar (Guidelines for design flood estimation for dams, in Swedish), <https://www.svk.se/siteassets/3.sakerhet-och-hallbarhet/dammsakerhet/rapporter-och-yttranden/riktlinjer-for-bestamning-av-dimensionerande-floden-for-dammanlaggningar-2022.pdf> (last access: 7 March 2022), 2022.
- Teutschbein, C. and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, *J. Hydrol.*, 456–457, 12–29, <https://doi.org/10.1016/j.jhydrol.2012.05.052>, 2012.
- Todini, E.: The ARNO rainfall-runoff model, *J. Hydrol.*, 175, 339–382, [https://doi.org/10.1016/S0022-1694\(96\)80016-3](https://doi.org/10.1016/S0022-1694(96)80016-3), 1996.

- University of Zurich: HBV-light model software, <https://www.geo.uzh.ch/en/units/h2k/Services/HBV-Model.html> (last access: 10 March 2022), 2020.
- van Meerveld, H. J. I., Vis, M. J. P., and Seibert, J.: Information content of stream level class data for hydrological model calibration, *Hydrol. Earth Syst. Sci.*, 21, 4895–4905, <https://doi.org/10.5194/hess-21-4895-2017>, 2017.
- Vehviläinen, B.: Operational spring time forecasting difficulties and improvements (Finland), *Nord. Hydrol.*, 17, 363–370, <https://doi.org/10.2166/nh.1986.0027>, 1986.
- Veijalainen, N., Lotsari, E., Alho, P., Vehviläinen, B., and Käyhkö, J.: National scale assessment of climate change impacts on flooding in Finland, *J. Hydrol.*, 391, 333–350, <https://doi.org/10.1016/j.jhydrol.2010.07.035>, 2010.
- Weiler, M., Seibert, J., and Stahl, K.: Magic components—why quantifying rain, snowmelt, and icemelt in river discharge is not easy, *Hydrol. Process.*, 32, 160–166, <https://doi.org/10.1002/hyp.11361>, 2018.
- WMO: Intercomparison of conceptual models used in operational hydrological forecasting, World Meteorological Organization, Geneva, Switzerland, Operational Hydrology Report No. 7, WMO No. 429, 172 pp., 1975.
- WMO: Intercomparison of Models of Snowmelt, World Meteorological Organization, Geneva, Switzerland, Operational Hydrology Report No. 23, WMO No. 646, 482 pp., 1986.
- WMO: Real-time Intercomparison of Hydrological Models: Report of the Workshop Held in Vancouver, Canada from 30 July to 8 August 1987, Technical reports to the Commission for Hydrology, Secretariat of the World Meteorological Organization, Geneva, Switzerland, 1988.
- Zhao, R.-J.: The Xinanjiang model applied in China, *J. Hydrol.*, 135, 371–381, [https://doi.org/10.1016/0022-1694\(92\)90096-E](https://doi.org/10.1016/0022-1694(92)90096-E), 1992.