Jan Seibert

Introduction to hydrological modelling and the HBV model

How do we see reality?

Figur 3.2. Vi väljer och omtolkar den information som vi sedan baserar våra subjektiva tankemoder på. (Av Nils Peterson, ur Forskning och Framsteg.)
How does a hydrologist see reality?

A system is a group of components which form a whole, interaction between the components leads to a certain system behavior.

<table>
<thead>
<tr>
<th>Force on system</th>
<th>System behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>Output</td>
</tr>
</tbody>
</table>

Driving variables

A model is a system which we have chosen to represent the essential characteristics of another system.
Why use a model?

- Forecast / Prediction
- Instead of expensive and time consuming measurements
- 'What happens if' questions
- Summarize and test knowledge
- Education
Classification of models

- Symbolic (mathematic) - scale-/ analog model
- Empirical - conceptual - physical
- Linear - non-linear
- Lumped - distributed
- Deterministic - stochastic
- Static - dynamic
- Time-variant - time-invariant
- Event based - continuous
- Complete - partial
- General - special purpose

The HBV model ...

- ... has been developed by Sten Bergström
- ... is a conceptual model for runoff simulation
- ... has a simple structure
- ... is semi-distributed, i.e., allows to divide the catchment into subbasins, elevation and vegetation zones
- ... is easy to understand, learn and apply
- ... has been applied to many catchments in Sweden and abroad
- ... provided good results in most applications
- ... has become a standard tool for runoff studies in the Nordic countries
- ... needs a moderate amount of input data
- ... can be run on a PC (286 or better)
- ... exists in different versions (Swedish, Norwegian, Finnish, Swiss, ...)
- ... (partly) used in other models
IHMS - Integrated Hydrological Modeling System

30+ years ago ...

First successful application of the HBV-model (1972)
... whereas the HBV-model is still in pretty good shape

**Input data**

- Areal precipitation
  - Weighted mean
  - Elevation zones (lapse rate ~10% per 100 m)
- Temperature
  - Weighted mean
  - Elevation zones (lapse rate 0.6 °C per 100 m)
- Potential evaporation
  - Penman formula or measurements
  - Usually long-term monthly mean values
Input data (cont.)

- Correction of long-term potential evaporation

\[ E_{\text{pot}}(t) = (1 + C_{\text{ET}}(T(t) - T_M)) \cdot E_{\text{pot},M} \]

(BUT: \( 2 \cdot E_{\text{pot},M} \geq E_{\text{pot}}(t) \geq 0 \))

- Potential evaporation at day \( t \) [mm d\(^{-1}\)]
- Correction factor \([\degree C^{-1}]\)
- Temperature at day \( t \) [\degree C]
- Long-term mean temperature for this day of the year [\degree C]
- Long-term mean evaporation for this day of the year [mm d\(^{-1}\)]

Model structure

<table>
<thead>
<tr>
<th>Routine</th>
<th>Input data</th>
<th>Output data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snow routine</td>
<td>Precipitation, temperature</td>
<td>Snow pack, snow-melt</td>
</tr>
<tr>
<td>Soil routine</td>
<td>Pot. evaporation, precipitation, snowmelt</td>
<td>Act. evaporation, 'soil moisture', groundwater recharge</td>
</tr>
<tr>
<td>Response function</td>
<td>Groundwater recharge, (pot. evaporation)</td>
<td>Runoff to stream, 'Groundwater level'</td>
</tr>
<tr>
<td>Routing routine</td>
<td>Runoff to stream</td>
<td>Runoff at outlet</td>
</tr>
</tbody>
</table>

PRECIPITATION

Snow routine

Soil moisture routine

Response function

Routing routine

RUNOFF
Snow routine

- Accumulation of precipitation as snow if temperature $< T_T$
- Degree-day method for snowmelt
  (degree-day factor $C_{FMAX}$ [mm d$^{-1}$ °C$^{-1}$])

\[ M = C_{FMAX} (T - T_T) \quad [\text{mm d}^{-1}] \]

$C_{FMAX}$ typically around 4 mm d$^{-1}$ °C$^{-1}$, lower values for forested areas compared to open areas

Snow routine (cont.)

- Snow pack retains melt water until amount exceeds a certain portion ($C_{WH}$, usually 0.1) of the water equivalent of the snow pack.
- When temperatures decreases below $T_T$ water refreezes again

\[ M = C_{FR} C_{FMAX} (T_T - T) \quad [\text{mm h}^{-1}] \]

$C_{FR} = \sim 0.05 \quad [-]$

- All precipitation which is simulated to be snow is multiplied by a correction factor, $S_{CF} [-]$
- Calculations are carried out separately for each elevation and vegetation zone
Effect of $T_T$

$Q \text{ [mm/day]}$

- $T_T = 0 \degree C$
- $T_T = -0.5 \degree C$

Effect of $C_{FMAX}$

$Q \text{ [mm/day]}$

- $C_{FMAX} = 3.5 \text{ mm/}(\degree \text{C day})$
- $C_{FMAX} = 2.5 \text{ mm/}(\degree \text{C day})$
Soil routine

- $F_C$ maximum soil moisture storage [mm]
- $\beta$ shape parameter [-]

$$\text{recharge} = \left( \frac{S_{mm}}{F_C} \right)^\beta$$

Soil routine (cont.)

- $F_C$ = maximum soil moisture storage [mm]
- $L_P$ = factor defining reduction of evaporation [-]
Effect of $FC$

$FC = 150$ mm

$FC = 200$ mm

Effect of $\beta$

$BETA = 2$

$BETA = 5$
Response function

\[ Q_0 = K_0 \cdot (SUZ - UZL) \]
\[ Q_1 = K_1 \cdot SUZ \]
\[ Q_2 = K_2 \cdot SLZ \]

- 5 parameter, two boxes
- Upper box: shallow ground water, lower box: deeper groundwater

**Response function (cont.)**

- Slope of the recession:
  - Peaks: \( K_0 + K_1 + K_2 \)
  - Intermediate: \( K_1 + K_2 \)
  - Baseflow: \( K_2 \)
- Thresholds:
  - \( Q(T_1) = P_{ERC} + K_1 UZL \)
  - \( Q(T_2) = P_{ERC} \)
Effect of $K_0$

Effect of $U_{ZL}$
Effect of $K_1$

- $K_1 = 0.08 \text{ 1/day}$
- $K_1 = 0.12 \text{ 1/day}$

Effect of $P_{ERC}$

- $P_{ERC} = 0.8 \text{ mm/day}$
- $P_{ERC} = 2 \text{ mm/day}$
Effect of $K_2$

![Graph showing the effect of $K_2$ on runoff.](image)

- $K_2 = 0.02 \text{ 1/day}$
- $K_2 = 0.005 \text{ 1/day}$

Routing routine

- One parameter: MAXBAS [d]
  base in an equilateral triangular weighting function
Effect of \textit{MAXBAS}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{effect_of_maxbas.png}
\caption{Effect of \textit{MAXBAS} on Q [mm/day] for different values of \textit{MAXBAS}.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{hbv_overview.png}
\caption{HBV overview diagram.}
\end{figure}
Calibration of the HBV model

• Trial-and-error or automatic calibration
• Different criteria can be used to assess the fit of simulated runoff to observed runoff:
  - visual inspection of plots with $Q_{\text{Sim}}$ and $Q_{\text{Obs}}$
  - statistical criteria
• The coefficient of efficiency, $R_{\text{eff}}$, is normally used for assessment of simulations by the HBV model:

$$R_{\text{eff}} = 1 - \frac{\sum(Q_{\text{sim}}(t) - Q_{\text{obs}}(t))^2}{\sum(Q_{\text{obs}}(t) - \bar{Q}_{\text{obs}})^2}$$

$R_{\text{eff}} = 1$ -> Perfect fit; $R_{\text{eff}} = 0$ -> Simulation as good (or poor) as the constant-value prediction; $R_{\text{eff}} < 0$ -> Very poor fit
$R_{\text{eff}}$ not equal $r^2$ !!!

Calibration of the HBV model (cont.)

• Calibration period should include a variety of hydrological events
• Normally 5 to 10 years sufficient to calibrate the model
• Split-sample test: test of model performance with calibrated parameters for an independent period
• Different objective functions (e.g., efficiency based on log Q, or ‘punishment’ for volume error)

• Problems
  - Parameter uncertainty
  - Internal model consistency
  - Parameter values for ungauged catchments
Applications of the HBV model

The HBV model is/can be used ... 
• ... to extend runoff data series (or filling gaps) 
• ... for data quality control 
• ... for water balance studies 
• ... for runoff forecasting (flood warning and reservoir operation) 
• ... to compute design floods for dam safety 
• ... to investigate the effects of changes within the catchment 
• ... to simulate discharge from ungauged catchments 
• ... to simulate climate change effects
1. Beginners exercises

Calibration using synthetic data (HBVland)

Calibrate the HBV model for the HBV-land catchment for the period 1981-01-01 to 1981-08-31 (warm-up period starting at 1981-01-01). This catchment behaves exactly as the HBV model sees the world, therefore you might be able to achieve a perfect fit (R^eff=1).

Real-world calibration

Try to calibrate the model. It is a good idea to start with the snow routine to get the spring flood right, then work on the soil routine parameters to get the water balance ok and finally fix the response function. You might have to do this in iterations.

During calibration look also on different variables, i.e. soil moisture, storage in the upper groundwater box, ...

Once you have reached a perfect fit (or have received the 'true' parameter values by kindly asking your teacher), you may again change parameter values and study the effects of different parameter values.

Change one (or two) of the following parameter: TT, CFMAX, FC, BETA, LP, K0, K1, K2, PERC, UZL, MAXBAS, SFCF.

Discuss - before running the model - what effect you expect (i.e. more runoff during spring, slower response to rain, ...).

Run the model and look on the deviation of the simulated runoff (red line) from the recorded runoff (blue line).

Make a note of each change of a parameter value and its effect to the simulation.

Change the parameter value back to its original value.

Continue with 3.
2. Advanced exercises

- **Calculation of design flood**
  - Calibration of model
  - Hypothetical precipitation and temperature
  - Simulation of design flood

- **Estimation of land-use change effect**

3. Expert exercises

- **Recode snow routine**
- **Design your own interception routine**

**Example:**

A synthetic sequence of extreme precipitation has been derived by meteorologists (Table 1). Now it is your task to simulate the flood that this sequence would cause for the River Fyris at Vattholma (Uppland). In other words, you should estimate a new synthetic sequence of extreme precipitation that is likely to cause a flood that is similar to the flood that occurred in 1975 to 1985. The model is then calibrated. However, the calibration can be quite sensitive to changes in the parameter set. The model is far from well-calibrated.

You have to complete three steps:

1) Calibration
   - Choose a period for which you replace the observed precipitation by the synthetic precipitation. Use the period 870901 to 870831 for calibration (with the warming-up period starting at 810901).

2) Simulation of flood
   - Before you can use your calibrated model for any prediction it is important that you test your parameter set for an independent time period. Use the period 870831 to 871131 for this test. Is this fit worse? Can you give an explanation? How will your design flood be affected?

3) Validation
   - Make a backup copy of the file.
   - Open the file plotted in a text editor (or Excel).
   - Choose a period for which you replace the observed precipitation by the synthetic precipitation. Use the period 870901 to 870831 for calibration (with the warming-up period starting at 810901).
   - Save the file (if you use Excel choose the format '*.csv' (colon-separated), but use the name 'ptq.dat').
   - Restart MATLAB and run the model. Check the peak value of your simulated flood.
   - Return to the backup file, choose a different period and continue with 3. Do this 5-10 times.

What influences the size of the simulated flood?

Under which conditions becomes the simulated flood largest/smallest?

Table 1

<table>
<thead>
<tr>
<th>Day</th>
<th>P [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>

**Example:**

Data from the Kassjön basin in Medelpad, Sweden is used in this exercise. In the file ex5_snow7376.dat you find precipitation [mm], temperature [°C] and depth of the snow pack [mm water equivalent].

A synthetic sequence of extreme precipitation has been derived by meteorologists. In other words, you should estimate a new synthetic sequence of extreme precipitation that is likely to cause a flood that is similar to the flood that occurred in 1975 to 1985. The model is then calibrated. However, the calibration can be quite sensitive to changes in the parameter set. The model is far from well-calibrated.

You have to complete three steps:

1) Calibration
   - Choose a period for which you replace the observed precipitation by the synthetic precipitation. Use the period 870901 to 870831 for calibration (with the warming-up period starting at 810901).

2) Validation
   - Before you can use your calibrated model for any prediction it is important that you test your parameter set for an independent time period. Use the period 870831 to 871131 for this test. Is this fit worse? Can you give an explanation? How will your design flood be affected?

3) Simulation of flood
   - Make a backup copy of the file.
   - Open the file plotted in a text editor (or Excel).
   - Choose a period for which you replace the observed precipitation by the synthetic precipitation. Use the period 870901 to 870831 for calibration (with the warming-up period starting at 810901).
   - Save the file (if you use Excel choose the format '*.csv' (colon-separated), but use the name 'ptq.dat').
   - Restart MATLAB and run the model. Check the peak value of your simulated flood.
   - Return to the backup file, choose a different period and continue with 3. Do this 5-10 times.

What influences the size of the simulated flood?

Under which conditions becomes the simulated flood largest/smallest?

Table 1

<table>
<thead>
<tr>
<th>Day</th>
<th>P [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>

**Example:**

Data from the Kassjön basin in Medelpad, Sweden is used in this exercise. In the file ex5_snow7376.dat you find precipitation [mm], temperature [°C] and depth of the snow pack [mm water equivalent].

A synthetic sequence of extreme precipitation has been derived by meteorologists. In other words, you should estimate a new synthetic sequence of extreme precipitation that is likely to cause a flood that is similar to the flood that occurred in 1975 to 1985. The model is then calibrated. However, the calibration can be quite sensitive to changes in the parameter set. The model is far from well-calibrated.

You have to complete three steps:

1) Calibration
   - Choose a period for which you replace the observed precipitation by the synthetic precipitation. Use the period 870901 to 870831 for calibration (with the warming-up period starting at 810901).

2) Validation
   - Before you can use your calibrated model for any prediction it is important that you test your parameter set for an independent time period. Use the period 870831 to 871131 for this test. Is this fit worse? Can you give an explanation? How will your design flood be affected?

3) Simulation of flood
   - Make a backup copy of the file.
   - Open the file plotted in a text editor (or Excel).
   - Choose a period for which you replace the observed precipitation by the synthetic precipitation. Use the period 870901 to 870831 for calibration (with the warming-up period starting at 810901).
   - Save the file (if you use Excel choose the format '*.csv' (colon-separated), but use the name 'ptq.dat').
   - Restart MATLAB and run the model. Check the peak value of your simulated flood.
   - Return to the backup file, choose a different period and continue with 3. Do this 5-10 times.

What influences the size of the simulated flood?

Under which conditions becomes the simulated flood largest/smallest?

Table 1

<table>
<thead>
<tr>
<th>Day</th>
<th>P [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>

**Example:**

Data from the Kassjön basin in Medelpad, Sweden is used in this exercise. In the file ex5_snow7376.dat you find precipitation [mm], temperature [°C] and depth of the snow pack [mm water equivalent].

A synthetic sequence of extreme precipitation has been derived by meteorologists. In other words, you should estimate a new synthetic sequence of extreme precipitation that is likely to cause a flood that is similar to the flood that occurred in 1975 to 1985. The model is then calibrated. However, the calibration can be quite sensitive to changes in the parameter set. The model is far from well-calibrated.

You have to complete three steps:

1) Calibration
   - Choose a period for which you replace the observed precipitation by the synthetic precipitation. Use the period 870901 to 870831 for calibration (with the warming-up period starting at 810901).

2) Validation
   - Before you can use your calibrated model for any prediction it is important that you test your parameter set for an independent time period. Use the period 870831 to 871131 for this test. Is this fit worse? Can you give an explanation? How will your design flood be affected?

3) Simulation of flood
   - Make a backup copy of the file.
   - Open the file plotted in a text editor (or Excel).
   - Choose a period for which you replace the observed precipitation by the synthetic precipitation. Use the period 870901 to 870831 for calibration (with the warming-up period starting at 810901).
   - Save the file (if you use Excel choose the format '*.csv' (colon-separated), but use the name 'ptq.dat').
   - Restart MATLAB and run the model. Check the peak value of your simulated flood.
   - Return to the backup file, choose a different period and continue with 3. Do this 5-10 times.

What influences the size of the simulated flood?

Under which conditions becomes the simulated flood largest/smallest?

Table 1

<table>
<thead>
<tr>
<th>Day</th>
<th>P [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>