Short to medium range probabilistic streamflow forecasting for Tel river using the Hydrologic Ensemble Forecast Service

M Sarath, V. V. Srinivas

Department of Civil Engineering, and Interdisciplinary Center for Water Research (ICWaR)
Indian Institute of Science, Bangalore, India
# Streamflow forecasting - Applications

<table>
<thead>
<tr>
<th>Lead time of forecast</th>
<th>Applications</th>
</tr>
</thead>
</table>
| Short range flood forecasts (up to 3 days) | • Flood warning  
• Flood control system management for saving lives and some assets |
| Medium range flood forecasts (3–15 days) | • Flood control system management for preserving additional livelihood assets |
| Long-range flood forecasts (e.g., monthly, seasonal flow outlooks) | • Planning and management of water resources (e.g., for agriculture, hydropower, Industry) |

Kerala floods 2018

Overtopping of dam in Kerala
Streamflow Forecasting Approaches

**Process-driven**

Meteorological forecast

Terrestrial Hydrologic Model

Ground-Water Hydrologic Model

Evapotranspiration

Hydrologic models are forced with meteorological forecasts from NWP models, considering basin initial conditions

(Source: essc.psu.edu)

**Data-driven/statistical**

Predictors

ARIMA

ANN

ANFIS

k-NN

Streamflow
Forecast uncertainty

- Causes
  - Complex nature of system
  - Imperfect models
  - Lack of data
  - Measurement errors, etc.

- Quantification of uncertainty
  - Ensemble (instead of single valued forecast)
Forecasting Flow in a River using Hydrologic Ensemble Prediction Systems (HEPS)

- Generate ensemble forecasts of streamflow to quantify forecast uncertainty
- Increasingly used by flood forecasting centres
### Forecasting Flow in a River using Hydrologic Ensemble Prediction Systems (HEPS)

<table>
<thead>
<tr>
<th>HEPS name</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Flood Awareness System (GloFAS)</td>
<td>Global</td>
</tr>
<tr>
<td>European Flood Awareness System (EFAS)</td>
<td>Europe</td>
</tr>
<tr>
<td>Hydrologic Ensemble Forecast Service (HEFS)</td>
<td>United States</td>
</tr>
<tr>
<td>Climate Forecast Applications in Bangladesh (CFAB)</td>
<td>Bangladesh</td>
</tr>
<tr>
<td>Joint Flood Forecasting System</td>
<td>UK</td>
</tr>
<tr>
<td>French Hydrometeorological Ensemble Prediction System</td>
<td>France</td>
</tr>
<tr>
<td>Watershed simulation and forecasting system (WSFS)</td>
<td>Finland</td>
</tr>
</tbody>
</table>

Operational/pre-operational HEPS source: HEPEX

Can the **HEPS** provide skilful streamflow forecasts for Indian rivers? Do we need India specific ensemble forecasting system?
Forecasting Flow in a River using Hydrologic Ensemble Prediction Systems (HEPS)

Studies in India

Map 1.1: The Ganges-Brahmaputra-Meghna River Basins

Source: World Bank, 2017

Foreseeing River Flow using Hydrologic Ensemble Prediction Systems (HEPS)

<table>
<thead>
<tr>
<th>HEPS Name</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Flood Awareness System (GloFAS)</td>
<td>Global</td>
</tr>
<tr>
<td>European Flood Awareness System (EFAS)</td>
<td>Europe</td>
</tr>
<tr>
<td>Hydrologic Ensemble Forecast Service (HEFS)</td>
<td>United States</td>
</tr>
<tr>
<td>Climate Forecast Applications in Bangladesh (CFAB)</td>
<td>Bangladesh</td>
</tr>
<tr>
<td>Joint Flood Forecasting System</td>
<td>UK</td>
</tr>
<tr>
<td>French Hydro-meteorological Ensemble Prediction System</td>
<td>France</td>
</tr>
<tr>
<td>Watershed simulation and forecasting system (WES)</td>
<td>Finland</td>
</tr>
</tbody>
</table>

Operational/pre-operational HEPS

source: HEPEX

HEFS (Hydrologic Ensemble Forecast Service)

Figure: Schematic diagram of HEFS

I: MEFP
- Meteorological forecasts
  - Precipitation, temperature

II: Hydrologic Processor
- Simulate streamflow, propagate uncertainty
- Raw ensemble streamflow forecast
- Ensembles meteorological forecasts
- Model meteorological uncertainty

III: EnsPost
- Postprocessed ensemble streamflow forecast
- Model hydrologic uncertainty
- Meteorological forecasts (precipitation, temperature)
I: MEFP (Meteorological Ensemble Forecast Processor)

- Models meteorological uncertainty
- Considers raw forecasts from multiple sources
- Corrects bias in raw forecasts
  - Preserves skill over multiple time scales (Canonical events).
  - Preserves space-time correlation structure (Shaake Shuffle).

Single-valued raw forecasts (ensemble mean)

- Short-range
  - RFC

- Medium-range
  - GEFSv2

- Long-range
  - CFSv2

- Climatology

Bias corrected seamless/merged ensemble forecast

**Figure:** Modified from Wu (2014)

RFC: River forecast centre
CFS: Climate Forecast System
GEFS: Global Ensemble Forecast System
I: MEFP (Meteorological Ensemble Forecast Processor)

Canonical events

- Time windows over which forecasts and observations are aggregated
- Used to capture skill of raw forecasts at multiple time scales

<table>
<thead>
<tr>
<th>Lead time (days)</th>
<th>Temp. base events</th>
<th>Precipitation base events</th>
<th>Modulation events</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>
I: MEFP (Meteorological Ensemble Forecast Processor)

Figure: Regression-based statistical processing method used in MEFP (Li et al., 2017)
MEFP - statistical model

- Use historical data for bias correction and uncertainty estimation
- Model joint distribution of forecasts \((X)\) and observations \((Y)\)
  \[ F(x,y) = P(X \leq x, Y \leq y) \]
- Assumed to be bivariate normal distribution.
- Given raw forecast \(x\), obtain conditional distribution of \(Y\)
  \[ F_{Y|X}(y|x) = P(Y \leq y | X = x) \]

\[
\mu_{Y|X} = \mu_Y + \frac{\sigma_y}{\sigma_x} \rho_{xy} (x - \mu_x)
\]

\[
\sigma_{Y|X}^2 = \sigma_y^2 \left(1 - \rho_{xy}^2\right)
\]
HEFS (Hydrologic Ensemble Forecast Service)

**Figure:** Schematic diagram of HEFS

I: MEFP
- Meteorological forecasts (precipitation, temperature)
- Model meteorological uncertainty
- Raw ensemble meteorological forecasts

II: Hydrologic Processor
- Simulate streamflow forecast
- Propagate uncertainty
- Postprocessed ensemble streamflow forecast

III: EnsPost
- Model hydrologic uncertainty
- Meteorological forecasts
- Predicted streamflow forecast

MEFP: Meteorological Ensemble Forecast Processor
II: Hydrologic processor

- Propagates meteorological uncertainty
  - Model initialised with basin conditions
  - Forced with forecasts from MEFP

- GR4J hydrologic model
  - Lumped daily rainfall-runoff model
  - A hybrid metric-conceptual model (Wheater et al., 1993; Young, 2001)

Four parameters

**Raw ensemble streamflow forecast (HEFS-RAW)**
### PET (Oudin et al., 2005)

\[
PET = \begin{cases} 
\frac{R_e \cdot T_a + 5}{100} & \text{if } T_a + 5 > 0 \\
0 & \text{otherwise}
\end{cases}
\] 

(\text{mm/day})

- \(R_e\): Extra-terrestrial radiation (MJ m\(^{-2}\) day\(^{-1}\))
- \(\lambda\): Latent heat flux (MJ kg\(^{-1}\))
- \(\rho\): Density of water (kg m\(^{-3}\))
- \(T_a\): Mean daily air temperature (°C)

If \(P \geq E\), then \(P_n = P - E\) and \(E_n = 0\)

otherwise, \(P_n = 0\) and \(E_n = E - P\)

- \(P_n\): Net excess precipitation
- \(E_n\): Net deficit in meeting PET demand

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Controls</th>
<th>Units</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>Maximum capacity of the Soil moisture accounting store</td>
<td>Higher values increase soil moisture, giving hydrograph longer memory</td>
<td>mm</td>
<td>&gt; 0</td>
</tr>
<tr>
<td>x2</td>
<td>Groundwater exchange coefficient</td>
<td>Higher values increase streamflow, violating water balance</td>
<td>mm</td>
<td>real no.</td>
</tr>
<tr>
<td>x3</td>
<td>Reference capacity of the routing store</td>
<td>Controls the discharge of baseflow</td>
<td>mm</td>
<td>&gt; 0</td>
</tr>
<tr>
<td>x4</td>
<td>Time parameter for unit hydrographs</td>
<td>Higher values delay and attenuate the hydrograph</td>
<td>days</td>
<td>&gt; 0.5</td>
</tr>
</tbody>
</table>
HEFS (Hydrologic Ensemble Forecast Service)

**Figure: Schematic diagram of HEFS**

1. **MEFP**: Meteorological Ensemble Forecast Processor
   - Raw ensemble streamflow forecast
   - Simulate streamflow, propagate uncertainty

2. **EnsPost**: Model hydrologic uncertainty
   - Postprocessed ensemble streamflow forecast

3. **II**: Hydrologic Processor
   - Model meteorological uncertainty
   - Meteorological forecasts (precipitation, temperature)
III: EnsPost (Hydrologic Ensemble Postprocessor)

- Account for hydrologic uncertainty
- Correct bias in streamflow simulation
- Account for autocorrelation

```
HEFS-RAW → EnsPost → HEFS-POST
```

- Raw ensemble streamflow forecast
- Postprocessed ensemble streamflow forecast

\[ \text{Simulated flow (mm/day)} \]

\[ \text{Observed flow (mm/day)} \]

Hydrologic error

Graph showing scatter plot of simulated vs. observed flow, with a line indicating expected perfect correlation.
**EnsPost - statistical model**

- Output of GR4J
  
  \( Y: \text{Observations} \)

- NQT

  \[
  Z = \Phi^{-1}(F_X(X)) \\
  W = \Phi^{-1}(F_Y(Y))
  \]

- First order autoregressive model with single exogeneous input
  
  \( \text{ARX}(1,1) \)

\[
W_{t+1} = (1 - b) W_t + b Z_{t+1} + E_{t+1}
\]

\[
E_{t+1} = \rho E_t + \varepsilon_{t+1}
\]
Case study using HEFS - Tel river in Mahanadi basin

- Daily streamflow forecasts are generated at Kantamal gauge (CA: 20,235 km²)
- Lead times: 1-day to 1-month

HEFS: Hydrologic Ensemble Forecast Service
Data

- Daily Precipitation (IMD): 1901-2013 (0.25° resolution)
- Daily Temperature (IMD): 1951-2013 (1° resolution)
- Daily forecasts from Global Ensemble Forecast System (GEFSv2) for 1 to 16 day lead times
  - Precipitation (1° resolution)
  - Temperature (1° resolution)
  - Period of record: 1985-2013
- Daily streamflow at Kantanāl gauge (India-WRIS)
  - Period of record: 1972-2011
HEFS (Hydrologic Ensemble Forecast Service)

**I: MEFP**
- Meteorological forecasts (precipitation, temperature)

**II: Hydrologic Processor**
- Simulate streamflow
- Propagate uncertainty

**III: EnsPost**
- Postprocessed ensemble streamflow forecast

**Ensemble meteorological forecasts**

**Raw ensemble streamflow forecast**
Results - Temperature from GEFS after bias correction by MEFP

MEFP: Meteorological Ensemble Forecast Processor
GEFS: Global Ensemble Forecast System

GEFS temperature forecast has skill over the study area.
Results - Precipitation from GEFS after bias correction by MEFP

GEFS precipitation forecast has low skill over the study area

MEFP: Meteorological Ensemble Forecast Processor
GEFS: Global Ensemble Forecast System
Validation/verification of forecasts

- **Split sample approach**
  - Parameter estimation – data up to 1999
  - Forecast verification – 2000-2011 (wet season, June-Oct)

- **Verification metrics**
  - Mean Error (ME)
  - Pearson correlation coefficient
  - CRPSS (Continuous Ranked Probability Skill Score)

\[
\text{CRPS} = \int \{ F_Y(q) - F_X(q) \}^2 dq
\]

\[
\text{CRPSS} = 1 - \frac{\text{CRPS}_{\text{main}}}{\text{CRPS}_{\text{ref}}}
\]

Ref: Climatological forecast

Main: HEFS/ESP/ARIMA

CRPS (www.met-learning.eu)
Streamflow simulation (GR4J model) using observed rainfall

<table>
<thead>
<tr>
<th>Measure</th>
<th>Calibration</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>0.90</td>
<td>0.92</td>
</tr>
<tr>
<td>RMSE (mm/day)</td>
<td>1.76</td>
<td>2.22</td>
</tr>
<tr>
<td>NSE</td>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td>$\text{KGE}_{\text{mod}}$</td>
<td>0.90</td>
<td>0.79</td>
</tr>
</tbody>
</table>

$x_1 = 104.5 \text{ mm} \quad x_2 = 1.2 \text{ mm} \quad x_3 = 186.7 \text{ mm} \quad x_4 = 1.4 \text{ days}$

KGE$_{mod}$: Modified Kling-Gupta efficiency criterion.
Results - Streamflow forecast

MODIS images (NASA)

Streamflow forecast generated using HEFS on September 15, 2008

HEFS: Hydrologic Ensemble Forecast Service
Results - Streamflow forecasts

lead time: 1
lead time: 2
lead time: 3
lead time: 5
lead time: 7
lead time: 16

Forecast flow (mm/d)

Observed flow (mm/d)
Results - Streamflow forecasts

CRPSS: (Continuous Ranked Probability Skill Score)

HEFS: Hydrologic Ensemble Forecast Service

CLIM: climatology/climatological
HEFS (Hydrologic Ensemble Forecast Service)

- **I: MEFP**
  - Meteorological forecasts (precipitation, temperature)

- **II: Hydrologic Processor**
  - Raw ensemble streamflow forecast
  - Simulate streamflow, propagate uncertainty

- **III: EnsPost**
  - Postprocessed ensemble streamflow forecast

Schematic diagram of HEFS
## Comparison of Ensemble Streamflow Forecasts

<table>
<thead>
<tr>
<th>Model</th>
<th>Initial Conditions</th>
<th>Forecast Type</th>
<th>Post-Processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESP: GR4J</td>
<td></td>
<td>Ensemble streamflow</td>
<td></td>
</tr>
<tr>
<td>CLIM</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>ARIMA</td>
<td>✓</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>ESP</td>
<td>✓</td>
<td>CLIM</td>
<td>x</td>
</tr>
<tr>
<td>HEFS-RAW</td>
<td>✓</td>
<td>MEFP</td>
<td>x</td>
</tr>
<tr>
<td>HEFS-POST</td>
<td>✓</td>
<td>MEFP</td>
<td>✓</td>
</tr>
</tbody>
</table>

### Notes:
- **CLIM**: Climatological
- **ARIMA**: Autoregressive Integrated Moving Average
- **ESP**: Ensemble Streamflow Prediction
- **HEFS-RAW**: Hybrid Ensemble Forecast System - Raw
- **HEFS-POST**: Hybrid Ensemble Forecast System - Post-processed

### Diagram:
- **Basin Initial Conditions**
- **ESP**
- **Past <-> Future Time**
- **Ensemble streamflow prediction**
  (source: Stokes, 2014)

**MEFP**: Meteorological Ensemble Forecast Processor
Comparison of Ensemble Streamflow Forecasts

HEFS: Hydrologic Ensemble Forecast Service
ESP: Ensemble streamflow prediction
MEFP: Meteorological Ensemble Forecast Processor
GEFS: Global Ensemble Forecast System

CRPSS (Continuous Ranked Probability Skill Score)

Comparison of Ensemble Streamflow Forecasts

Daily time scale

Aggregated time scale

CRPSS (Continuous Ranked Probability Skill Score)

Lead time

HEFS-POST
HEFS-RAW
Precipitation forecast (MEFP-GEFS)
ESP
ARIMA

Hydrologic postprocessing
Basin initial conditions
## Conclusions & Future scope

**Conclusions**

- HEFS forecasts have higher skill than ARIMA forecasts.
- At the daily time scale, most of the skill of HEFS at short lead times is due to the contribution from basin initial conditions.

**Future scope**

- Alternate hydrologic, hydraulic models
- Rainfall forecasts from other sources (IMD, WRF)
- Data assimilation
- Other basins
KGE\textsubscript{mod} : Modified Kling-Gupta efficiency criterion

\[ KGE_{mod} = 1 - \sqrt{(r - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \]

\[ \beta = \frac{\mu_s}{\mu_0} \quad \text{and} \quad \gamma = \frac{CV_s}{CV_o} = \frac{\sigma_s/\mu_s}{\sigma_o/\mu_o} \]

\( r \) : Pearson product moment correlation coefficient between the simulated and observed streamflows
\( \beta \) : Bias ratio
\( \gamma \) : Variability ratio
\( \mu_0 \) and \( \mu_s \) represent the mean, \( CV_o \) and \( CV_s \) the coefficient of variation, \( \sigma_o \) and \( \sigma_s \) the standard deviation corresponding to the observed and simulated streamflows respectively.
FIGURE 1.6. Forecasting Tools and Methods for Different Types of Floods

Sources: Modified from Hopson et al. 2015b; axis annotations adopted from ESCAP 2016b.

Note: Various forecasting tools provide information at different scales, thereby providing different lead-time contributions. The NCAR model, which focuses on large river basin flooding, relies on weather forecasts (blue), simulation models (orange), and measurements (black). The specific NCAR tools are highlighted with asterisks. GLOFs = glacial lake outburst floods.
Normal quantile transform (NQT)

- Forecasts/simulations ($X$) and observations ($Y$)
- Skewness – convert to normal space
- fit empirical CDF to $X$ and $Y$

$$Z = \Phi^{-1}(F_X(X)) \quad W = \Phi^{-1}(F_Y(Y))$$

(Brown, 2014)
NCAR flood forecasting

Weather Forecasts (from 8 centers)
- Observed Rainfall (gauges, satellite)
- Observed River flow (stream gauges, satellites)

Hydrologic model (lumped catchment)
Hydrologic model (upstream stage-discharge routing)

“Multi-model” (hydrologic forecast combination)

Final water-level forecast (probability-based)
Forecast dissemination (online viewing, warnings, maps)

(I) Initial Data Input
- ECMWF Forecasts (F)
  - Daily
  - 3-member
- Precipitation Estimates (S)
  - NOAA & NASA satellite
  - NOAA interpolated gauge
- River Discharge (Q)
  - Daily
  - Distribute hydro model

(II) Statistical Rendering
- Downscale forecasts
- Statistical correction

(III) Hydrologic Modeling
- Semi-distributed Model (SDM)
  - Hindcast/Forecast Generation
  - Update soil moisture states
  - Update in-stream flows

Multi-Model (M) Hindcast/Forecast Generation
- Calibrate multi-model
  - Generate forecasts
  - Generate hindcasts

(IV) Generation of Probabilistic Discharge Forecasts
- Accounting for Uncertainties and Final Error Correction
  - Generate model error PDF
  - Convolute multi-model ensemble forecasts & model error PDF
  - Calculate above-critical-level probabilities

(V) Dissemination
- Above-critical-level forecast probabilities transferred to Bangladesh
ARIMA forecast
ARIMA forecast