PS2: Hydroclimate Risk Assessment for the Colorado River Basin (Due Nov 17)

End-to-end modeling: snow–soil–runoff, Bayesian calibration, and climate downscaling

CEVE 543 Fall 2025

2025-10-06

This problem set is the capstone assignment for Module 2 and spans approximately six weeks of class time. It requires you to synthesize concepts from the entire course—from Bayesian inference and numerical methods to climate model analysis. The deliverable is a complete, end-to-end analysis of future drought risk for the Colorado River Basin, written as a technical report.

> **i Note**
>
> Assigned: Oct 6, 2025
> Due: Nov 17, 2025

1 Provided

- Historical data: `colorado_river_data.csv` — daily, basin-averaged (1950–2022)
  - ‣ `date`: Date of record
  - ‣ `prcp`: Precipitation (mm/day)
  - ‣ `tavg`: Average temperature (°C)
  - ‣ `streamflow_obs`: Observed streamflow at Lees Ferry (mm/day, normalized by basin area)
- GCM data: `cesm_le_rcp85.csv` — raw daily GCM output (2050–2080) from one CESM-LE member under RCP 8.5 (`date`, `prcp`, `tavg`)
- Boilerplate Quarto template: `PS2-template.qmd` — suggested report structure, data loading helpers, plotting examples

> **! Important**
>
> Data availability
>
> Links/paths will be posted on Canvas or added to this repo when finalized. If these files are not yet present in your local checkout, proceed with stubs and clearly document assumptions. Replace with provided files once available.

1.1 Julia setup

You will likely need the following packages (non-exhaustive): `CSV`, `DataFrames`, `Dates`, `StatsBase`, `Distributions`, `Turing`, `MCMCChains`, `Random`, `Plots` or `CairoMakie`.

1. Open the Julia REPL
2. Type `]` to enter package mode
3. Install packages, e.g., `add CSV DataFrames Distributions Turing MCMCChains`

2 Overview: The modeling chain

You will act as a hydroclimate risk analyst tasked with projecting future water supply in the Colorado River Basin. Implement the complete modeling chain:

1. Build a model: Implement a simple, process-based hydrological model from governing equations.
2. Calibrate the model: Use Bayesian MCMC to calibrate against the historical record.
3. Downscale climate projections: Process raw GCM output and develop multiple future climate scenarios using different downscaling techniques.
4. Project future risk: Force your calibrated model with future scenarios and synthesize results, including propagation of parameter uncertainty.

3 Part 1: The hydrological model — theory and implementation

3.1 Task 1.1: Model theory

Implement a lumped conceptual snowmelt–runoff model designed to capture first-order hydroclimate processes in the Colorado River Basin: seasonal accumulation of mountain snowpack and spring melt.

The model has three conceptual modules:

1. Snow module (temperature-index/degree-day)

- Accumulation: when basin-mean temperature is below a threshold T_{thresh}, all precipitation is snow and accumulates in snow water equivalent (SWE).
- Melt: when temperature is above the threshold, existing snow melts at rate $M = C_m \max(0, T - T_{\text{thresh}})$, bounded by available SWE. Melt plus rain forms liquid water input to the soil module.

2. Soil moisture module ("leaky bucket")

- Storage: incoming liquid water fills a conceptual bucket up to capacity S_{max}.
- Evapotranspiration: actual ET is a fraction of potential ET (estimated from temperature), scaled by relative soil water: $E_a = E_p (S/S_{\text{max}})$.
- Runoff generation: overflow beyond S_{max} becomes excess runoff.

3. Runoff routing module (linear reservoir)

- The routing store G represents aggregate river network and shallow groundwater effects; outflow is proportional to storage: $Q = k G$.

State variables:

- SWE (t): snow water equivalent (mm)
- $S(t)$: soil moisture content (mm)
- $G(t)$: routing storage (mm)

Governing ODEs:

$$\frac{d\,\mathrm{SWE}}{dt} = P_s(t) - M(t)$$

$$\frac{dS}{dt} = W_{\mathrm{in}}(t) - E_a(t) - R(t)$$

$$\frac{dG}{dt} = R(t) - Q(t), \quad Q(t) = k\,G(t)$$

Fluxes like melt $M(t)$ and runoff $R(t)$ are functions of states and meteorological inputs.

3.2 Task 1.2: Numerical discretization

Use forward Euler with daily time step ($\Delta t = 1$ day). For day i:

1. Snow module

- $M_i = \min\left(\,\mathrm{SWE}_{i-1}, C_m \max(0, T_i - T_{\mathrm{thresh}})\,\right)$
- $\mathrm{SWE}_i = \mathrm{SWE}_{i-1} + P_{\mathrm{snow},i} - M_i$
- $W_{\mathrm{in},i} = P_{\mathrm{rain},i} + M_i$

2. Soil module

- $E_{a,i} = E_{p,i}\,(S_{i-1}/S_{\mathrm{max}})$ (with $E_{p,i}$ from, e.g., Hamon)
- $S'_i = S_{i-1} + W_{\mathrm{in},i} - E_{a,i}$
- $R_i = \max(0, S'_i - S_{\mathrm{max}})$
- $S_i = S'_i - R_i$

3. Routing module

- $Q_i = k\,(G_{i-1} + R_i)$
- $G_i = (G_{i-1} + R_i) - Q_i$

3.3 Task 1.3: Implementation

Implement a Julia function with signature `run_hydromodel(params, climate_data)` that returns a vector of simulated daily streamflow.

Parameters to calibrate (params):

- T_{thresh} (°C): snow/rain temperature threshold
- C_m (mm/°C/day): degree-day melt factor
- S_{max} (mm): maximum soil water capacity
- k (1/day): linear reservoir recession coefficient

> **💡 Tip**
>
> Function contract
>
> Inputs
>
> - `params::NamedTuple` or `Vector{<:Real}` with fields/positions (T_thresh, C_m, S_max, k)
> - `climate_data::DataFrame` with columns `:date`, `:prcp`, `:tavg` (and optionally `:streamflow_obs` during calibration)
>
> Output
>
> - `Vector{Float64}` of daily streamflow Q_i (mm/day) aligned to input dates

4 Part 2: Bayesian model calibration

4.1 Task 2.1: MCMC setup

Using `Turing.jl`, build a Bayesian model to estimate the four physical parameters plus an observation error scale σ.

- Priors: choose physically plausible distributions (e.g., T_{thresh} centered near 0°C, $S_{\text{max}} > 0$, $k \in (0, 1)$). Justify choices.
- Likelihood: assume IID Gaussian residuals for streamflow with unknown σ.

4.2 Task 2.2: Calibration and assessment

- Run MCMC to obtain at least 2,000 posterior samples after warmup; check convergence diagnostics.
- Deliverables:
  1. Plot of prior vs. posterior for the four physical parameters.
  2. Hydrograph overlay: observed streamflow and simulated streamflow using the MAP (or posterior mean) parameters.
  3. A 1–2 paragraph interpretation of posteriors (e.g., what S_{max} implies about storage) and a critical assessment of hydrograph fit.

5 Part 3: Future climate scenarios

Create three future climate scenarios for 2050–2080.

5.1 Task 3.1: Raw GCM forcing

- Analyze raw GCM vs. historical: compare distributions (e.g., CDFs or violin plots) for daily precipitation and temperature.
- Force the calibrated hydrologic model with raw GCM data; comment on biases in resulting streamflow.

5.2 Task 3.2: Univariate bias correction

- Implement simple quantile mapping to bias-correct daily precipitation and temperature independently against historical observations.
- Force the hydrologic model with the bias-corrected series; compare with raw-forced results.

5.3 Task 3.3: Statistical downscaling with a weather generator

- Develop a 2-state (wet/dry) Hidden Markov Model (HMM) weather generator.

- Train HMM parameters (transition probabilities; emission distributions for precipitation amount and temperature) on historical data.
- Adjust trained HMM parameters using GCM-projected delta changes in mean precipitation and temperature.
- Generate a 30-year synthetic weather sequence and force the hydrologic model.

6 Part 4: Synthesis and uncertainty quantification

6.1 Task 4.1: Scenario comparison
- Create a summary plot showing mean monthly hydrographs for the historical period and the three future scenarios (Raw GCM, Bias-Corrected, HMM).

6.2 Task 4.2: Propagating parameter uncertainty
- For the HMM scenario, quantify parameter uncertainty by running a 100-member ensemble:
  1. Draw a unique parameter set from the MCMC posterior for each member.
  2. Run the hydrologic model with this parameter set using the HMM-generated weather.

6.3 Task 4.3: Final report
Deliverables:

1. The mean monthly hydrograph comparison plot from Task 4.1.
2. A synthesis plot for the HMM scenario showing the mean monthly hydrograph with a 90% uncertainty band from the 100-member ensemble.
3. An executive summary (2–3 paragraphs) for a water manager: explain the cascade of uncertainty and argue why the HMM-forced, uncertainty-quantified projection is most credible and useful for long-term planning, citing your results.

7 Notes and guidance
- Units: keep all fluxes in mm/day; normalize streamflow by basin area to compare with observations.
- Potential ET: you may use Hamon's method or a comparable temperature-based proxy; justify your choice.
- Numerical stability: enforce non-negativity of states; cap melt by available SWE.
- Reproducibility: fix RNG seeds when appropriate and document all assumptions and choices.

Bibliography