EGU26 GENERAL ASSEMBLY

Sharing is encouraged

Communicating hydrological model calibration with toy examples

Georgia Papacharalampous, Francesco Marra, Eleonora Dallan, and Marco Borga

georgia.papacharalampous@unipd.it, papacharalampous.georgia@gmail.com

TESAF

Acknowledgements: The present study was developed in the framework of the ATLAS project, with financial support provided by the Autonomous Province of Bolzano/Bozen, Italy.

1. Loss and identification functions

Foundations and domain

- z: prediction
- y: observation
- $\boldsymbol{z}_n = (z_1, \dots, z_n)^{\mathrm{T}}$: vector of predictions in train or test set
- $\boldsymbol{y}_n = (y_1, \dots, y_n)^{\mathrm{T}}$: vector of observations in train or test set
- $L(z, y)$: loss function when we predict z and y realizes
- $\bar{L}(\boldsymbol{z}_n, \boldsymbol{y}_n) = (1/n)\sum_{i=1}^{n} L(z_i, y_i)$: average loss in train or test set when we predict $\boldsymbol{z}_n$ and $\boldsymbol{y}_n$ realizes
- $V(z, y)$: identification function when we predict z and y realizes
- $\bar{V}(\boldsymbol{z}_n, \boldsymbol{y}_n) = (1/n)\sum_{i=1}^{n} V(z_i, y_i)$: empirical identification function in train or test set when we predict $\boldsymbol{z}_n$ and $\boldsymbol{y}_n$ realizes
- With monotonicity and convexity hereinafter referring to the domain $z, y > 0$

Predictive mean

- Bregman loss function:

$$L_{\mathrm{Br}}(z, y; \varphi) = \varphi(y) - \varphi(z) - \varphi'(z)(y - z), \text{ with:}$$

  φ convex implying L consistent and φ strictly convex implying L strictly consistent
- For $\varphi(t) = |t|^b, t > 0$, we have the following family of loss functions:

$$L_{\mathrm{Br}}(z, y; b) = (1/b(b-1))\left(y^b - z^b\right) - (1/(b-1))z^{b-1}(y - z), b \in \mathbb{R}\backslash\{0,1\}$$

- Special cases arise for specific values of b:

$$L_{\mathrm{Br}}(z, y) = \begin{cases} L_{\text{Bregman 1 (QLIKE)}}(z, y) = (y/z) - \log(y/z) - 1, b = 0 \\ L_{\text{Bregman 2 (Patton)}}(z, y) = y\log(y/z) - y + z, b = 1 \\ L_{\text{Bregman 3 (square)}}(z, y) = (1/2)(z - y)^2, b = 2 \\ L_{\text{Bregman 4}}(z, y) = (1/12)(y^4 - z^4) - (1/3)z^3(y - z), b = 4 \end{cases}$$

- A special case of the Bregman loss function for evaluating predictions of extremes:

$$L_{\text{Bregman 5 (Taggart) }[\alpha \in \mathbb{R}]}(z, y) = (y - a)^2 \mathbb{1}\{y \geq a\} + [(y - z)^2 - (y - a)^2]\mathbb{1}\{z \geq a\}$$

- Mean identification function:

$$V_{\mathrm{mean}}(z, y) = z - y$$

Predictive median

- Generalized piecewise linear (GPL) loss function for the median:

$$1L_{\mathrm{GPL}\,[\tau=0.5]}(z, y; g) = |g(z) - g(y)|, \text{ with:}$$

  g non-decreasing implying L consistent and g strictly increasing implying L strictly consistent
- Special cases arise for specific functions g:

$$L_{\mathrm{GPL\,1}\,[\tau\,=\,0.5]}(z, y; g = \log(t)) = |\log(z) - \log(y)|$$

$$L_{\mathrm{GPL\,2}\,[\tau\,=\,0.5]\,\mathrm{(absolute)}}(z, y; g = t) = |z - y|$$

$$L_{\mathrm{GPL\,3}\,[\tau\,=\,0.5]}(z, y;\ g = t^2) = |z^2 - y^2|$$

$$L_{\mathrm{GPL\,4}\,[\tau\,=\,0.5]}(z, y;\ g = t^3) = |z^3 - y^3|$$

- A special case of the GPL loss function for the median for evaluating predictions of extremes:

$$L_{\mathrm{GPL\,5\,(Taggart)}\,[\tau\,=\,0.5,\,a\,\in\,\mathbb{R}]}(z, y) = |(z - a)\mathbb{1}\{z \geq a\} - (y - a)\{y \geq a\}|$$

- Quantile identification function for the median:

$$V_{\mathrm{median}}(z, y) = \mathbb{1}\{z \geq y\} - 0.5$$

Predictive τ-quantile

- GPL loss function:

$$L_{\mathrm{GPL}}(z, y; g, \tau) = (\mathbb{1}\{z \geq y\} - \tau)(g(z) - g(y)), \text{ with:}$$

  g non-decreasing implying L consistent and g strictly increasing implying L strictly consistent
- Special cases arise for specific functions g:

$$L_{\mathrm{GPL\,1}\,[\tau]}\big(z, y; g = \log(t)\big) = (\mathbb{1}\{z \geq y\} - \tau)(\log(z) - \log(y))$$

$$L_{\mathrm{GPL\,2\,(pinball)}\,[\tau]}(z, y; g = t) = (\mathbb{1}\{z \geq y\} - \tau)(z - y)$$

$$L_{\mathrm{GPL\,3}\,[\tau]}(z, y;\ g = t^2) = (\mathbb{1}\{z \geq y\} - \tau)(z^2 - y^2)$$

$$L_{\mathrm{GPL\,4}\,[\tau]}(z, y;\ g = t^3) = (\mathbb{1}\{z \geq y\} - \tau)(z^3 - y^3)$$

- A special case of the GPL loss function for evaluating predictions of extremes:

$$L_{\mathrm{GPL\,5\,(Taggart)}\,[\tau,\,\alpha\,\in\,\mathbb{R}]}(z, y) = (\mathbb{1}\{z \geq y\} - \tau)((z - a)\mathbb{1}\{z \geq a\} - (y - a)\{y \geq a\})$$

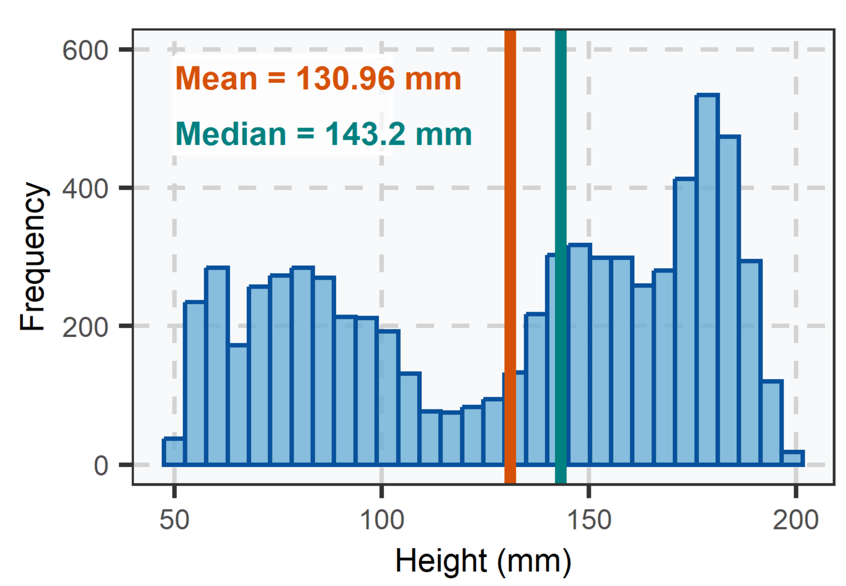
- Quantile identification function:

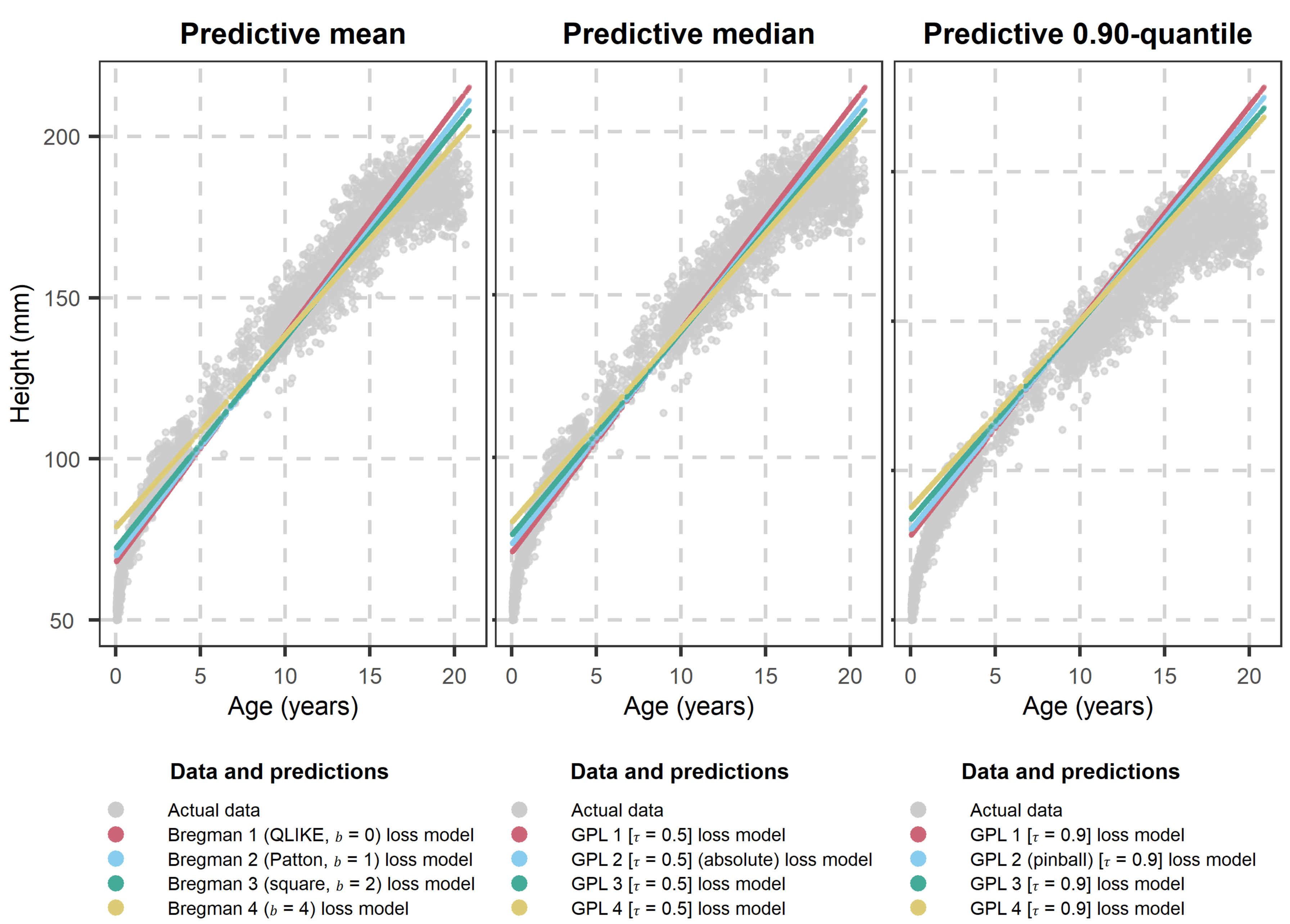
$$V_{\mathrm{quantile}}(z, y; \tau) = \mathbb{1}\{z \geq y\} - \tau$$

2. Toy modelling

Experimental design

- **Model:** Linear with one predictor
- **Dataset:** `dutchboys`
- **Predictand:** Height
- **Predictor:** Age
- **Target functionals:** Mean, median, 0.90-quantile
- **Calibration:** Using all non-extreme loss functions
- **Evaluation:** Using all loss and identification functions
- **Extremes:** Defined through specific thresholds for the Taggart loss functions (a = 180, 185, 190, 195 mm)





3. Predictive performance

Rank (1 = best): 1, 2, 3, 4, NA

Predictive mean — Loss or identification function (rows) × Model (columns)

Loss or identification function	Train: Bregman 1 (QLIKE, b = 0) loss model	Train: Bregman 2 (Patton, b = 1) loss model	Train: Bregman 3 (square, b = 2) loss model	Train: Bregman 4 (b = 4) loss model	Test: Bregman 1 (QLIKE, b = 0) loss model	Test: Bregman 2 (Patton, b = 1) loss model	Test: Bregman 3 (square, b = 2) loss model	Test: Bregman 4 (b = 4) loss model
Bregman 1 (QLIKE, b = 0) loss	0	0	0	0.01	0	0	0	0.01
Bregman 2 (Patton, b = 1) loss	0.41	0.4	0.41	0.55	0.43	0.41	0.43	0.58
Bregman 3 (square, b = 2) loss	110.88	100.28	97.24	112.79	114.68	103.18	100.02	117.18
Bregman 4 (b = 4) loss	1468793.86	1200074.7	1060980.68	956663.5	1530566.69	1236232.79	1083634.4	971432.27
Bregman 5 (Taggart) [a = 180] loss	51.51	36.55	27.82	18.02	57.11	40.38	30.53	19.64
Bregman 5 (Taggart) [a = 185] loss	45.81	31.09	22.49	12.86	50.68	34.28	24.77	14.19
Bregman 5 (Taggart) [a = 190] loss	37.52	23.99	16.25	7.86	41.28	26.36	17.62	8.11
Bregman 5 (Taggart) [a = 195] loss	27.31	15.95	9.69	3.34	29.68	16.69	9.64	2.93
Mean identification	0.69	0	0	1.56	1.04	0.35	0.35	1.9
Quantile identification [τ = 0.5]	-0.04	-0.06	-0.05	0.03	-0.03	-0.04	-0.04	0.03

Predictive median — Loss or identification function (rows) × Model (columns)

Loss or identification function	Train: GPL 1 [τ = 0.5] loss model	Train: GPL 2 [τ = 0.5] (absolute) loss model	Train: GPL 3 [τ = 0.5] loss model	Train: GPL 4 [τ = 0.5] loss model	Test: GPL 1 [τ = 0.5] loss model	Test: GPL 2 [τ = 0.5] (absolute) loss model	Test: GPL 3 [τ = 0.5] loss model	Test: GPL 4 [τ = 0.5] loss model
GPL 1 [τ = 0.5] loss	0.07	0.07	0.07	0.08	0.07	0.07	0.07	0.08
GPL 2 [τ = 0.5] (absolute) loss	7.97	7.82	7.96	8.52	8.16	7.99	8.13	8.68
GPL 3 [τ = 0.5] loss	2229.81	2114.17	2080.69	2123.71	2275.38	2144.6	2103.79	2143.65
GPL 4 [τ = 0.5] loss	522458.38	481917.63	462145.57	454121.99	535411.61	489478.31	466944.05	457625.05
GPL 5 (Taggart) [τ = 0.5, a = 180] loss	2.24	1.82	1.56	1.34	2.51	2.02	1.72	1.48
GPL 5 (Taggart) [τ = 0.5, a = 185] loss	1.72	1.31	1.05	0.84	1.89	1.42	1.14	0.91
GPL 5 (Taggart) [τ = 0.5, a = 190] loss	1.15	0.8	0.58	0.41	1.26	0.87	0.63	0.43
GPL 5 (Taggart) [τ = 0.5, a = 195] loss	0.67	0.4	0.25	0.14	0.71	0.41	0.24	0.13
Mean identification	1.61	1.31	1.55	2.5	1.96	1.66	1.89	2.83
Quantile identification [τ = 0.5]	0	0	0.02	0.06	0.01	0.01	0.03	0.06

Predictive 0.90-quantile — Loss or identification function (rows) × Model (columns)

Loss or identification function	Train: GPL 1 [τ = 0.9] loss model	Train: GPL 2 (pinball) [τ = 0.9] loss model	Train: GPL 3 [τ = 0.9] loss model	Train: GPL 4 [τ = 0.9] loss model	Test: GPL 1 [τ = 0.9] loss model	Test: GPL 2 (pinball) [τ = 0.9] loss model	Test: GPL 3 [τ = 0.9] loss model	Test: GPL 4 [τ = 0.9] loss model
GPL 1 [τ = 0.9] loss	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
GPL 2 (pinball) [τ = 0.9] loss	1.52	1.5	1.53	1.6	1.54	1.51	1.53	1.6
GPL 3 [τ = 0.9] loss	449.51	434	427.56	434.16	452.8	433.37	425.94	432.42
GPL 4 [τ = 0.9] loss	110142.3	104387.87	100362.5	99171.27	111200.53	104555.54	100362.66	99073.46
GPL 5 (Taggart) [τ = 0.9, a = 180] loss	0.52	0.46	0.42	0.39	0.56	0.51	0.47	0.44
GPL 5 (Taggart) [τ = 0.9, a = 185] loss	0.43	0.38	0.33	0.3	0.46	0.41	0.36	0.33
GPL 5 (Taggart) [τ = 0.9, a = 190] loss	0.33	0.28	0.24	0.21	0.36	0.31	0.26	0.23
GPL 5 (Taggart) [τ = 0.9, a = 195] loss	0.24	0.2	0.16	0.13	0.27	0.22	0.17	0.14
Quantile identification [τ = 0.9]	0	0	0.01	0.02	0	0	0.02	0.03

4. Key findings and takeaways

See also:

Overall calibration

- The models were well-calibrated, as indicated by the computed identification functions.

Objective alignment

- A model trained with a specific loss function consistently achieves the best performance (rank 1) when evaluated using the same loss function.
- There is no "universally best" model across loss functions. There are only "fit-for-purpose" models.
- A specific directive (target functional) during calibration is not sufficient on its own for optimal performance. The target loss function is needed.

Performance on extremes

- Loss functions stemming from higher-order φ and g functions consistently demonstrate better performance on extremes.
- This aligns with established intuition in hydrological literature.
- However, in this work, the transformations are introduced directly into the Bregman and GPL functions rather than their special cases, resulting in consistent loss functions for pre-specified target functionals of engineering relevance (e.g., mean, median, other quantiles).

The value of abstraction

- Toy models serve as essential tools for building the core intuition needed to effectively engage with more complex, real-world hydrological systems.