Supplement of Hydrol. Earth Syst. Sci., 29, 5005–5029, 2025
https://doi.org/10.5194/hess-29-5005-2025-supplement

Hydrology and Earth System Sciences
Open Access

Supplement of

How well do process-based and data-driven hydrological models learn from limited discharge data?

Maria Staudinger et al.

Correspondence to: Maria Staudinger (maria.staudinger@geo.uzh.ch)

S1 Parameter ranges for the different process-based models

Table S1. Parameter ranges used for the GR4J model, plus snow parameters

Parameter	Min	Max	Description
x01	0.01	1.5	Capacity of production store
x02	-5.0	5.0	Water exchange coefficient
x03	1.0	500	Capacity of routing store
x04	0.5	5.0	Time parameter of unit hydrograph
x05	0.0	500	Average annual snow
x06	0.01	0.99	Air snow coefficient

Table S2. Parameters ranges of the HBV model.

Parameter	Min	Max	Description
TT	-2.5	2.5	Threshold temperature [$^\circ$ C]; threshold defining over which air temperature snow is melting and under which snow is accumulating
SFCF	0.8	1.2	Snow correction factor[-]; correcting the snow input to account for gauge undercatch (overcatch)
CFMAX	2.0	15	degree day factor [mm/d $^\circ$ C]; defining the rate of snowmelt per degree temperature
LP	0.0	1.0	threshold reduction ETP [-]; reducing the potential ET to estimate actual ET
FC	50	1000	Maximum storage in soil box [mm]; defines the size of the soil bucket
BETA	1.0	5.0	Shape coefficient [-]; shapes the relation between soil moisture [mm] and fraction of rain (or snowmelt) and thus its relative contribution to runoff
Alpha	0.0	1.0	Shape coefficient [-]; shapes the relation between water storage in the upper groundwater (gw) bucket and drainage
K1	0.01	0.4	Recession coefficient (upper gw bucket) [1/d]
PERC	0.0	6.0	Max. flow from upper to lower gw box [-]
K2	0.0001	0.1	Recession coefficient (lower gw bucket) [1/d]
MAXBAS	1.0	5.0	Factor triangular weighting [d]

Table S3. Parameter ranges used for the SWAT+ model.

Parameter	Min	Max	Description
snomelt_tmp	-2.0	2.0	Snowmelt temperature [$^\circ$ C]
esco	0.0	1.0	Soil evaporation compensation factor
epco	0.0	1.0	Plant uptake compensation factor
cn3_swf	-0.5	0.5	Soil water factor for cn3
awc	0.0	0.2	Available water capacity of soil layer
k	0.5	2.0	Saturated hydraulic conductivity
lat_ttime	0.5	50	Lateral flow travel time
perco	0.0	1.0	Percolation from upper to lower tank
alpha	0.0	1.0	Baseflow factor
surlag	0.2	8.0	surface runoff lag coefficient
cn2	-10	10	SCS curve number for moisture condition II

S2 Settings Latin Hypercube Sampling

For the Latin Hypercube Sampling, all model parameters were sampled in their provided range in 1,000 repetitions. All resulting 1,000 model runs, can be considered at once to cover all 10 sample sizes since the sampling strategy is independent of the model performance of previous runs. In contrast, evolutionary algorithms such as dynamically dimensioned search (DDS, Tolson and Shoemaker (2007)) or shuffled complex evolution algorithm (SCE-UA, Duan et al. (1992)) define their search based on the model performance value of previous runs, the sampling would be unique for each sample size. Therefore, each of the 30 repetitions would have to run separately for each sample size (10 times more computational effort). The Latin Hypercube Sampling was performed by using the Python framework SPOTPY (Houska et al., 2015).

S3 Example of the Douglas Peucker sampling scheme

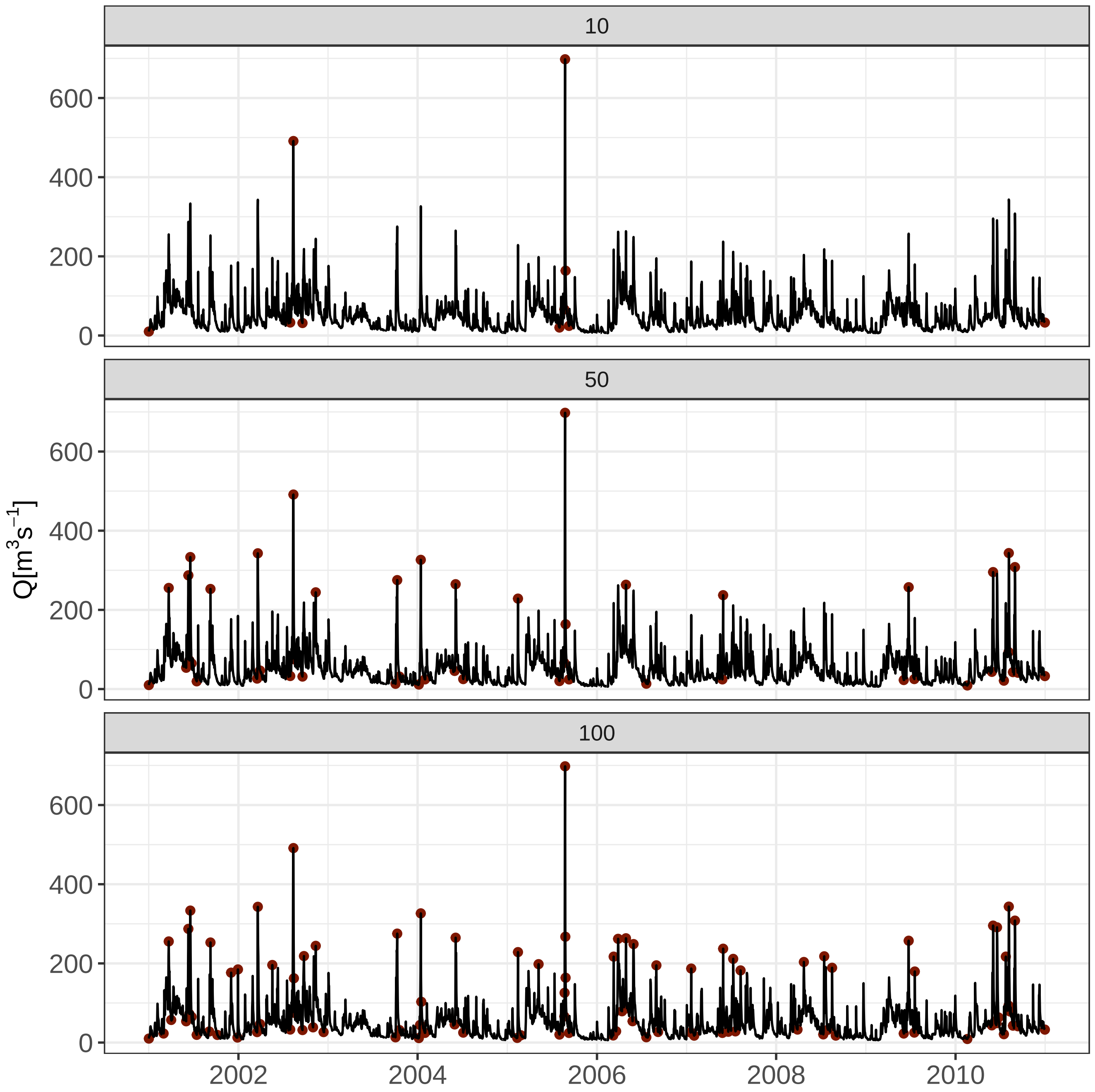

Figure S1. Douglas-Peucker sampling points for sample size 10, 50, and 100 for the training period of the Iller catchment.

S4 Comparison learning curves using the Kling Gupta efficiency (KGE)

Figure S2 compares the learning curves of the different data- and process-based models for the three study catchments Iller, Saale and Selke using the KGE as evaluation metric. Note that the learning curves are made for the validation period, not the calibration period.

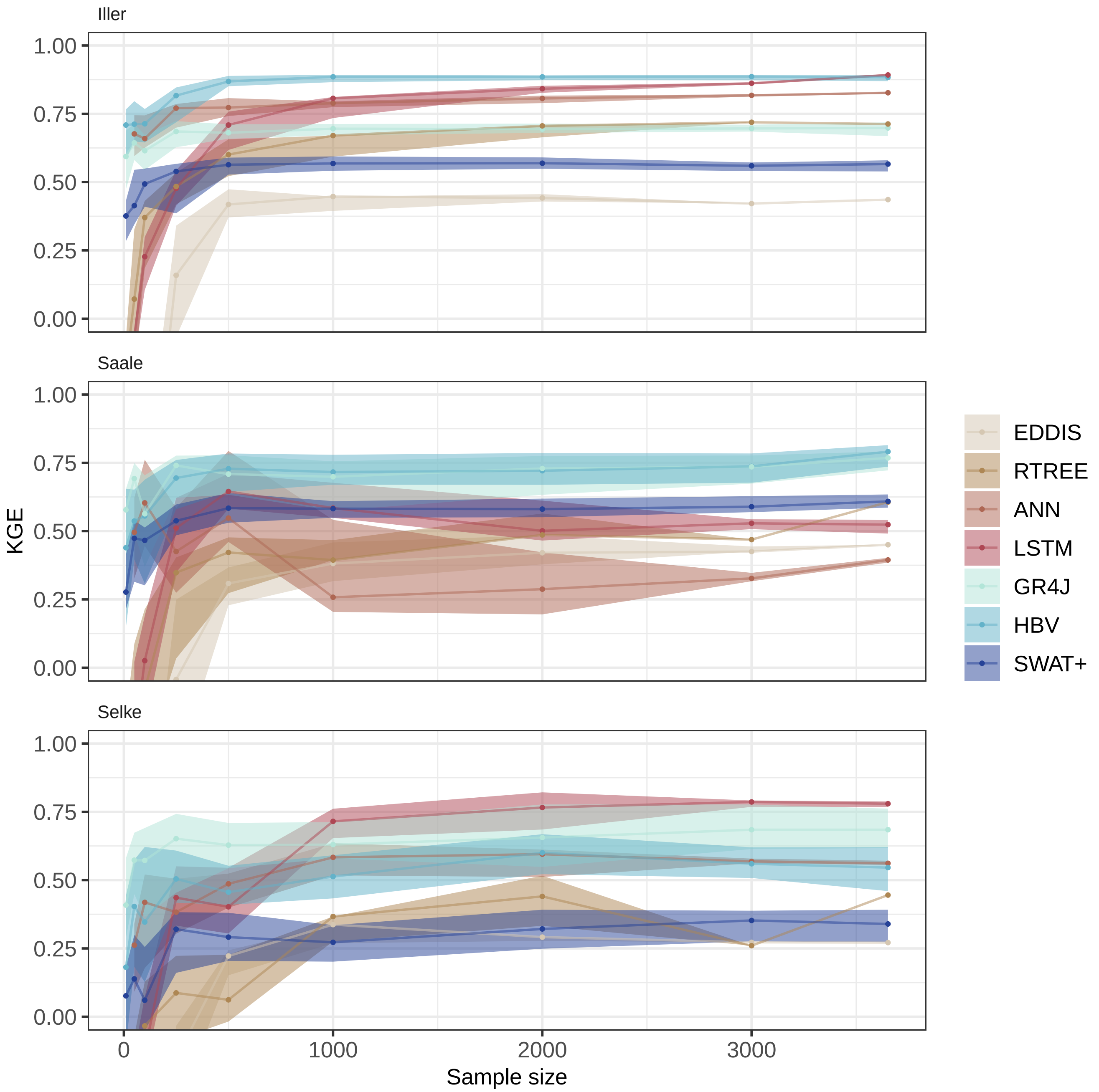

Figure S2. Learning curve using the continuous random sampling strategy for the different models and catchments, Kling Gupta efficiency, KGE. The closer the values to 1 the better the model performance.

S5 Comparison sampling schemes using the Kling Gupta efficiency (KGE)

Figure S3 compares the learning curves of the HBV model for three different sample schemes (optimal = Douglas-Peucker, random, random consecutive) for the three study catchments Iller, Saale and Selke using the KGE as evaluation metric. Note that the learning curves are made for the validation period, not the calibration period.

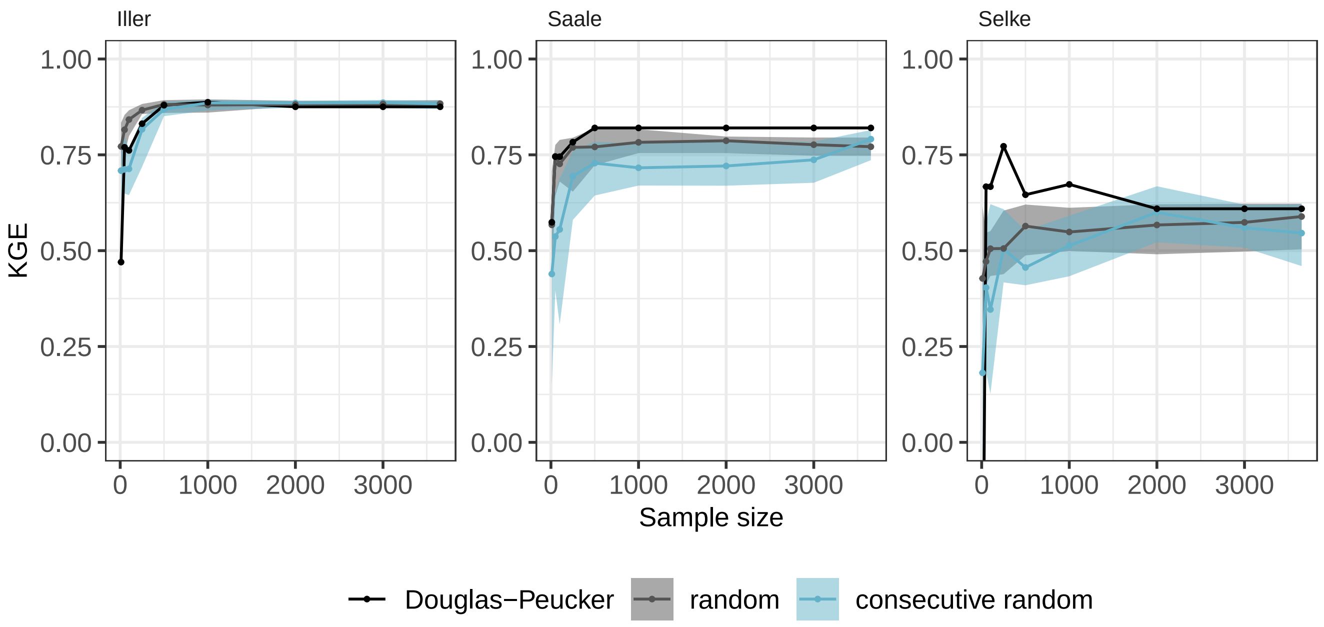

Figure S3. Learning curves using Kling Gupta efficiency, KGE. Three different sampling schemes are compared for the HBV model and for the three study catchments Iller, Saale and Selke. The closer the values to 1 the better the model performance.

S6 Comparison model performance with different level of discretization of model forcing using the Kling Gupta efficiency (KGE)

Figure S4 compares the KGE values when forcing the HBV model once with lumped and once with semi-distributed, i.e. sub-catchment wise, meteorological input for the three study catchments Iller, Saale and Selke.

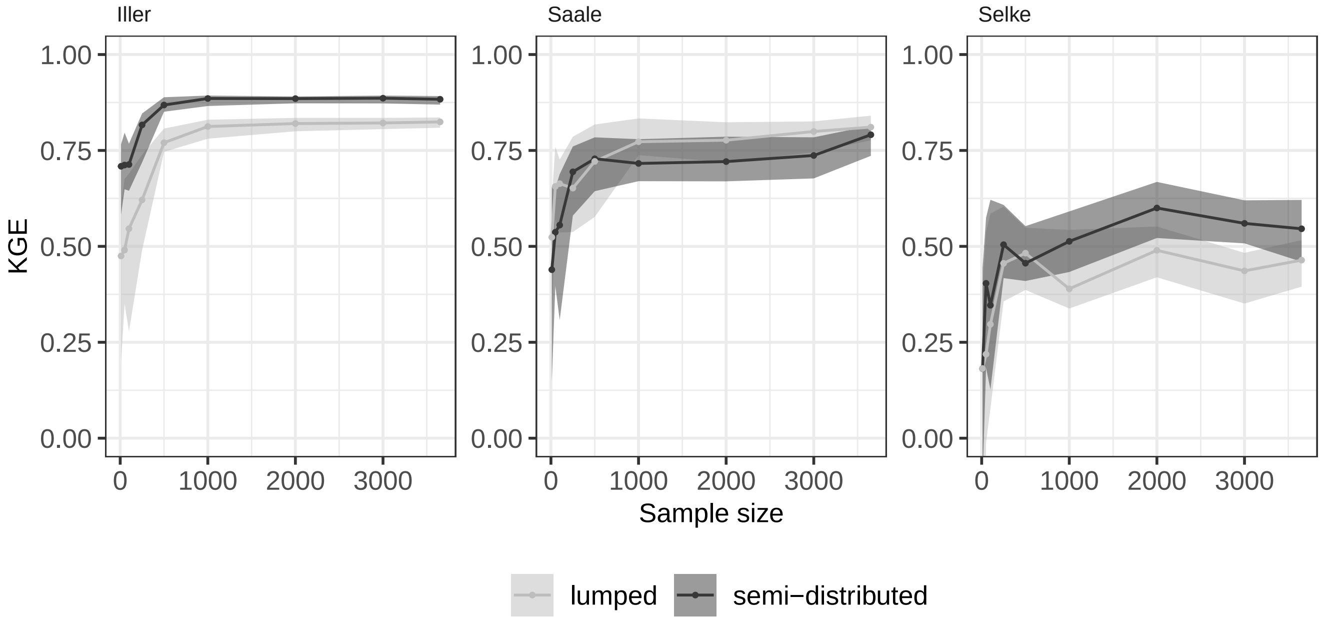

Figure S4. Learning curves using Kling Gupta efficiency, KGE. Lumped and distributed forcing are compared for the HBV model and for the three study catchments Iller, Saale and Selke. The closer the values to 1 the better the model performance.

S7 Examples of observed and simulated hydrographs

S7.1 Process-based models

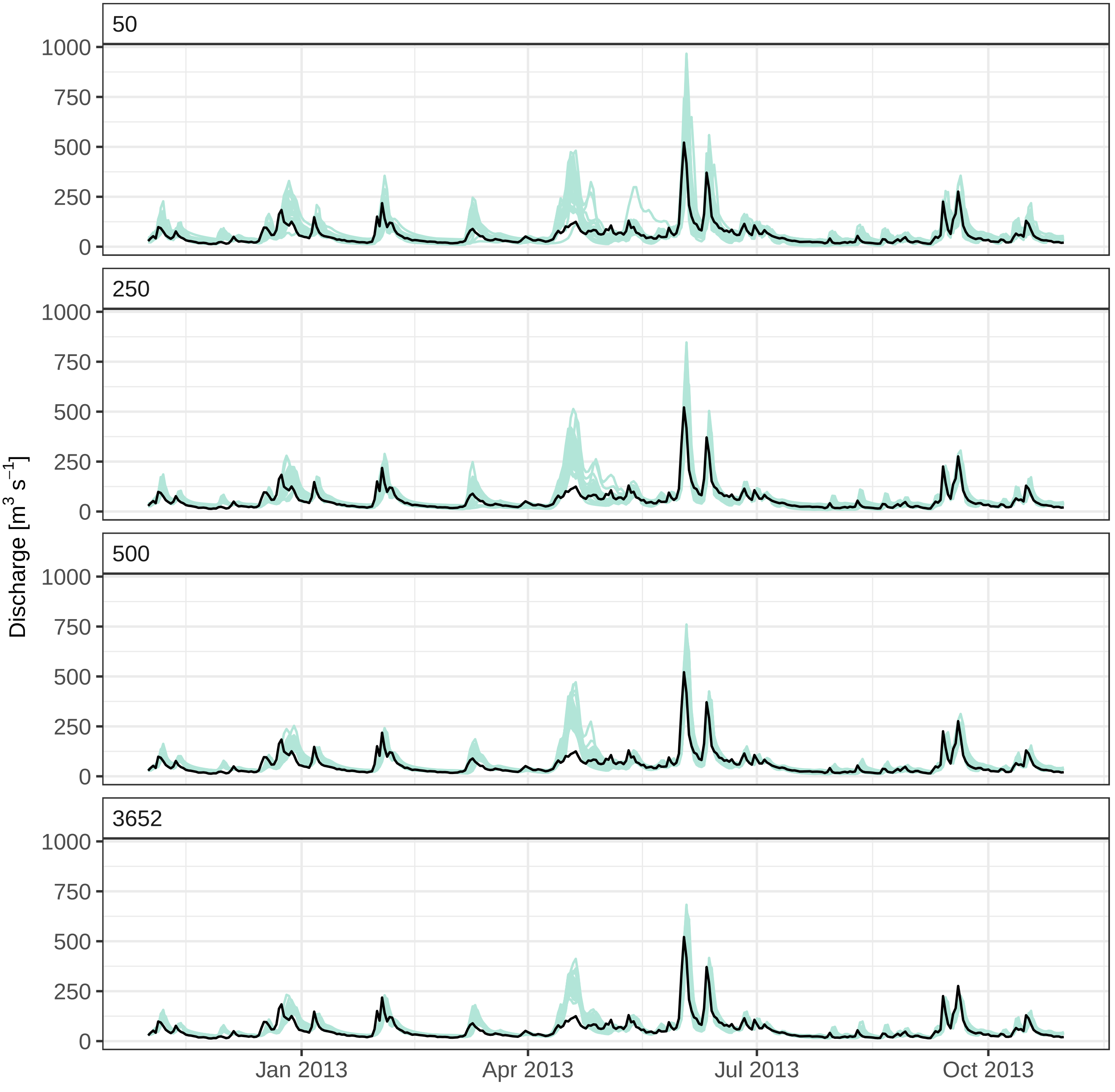

Figure S5. Hydrograph of the Iller at Wiblingen for the hydrological year 2013 (i.e. during validation) for three different sample sizes and all 30 repetitions for the GR4J model

S7.2 Data-driven models

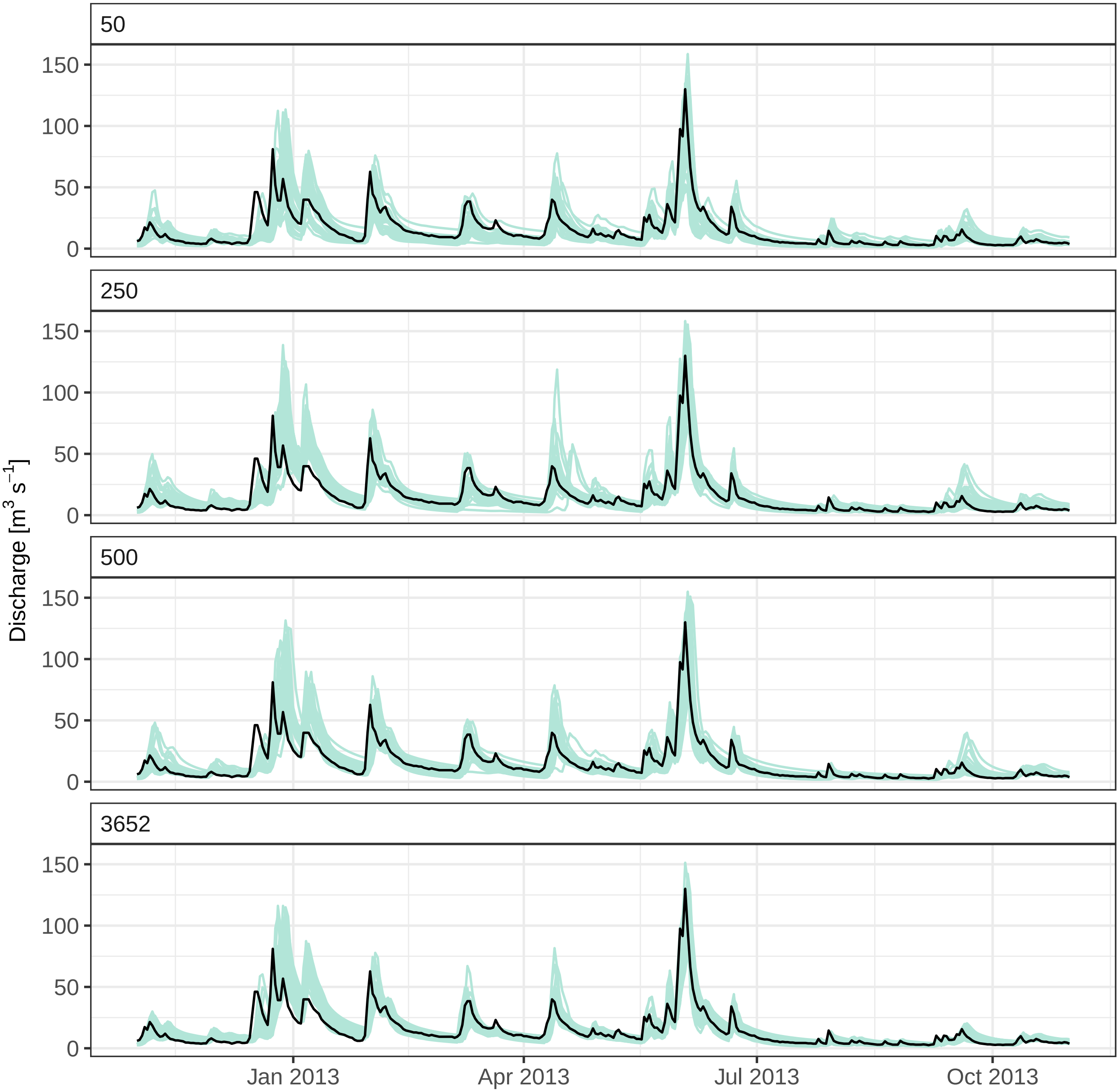

Figure S6. Hydrograph of the Saale at Blankenstein for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the GR4J model

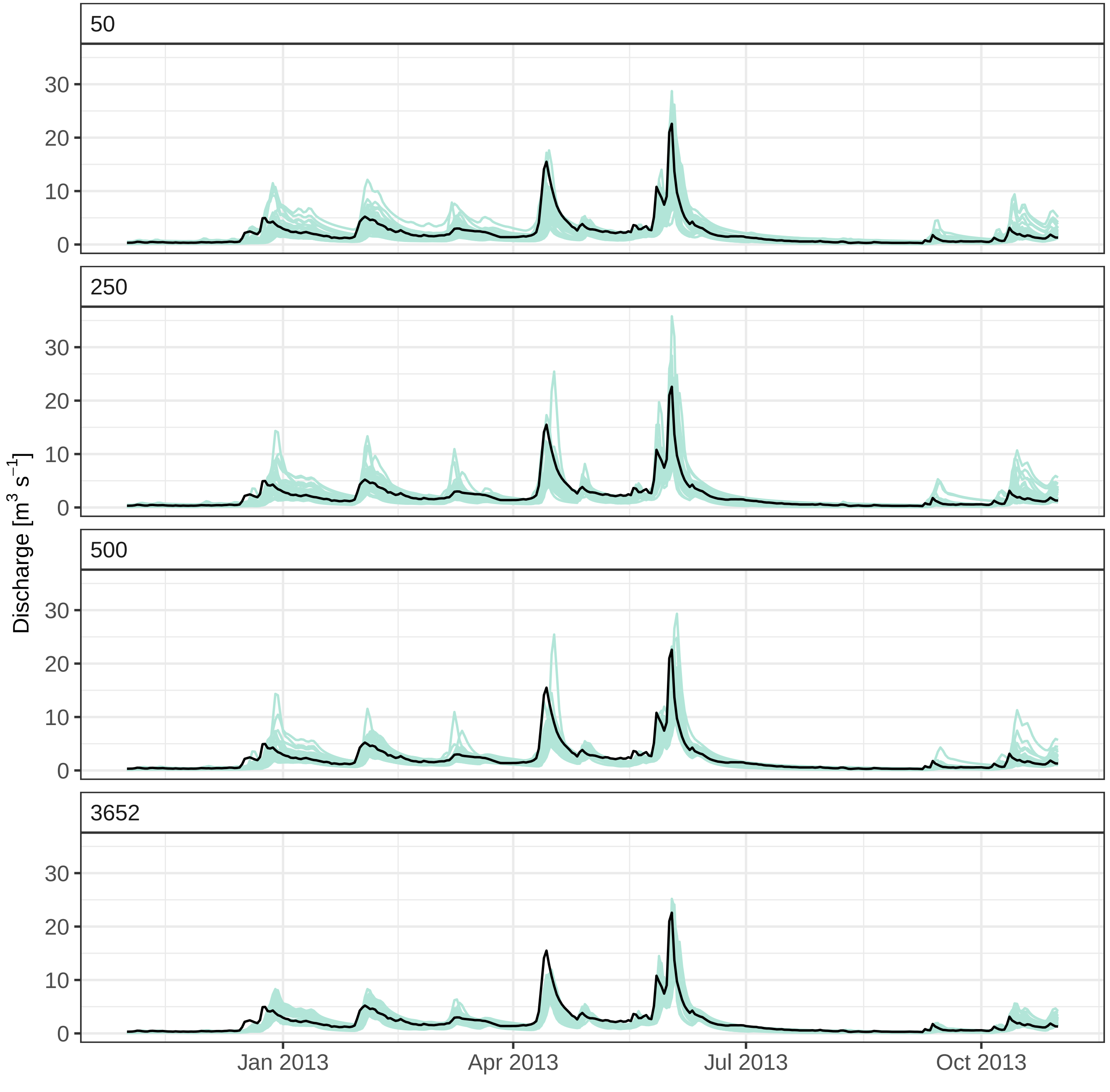

Figure S7. Hydrograph of the Selke at Hausneindorf for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the GR4J model

References

Duan, Q., Sorooshian, S., and Gupta, V.: Effective and efficient global optimization for conceptual rainfall-runoff models, Water Resources Research, 28, 1015–1031, https://doi.org/https://doi.org/10.1029/91WR02985, 1992.

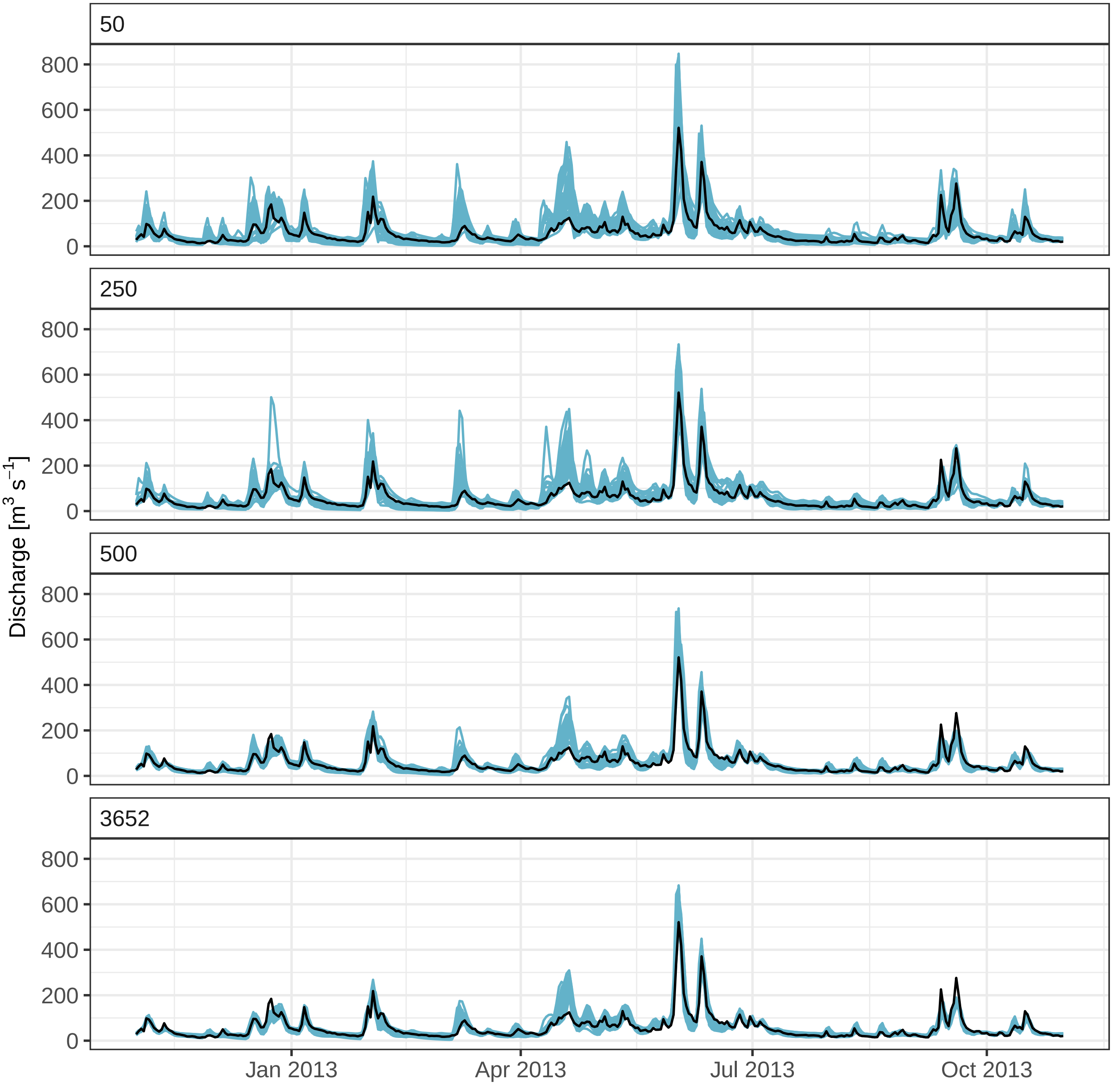

Figure S8. Hydrograph of the Iller at Wiblingen for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the HBV model

Houska, T., Kraft, P., Chamorro-Chavez, A., and Breuer, L.: SPOTting Model Parameters Using a Ready-Made Python Package, PLOS ONE, 10, 1–22, https://doi.org/10.1371/journal.pone.0145180, 2015.

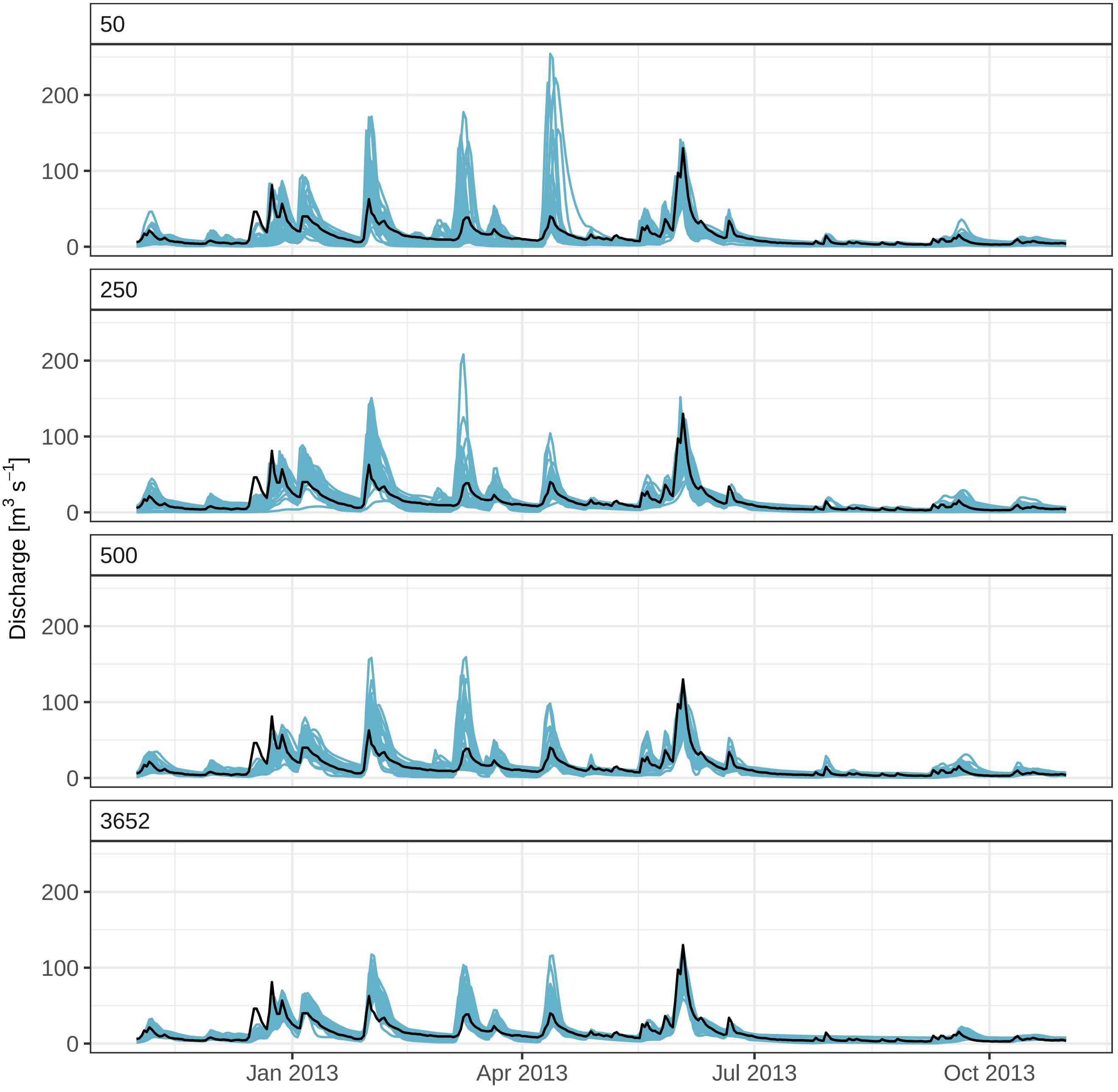

Figure S9. Hydrograph of the Saale at Blankenstein for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the HBV model

Tolson, B. A. and Shoemaker, C. A.: Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, Water Resources Research, 43, https://doi.org/10.1029/2005WR004723, cited by: 592, 2007.

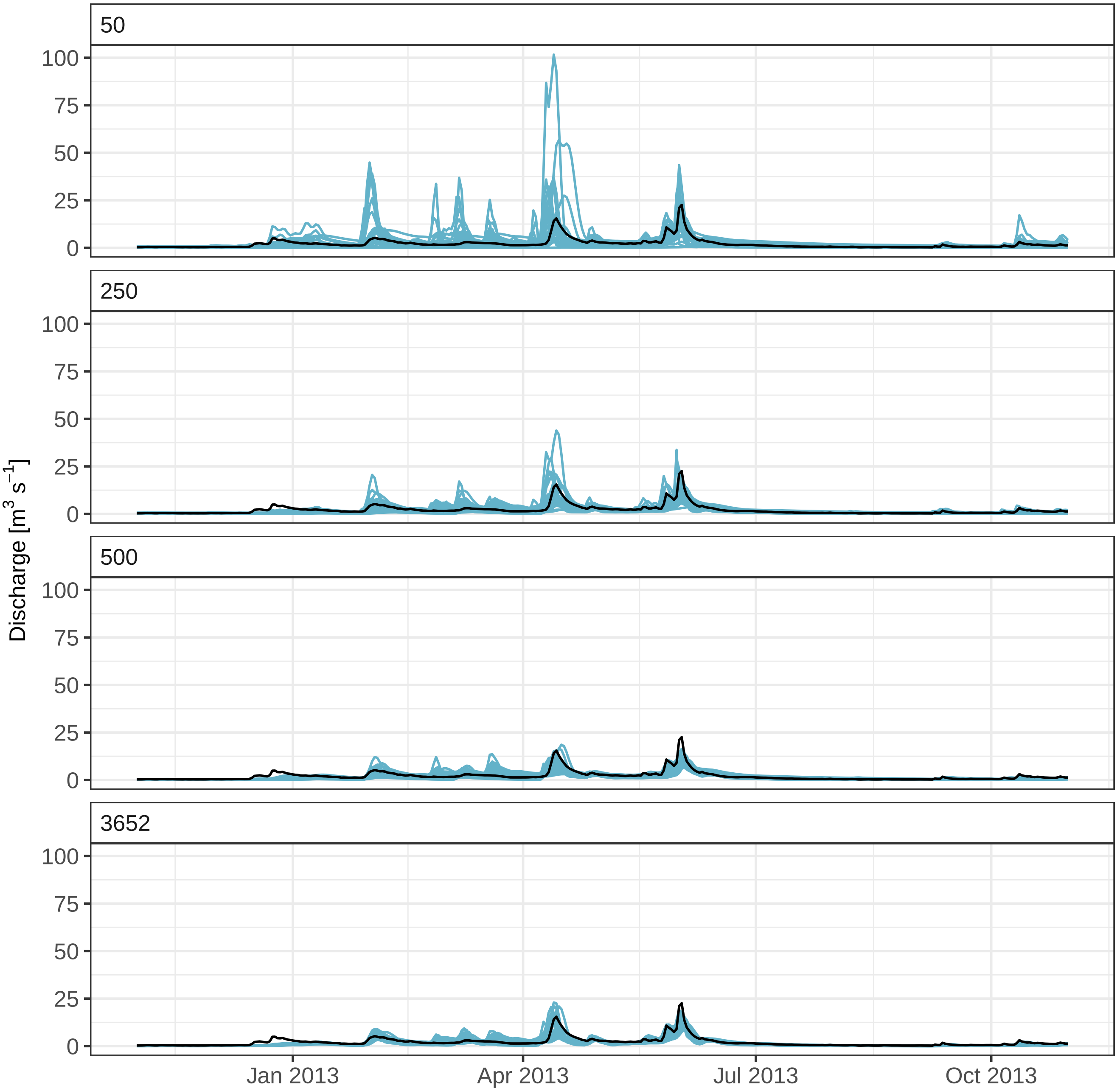

Figure S10. Hydrograph of the Selke at Hausneindorf for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the HBV model

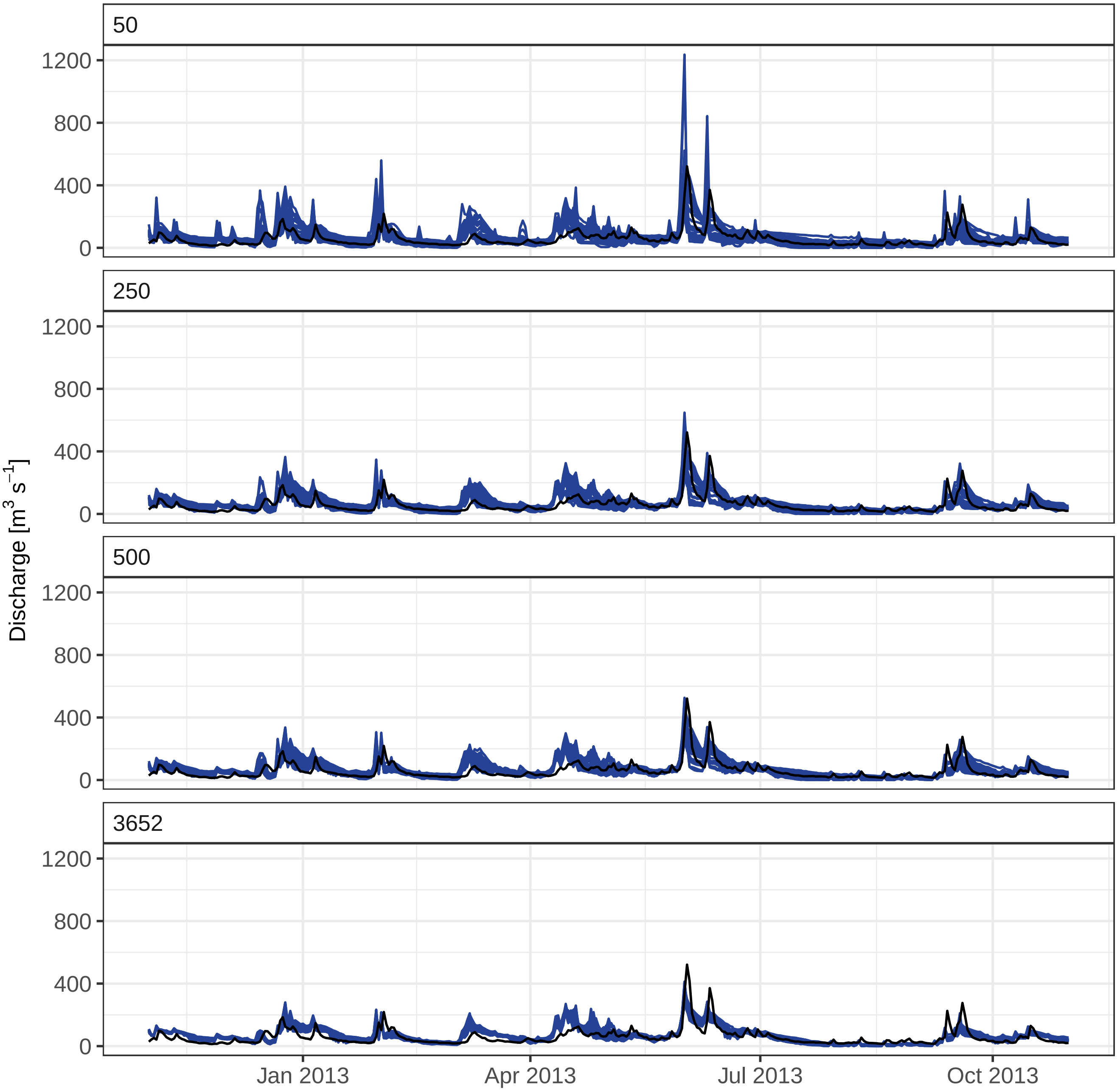

Figure S11. Hydrograph of the Iller at Wiblingen for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the SWAT+ model

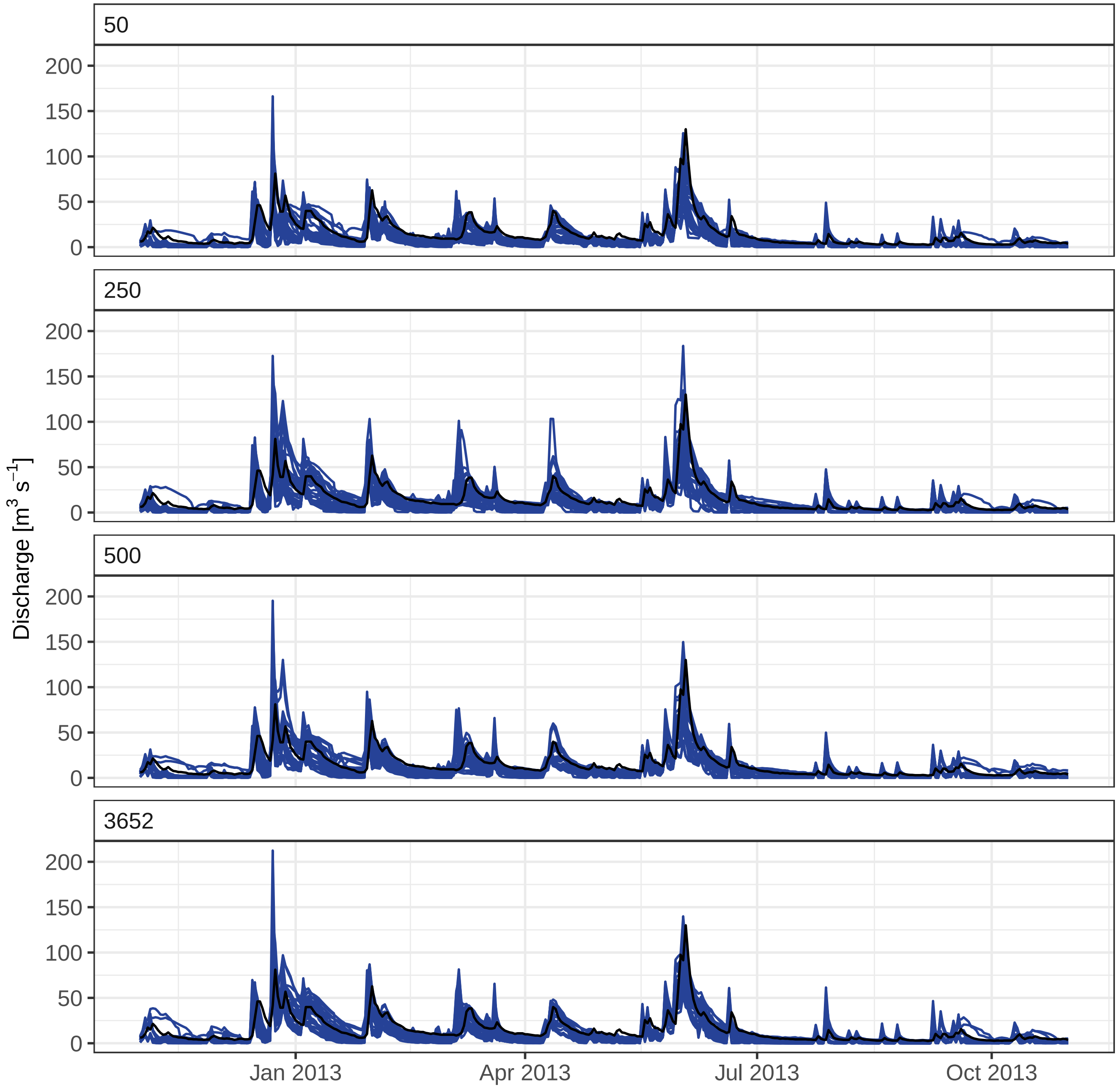

Figure S12. Hydrograph of the Saale at Blankenstein for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the SWAT+ model

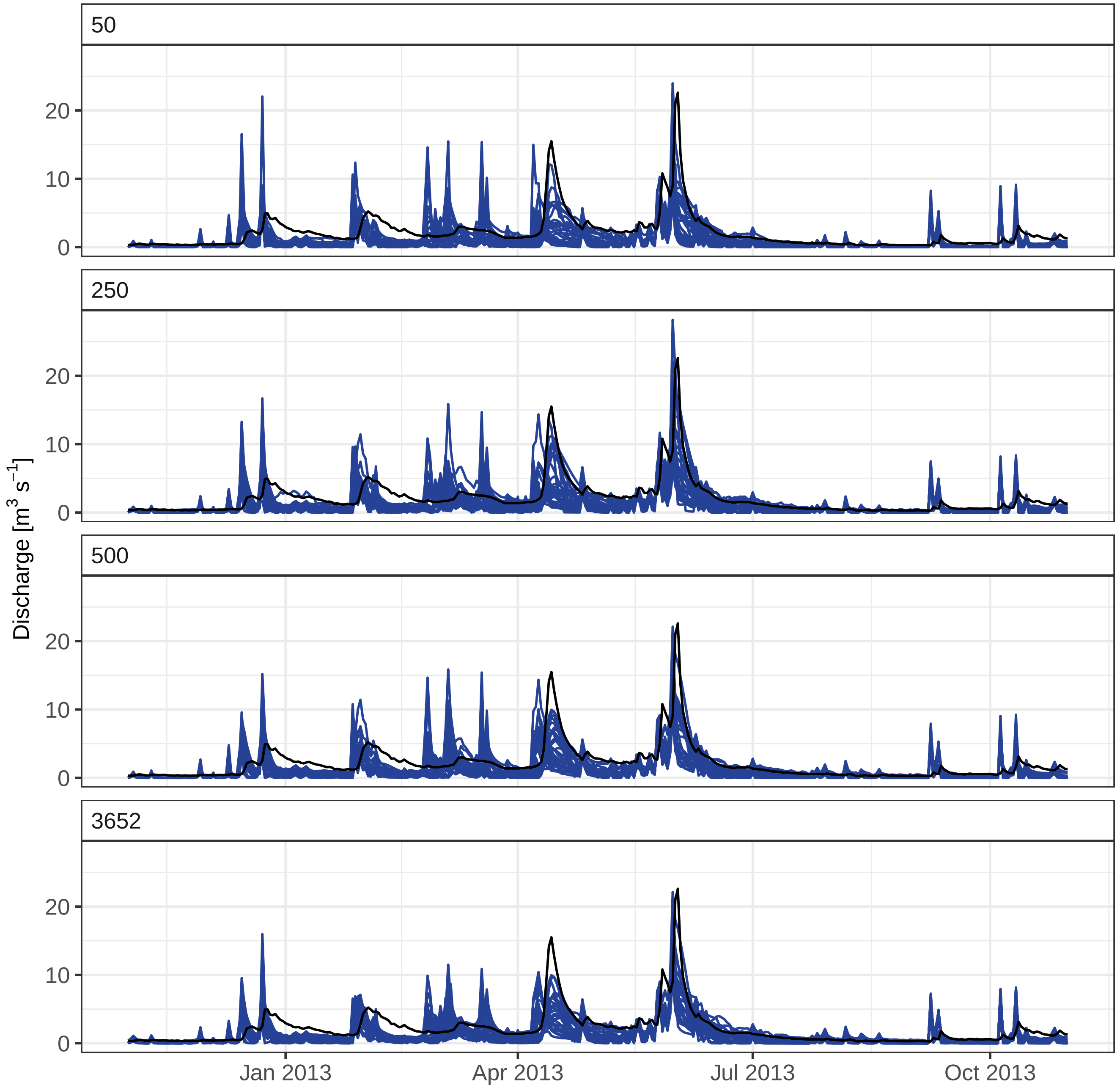

Figure S13. Hydrograph of the Selke at Hausneindorf for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the SWAT+ model

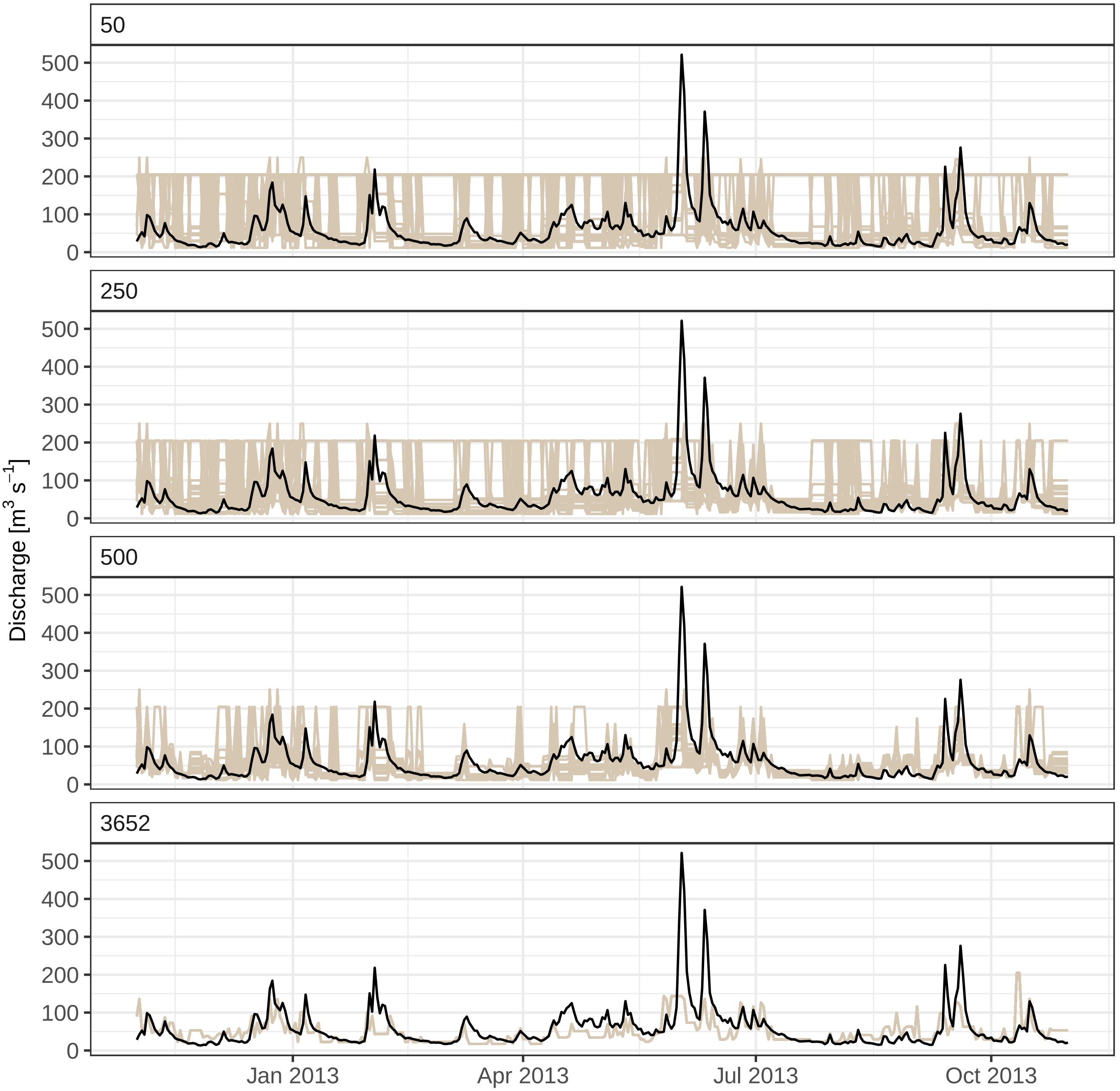

Figure S14. Hydrograph of the Iller at Wiblingen for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the EDDIS model

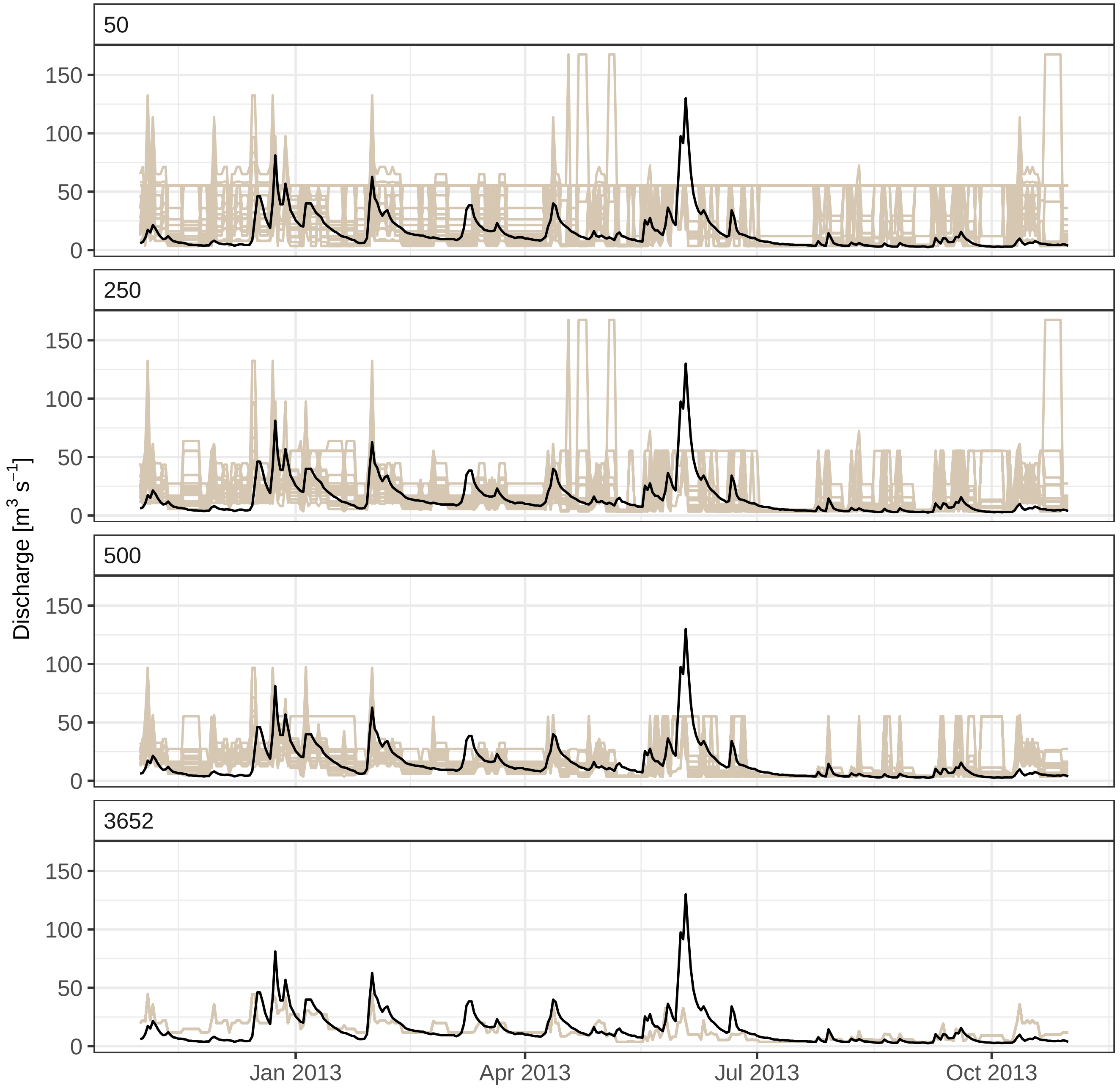

Figure S15. Hydrograph of the Saale at Blankenstein for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the EDDIS model

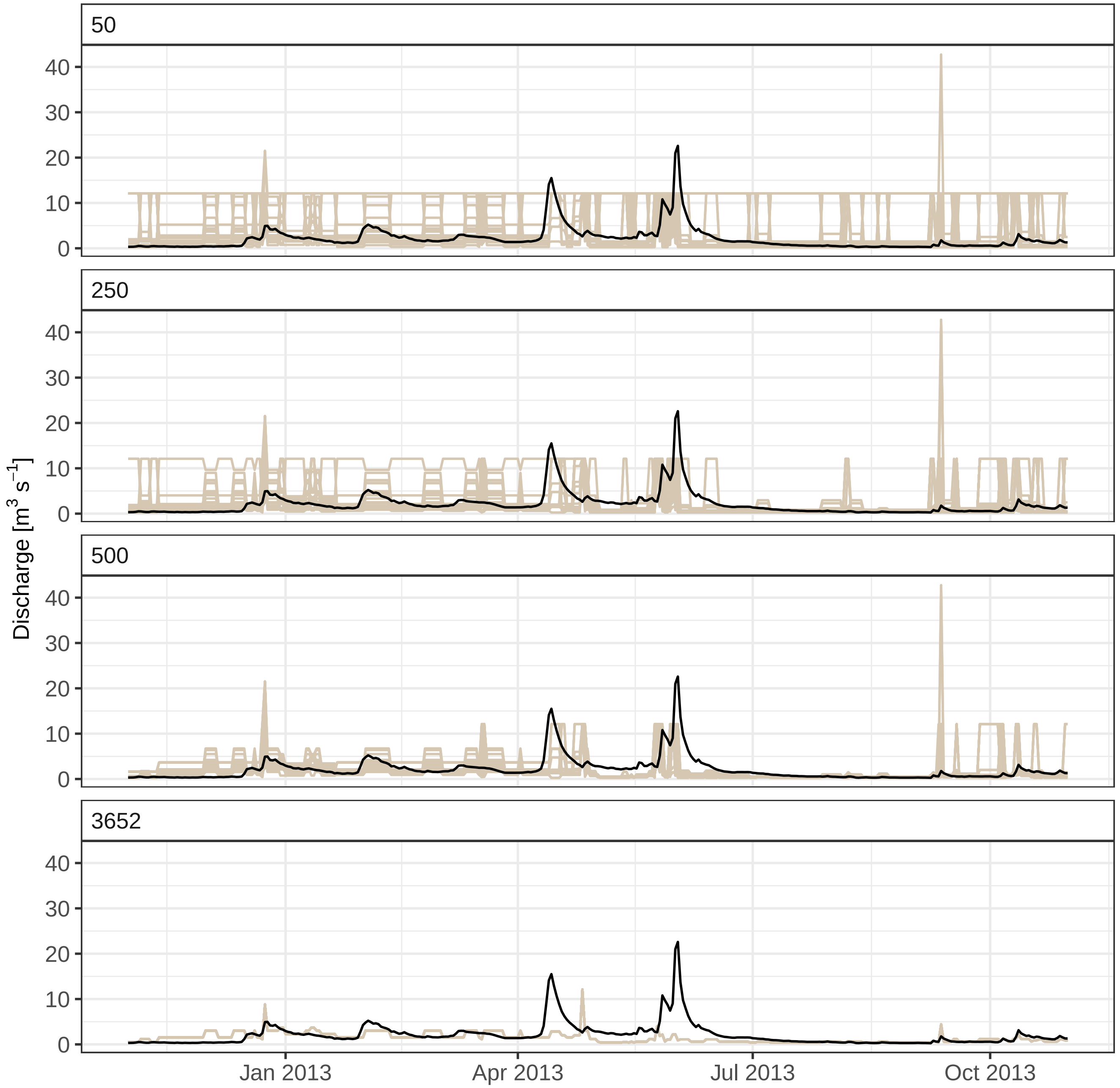

Figure S16. Hydrograph of the Selke at Hausneindorf for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the EDDIS model

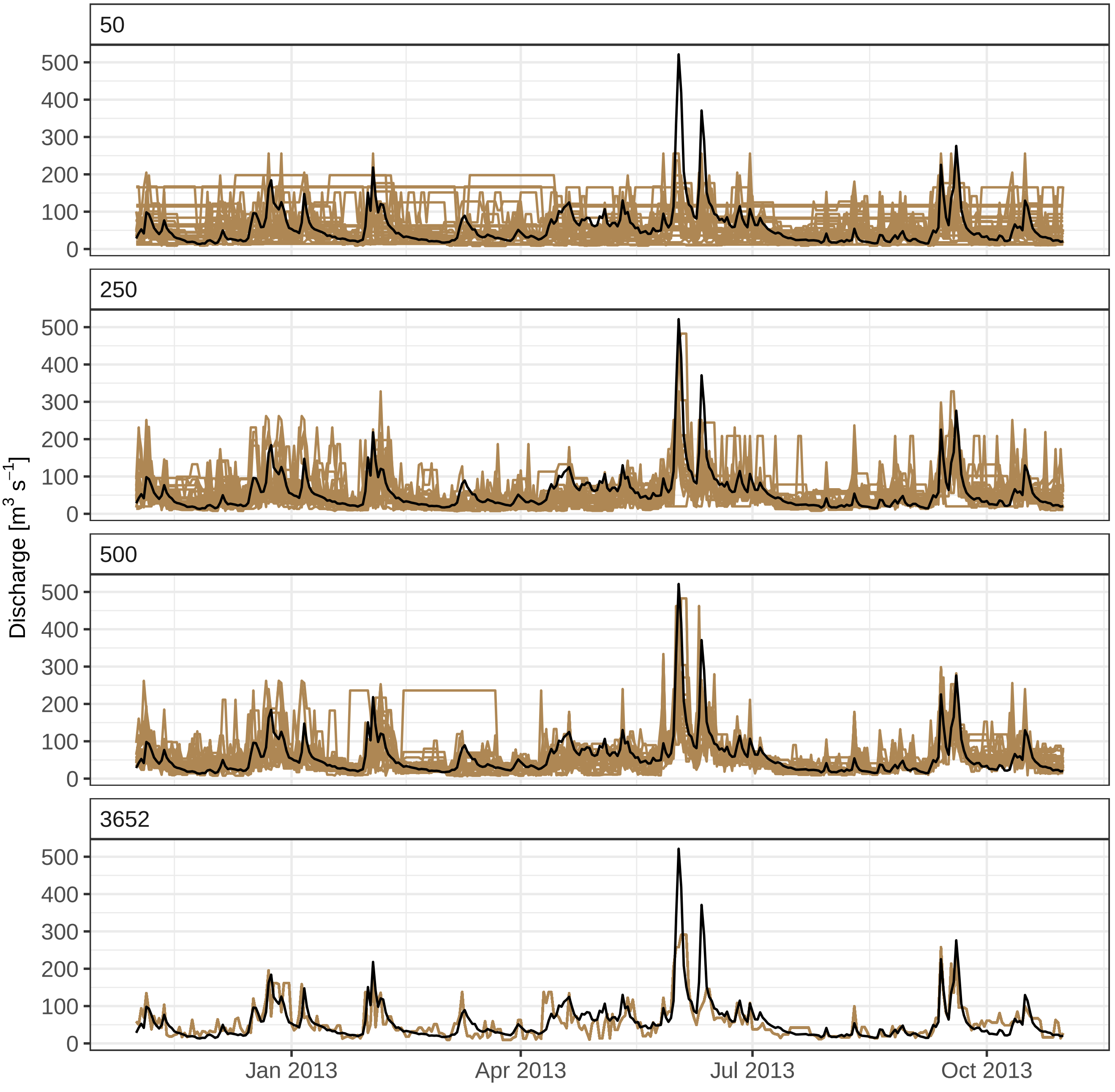

Figure S17. Hydrograph of the Iller at Wiblingen for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the RTREE model

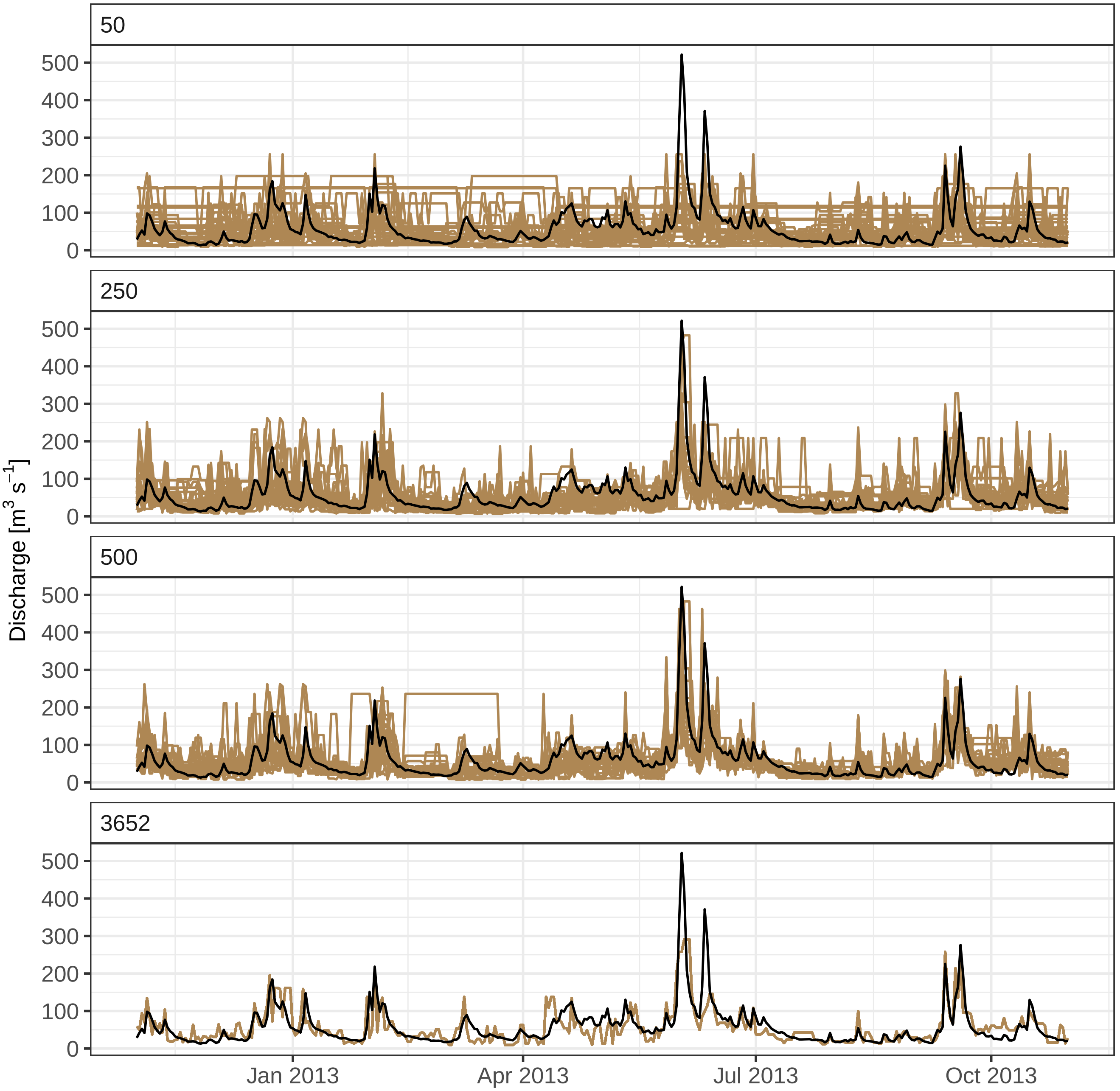

Figure S18. Hydrograph of the Saale at Blankenstein for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the RTREE model

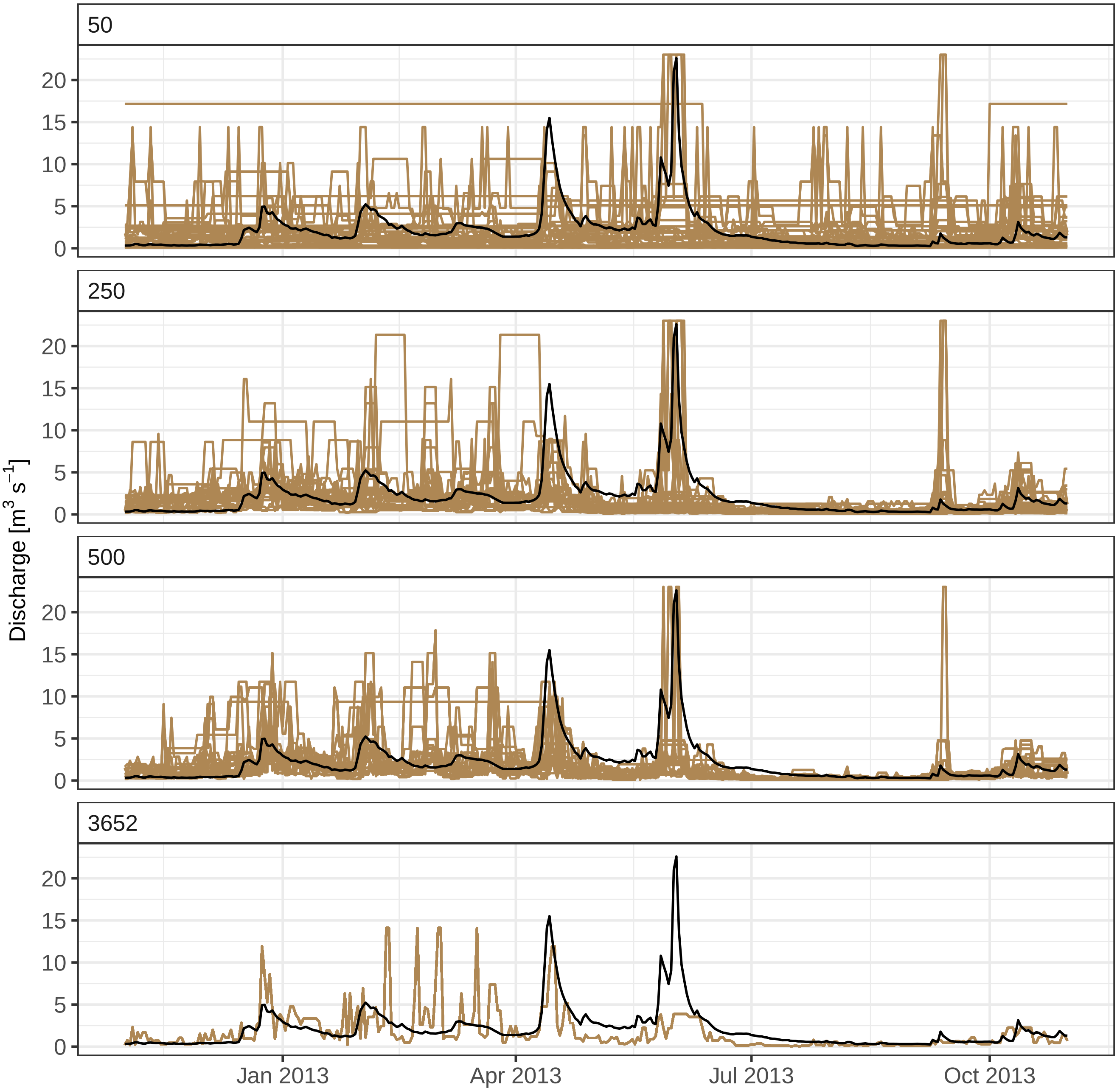

Figure S19. Hydrograph of the Selke at Hausneindorf for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the RTREE model

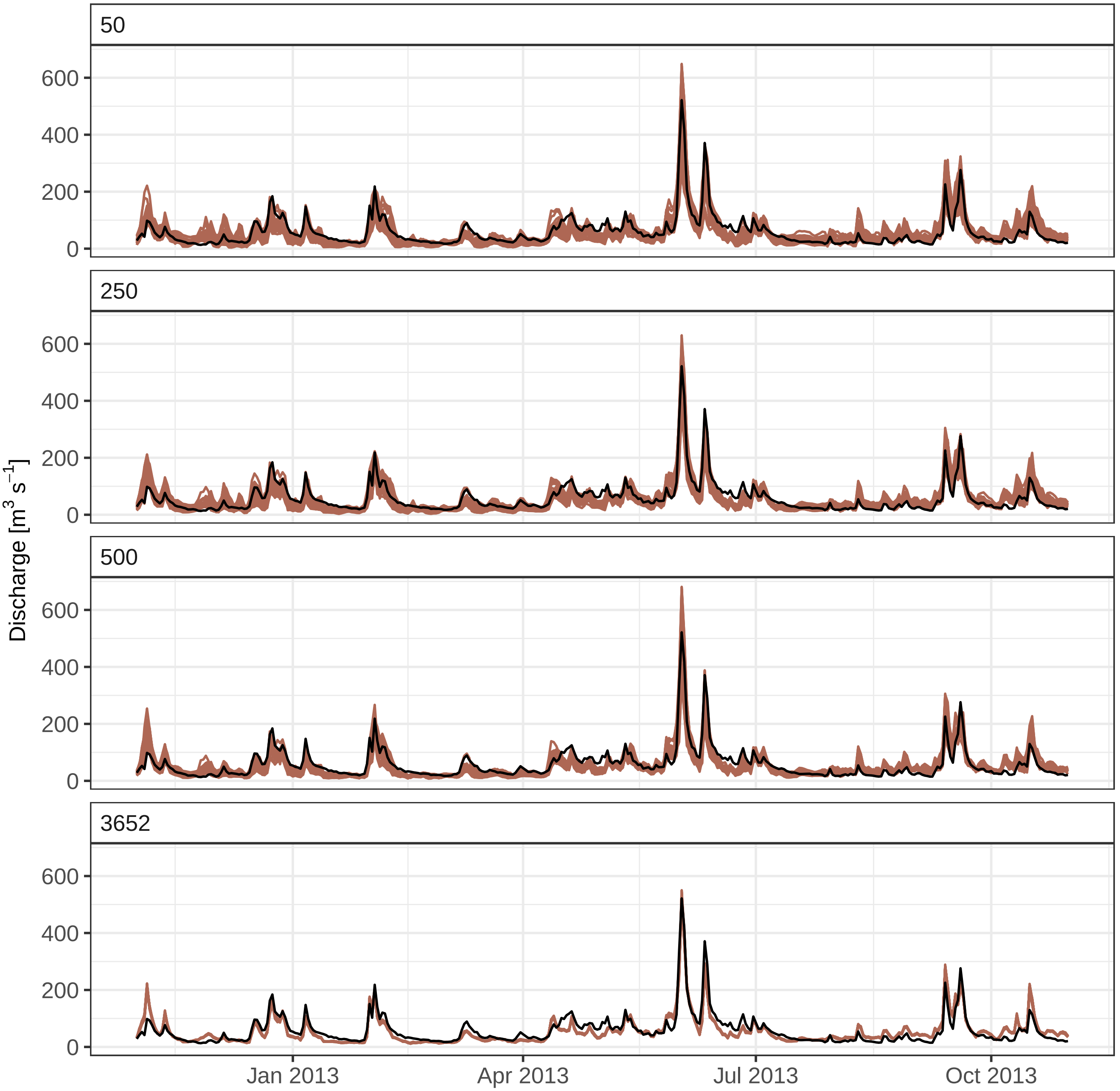

Figure S20. Hydrograph of the Iller at Wiblingen for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the ANN model

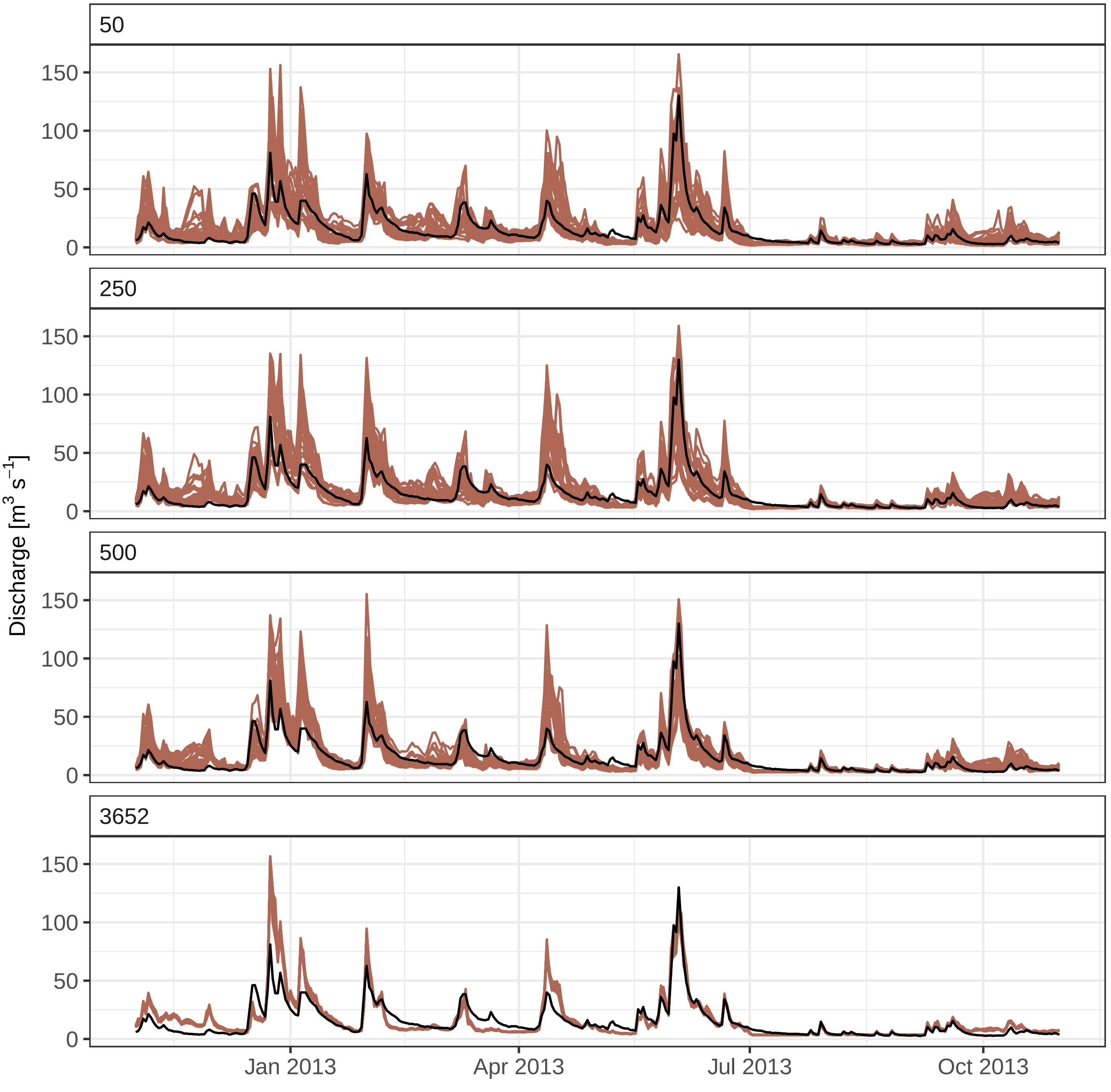

Figure S21. Hydrograph of the Saale at Blankenstein for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the ANN model

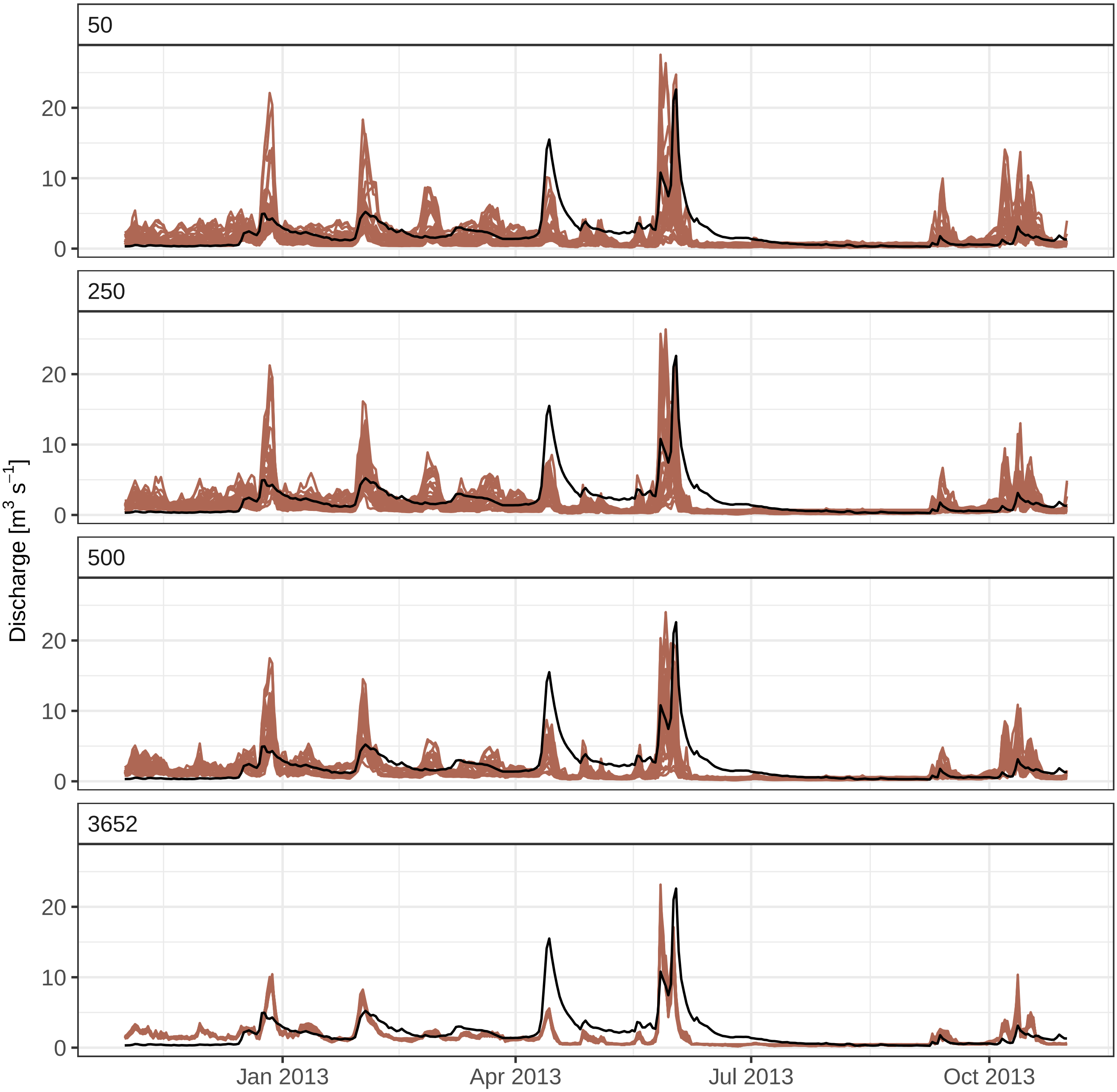

Figure S22. Hydrograph of the Selke at Hausneindorf for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the ANN model

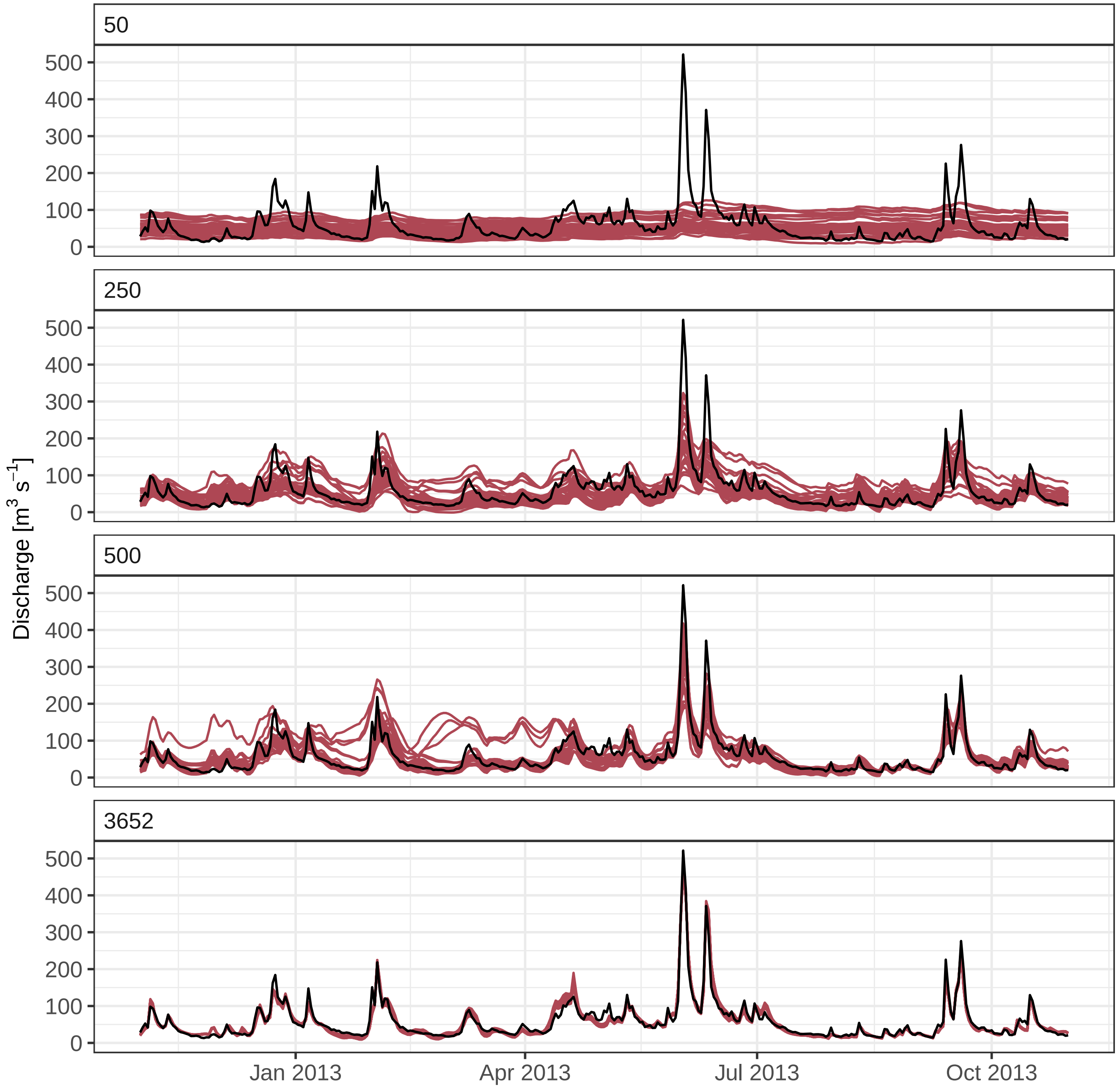

Figure S23. Hydrograph of the Iller at Wiblingen for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the LSTM model

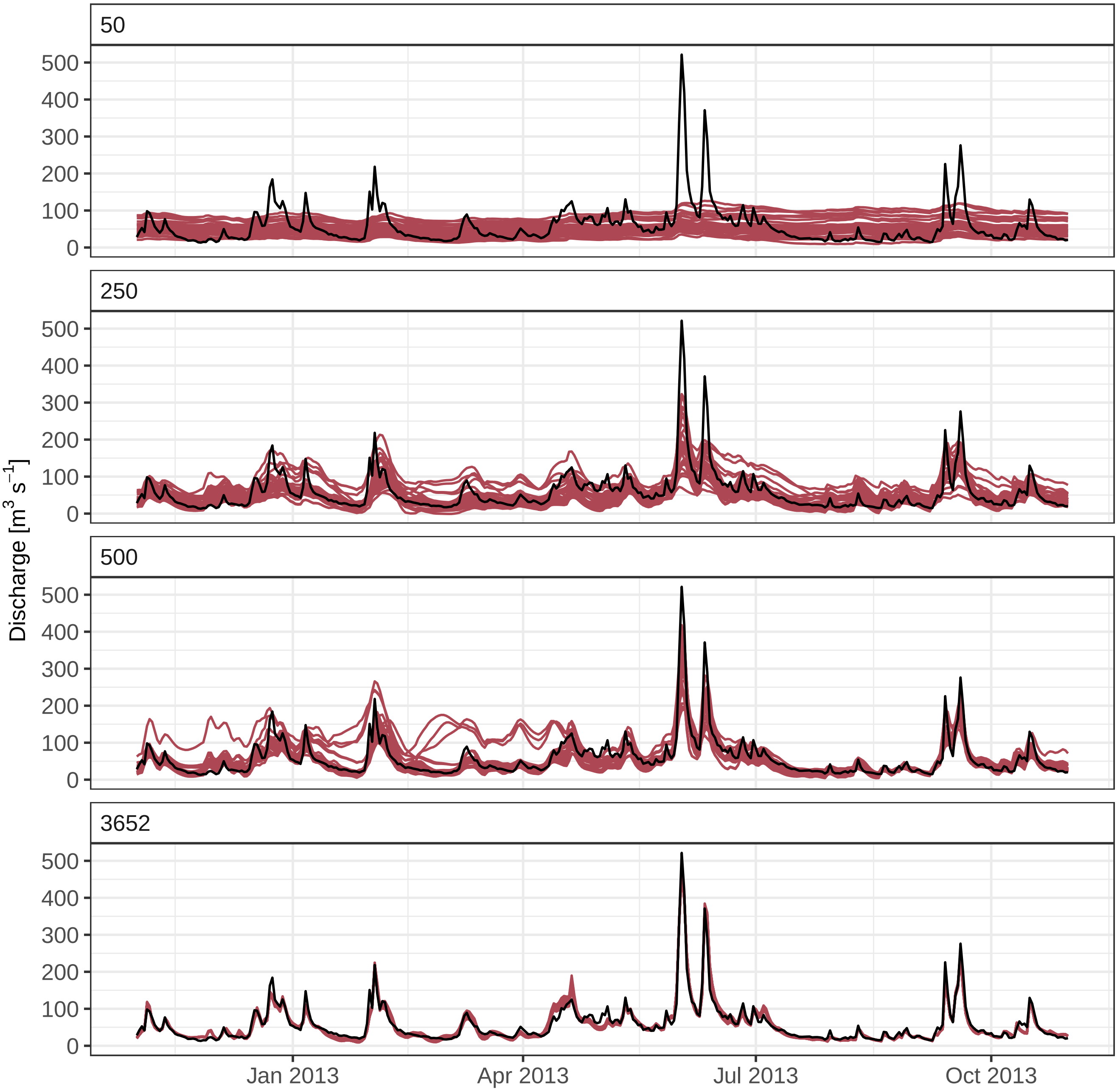

Figure S24. Hydrograph of the Saale at Blankenstein for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the LSTM model

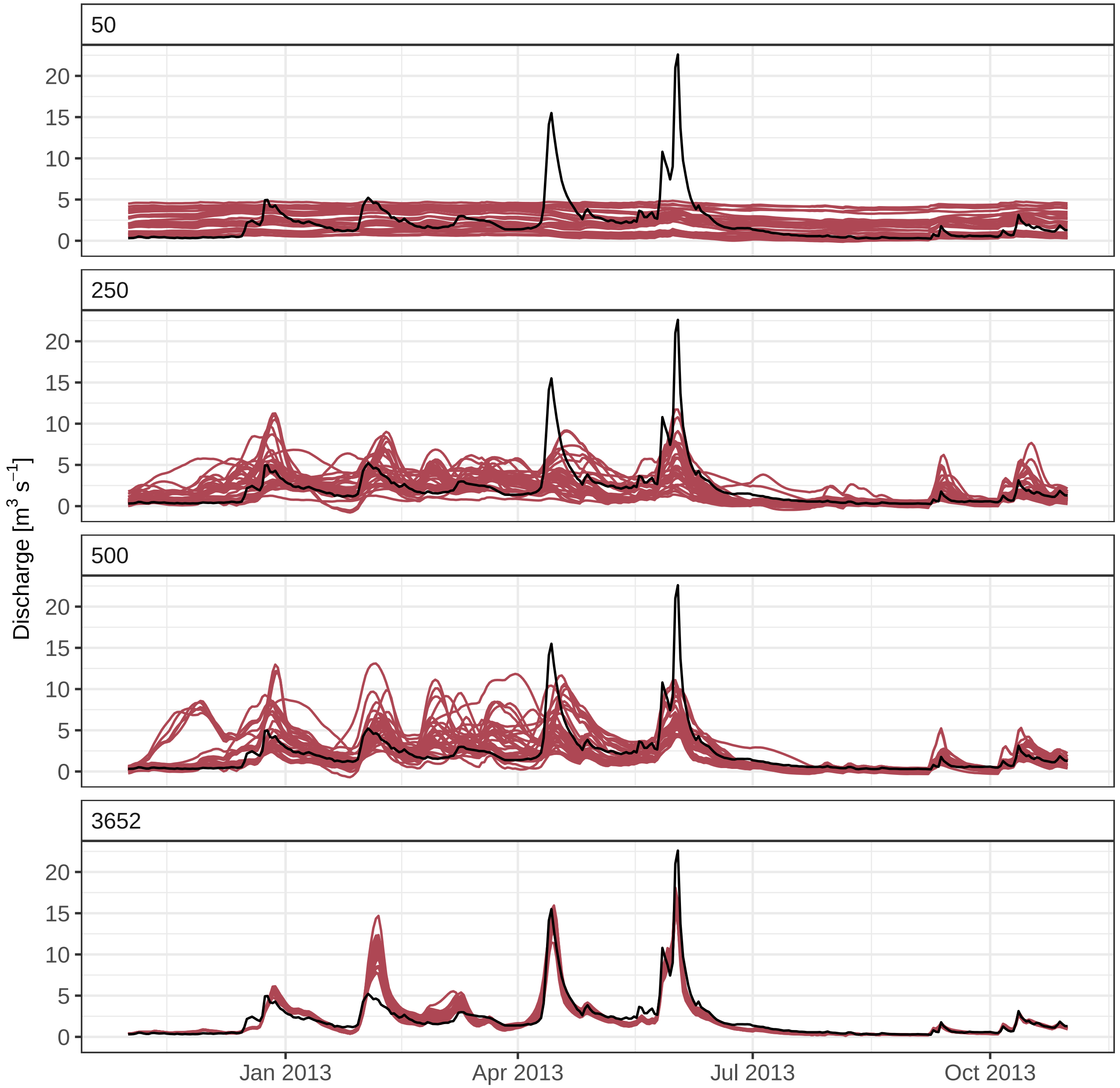

Figure S25. Hydrograph of the Selke at Hausneindorf for the hydrological year 2013 (i.e. during validation) for four different sample sizes and all 30 repetitions for the LSTM model