



Figure 11. Calibration and test MSE ratio between Bagging and SMOTER-Bagging models for the Bow (a) and (c) and Don (b) and (d) Rivers across high stage threshold values ranging from 50% to 90%.

“As discussed in Sect. 2.1, a fixed threshold is used to distinguish between high and typical stages. Fig. 11 shows the effects of the fixed threshold increasing from the 50th to 90th percentile of the stage distribution. These plots show the relative effects of SMOTER-Bagging compared to simple Bagging. A performance ratio greater than 1 indicates that the SMOTER-Bagging model has greater error compared to the Bagging model, 1 indicates that they have the same performance, and less than 1, improved performance. The error (MSE) is presented for all stages as well as the TS and HS subsets. The calibration plots illustrate an asymmetric trade-off between HS and TS error. For a given θ_{HS} value, the error ratio of the TS subset increases more than the decline in HS error. More importantly, the improvements in HS performance obtained in calibration are considerably less pronounced in the test dataset, despite a loss in TS performance”.

Table RC3-1: Bow River 2005 performance for the testing fold

Label	ce_cal	ce_tf_cal	ce_hf_cal	pi_cal	pi_tf_cal	pi_hf_cal	ce_test	ce_tf_test	ce_hf_test	pi_test	pi_tf_test	pi_hf_test
Single	0.98	0.96	0.89	0.23	0.18	0.28	0.93	0.95	0.69	0.14	0.25	0.11
RWB	0.98	0.96	0.91	0.30	0.22	0.39	0.93	0.96	0.67	0.12	0.30	0.07
RUS-RWB	0.98	0.96	0.91	0.27	0.15	0.40	0.93	0.95	0.66	0.09	0.27	0.05
ROS-RWB	0.98	0.96	0.91	0.28	0.16	0.42	0.92	0.95	0.64	0.03	0.26	-0.03
SMOTER-RWB	0.98	0.96	0.92	0.30	0.17	0.44	0.93	0.96	0.68	0.13	0.29	0.09
Bagging	0.98	0.96	0.91	0.29	0.22	0.37	0.92	0.96	0.62	-0.01	0.31	-0.09
RUS-Bagging	0.98	0.96	0.91	0.28	0.16	0.41	0.92	0.96	0.64	0.04	0.30	-0.03
ROS-Bagging	0.98	0.96	0.92	0.29	0.15	0.43	0.92	0.95	0.61	-0.04	0.27	-0.11
SMOTER-Bagging	0.98	0.96	0.92	0.28	0.16	0.42	0.93	0.96	0.65	0.07	0.29	0.02
AdaBoost	0.98	0.96	0.91	0.29	0.20	0.38	0.92	0.96	0.64	0.04	0.31	-0.02
RUS-AdaBoost	0.98	0.96	0.91	0.27	0.14	0.39	0.91	0.95	0.58	-0.08	0.28	-0.17
ROS-AdaBoost	0.98	0.95	0.91	0.25	0.10	0.41	0.93	0.95	0.66	0.08	0.26	0.03
SMOTER-AdaBoost	0.98	0.96	0.92	0.29	0.17	0.43	0.92	0.96	0.61	-0.01	0.31	-0.09
LSBoost	0.98	0.96	0.91	0.29	0.19	0.38	0.93	0.96	0.69	0.15	0.30	0.11
RUS-LSBoost	0.98	0.95	0.92	0.28	0.09	0.48	0.92	0.94	0.64	0.01	0.12	-0.02
ROS-LSBoost	0.98	0.95	0.92	0.26	0.07	0.45	0.86	0.95	0.32	-0.70	0.24	-0.93
SMOTER-LSBoost	0.98	0.95	0.91	0.25	0.10	0.40	0.93	0.96	0.67	0.12	0.29	0.07