

Supplementary Materials: A Robust Prediction Model for Species Distribution Using Bagging Ensemble with Deep Neural Networks

Jehyeok Rew, Yongjang Cho and Eenjun Hwang*

1. Our source code of pseudo-absence generation using ‘mopa’ R software package located in “pre-processing”. Our source code can be found in <https://github.com/dydwkd486/SDMsBaggingEnsemble/tree/main/preprocessing>.
2. Our source code of bootstrapping using ‘boot’ R software package located in “pre-processing”. Our source code can be found in https://github.com/dydwkd486/SDMsBaggingEnsemble/blob/main/preprocessing/boot_SDM.r.
3. Experimental codes of our proposed model and comparison models including GLM, GBM, CTA, SNN, FDA, MARS, RF, SRE, and MAXENT are located in “ensemble-DNN”. Source code of our proposed model was located in <https://github.com/dydwkd486/SDMsBaggingEnsemble/tree/main/ensemble-DNN/python/sdmdl>. Comparison codes in <https://github.com/dydwkd486/SDMsBaggingEnsemble/tree/main/ensemble-DNN/R> were used in our experiments.

Table S1. Selected environmental parameters using stepwise VIF algorithm runs (VIF < 10).

Habitats characteristic of our target species (*Hynobius leechii*, *Cyanopica cyanus*, *Platalea minor*, *Hypsipetes amaurotis* and *Hypsipetes amaurotis*) are closely related to the aquatic condition. Therefore, we used a GlobCover_14 (Water bodies) as a retained parameter until stepwise process is done.

Parameter name	Description	Scale	VIF value
Climate_03	Isothermality	30 seconds	1.556019
Climate_04	Temperature seasonality	30 seconds	5.994098
Climate_05	Mean temperature of wettest quarter	30 seconds	1.78158
Climate_12	Annual precipitation	30 seconds	6.850899
Climate_13	Precipitation of wettest month	30 seconds	6.646591
Climate_14	Precipitation of driest month	30 seconds	2.180855
GlobCover_02	Mosaic cropland (50%–70%)/vegetation (20%–50%)	300 meters	1.052725
GlobCover_03	Mosaic vegetation (50%–70%)/cropland (20%–50%)	300 meters	1.05982
GlobCover_04	Closed (>40%) broadleaved deciduous forest (>5 m)	300 meters	6.85911
GlobCover_05	Closed (>40%) needle leaved evergreen forest (>5 m)	300 meters	1.201699
GlobCover_06	Open (15%–40%) needle leaved deciduous or evergreen forest (>5 m)	300 meters	3.888618
GlobCover_07	Closed to open (>15%) mixed broadleaved and needle leaved forest (>5 m)	300 meters	1.458814
GlobCover_08	Mosaic forest or shrubland (50%–70%)/grassland (20%–50%)	300 meters	6.09256
GlobCover_09	Mosaic grassland (50%–70%)/forest or shrubland (20%–50%)	300 meters	2.447753
GlobCover_10	Closed to open (>15%) herbaceous vegetation	300 meters	1.119228
GlobCover_11	Sparse (<15%) vegetation	300 meters	4.744607
GlobCover_12	Artificial surfaces and associated areas (urban areas > 50%)	300 meters	1.097018
GlobCover_13	Bare areas	300 meters	1.232985
GlobCover_14	Water bodies	300 meters	1.545172

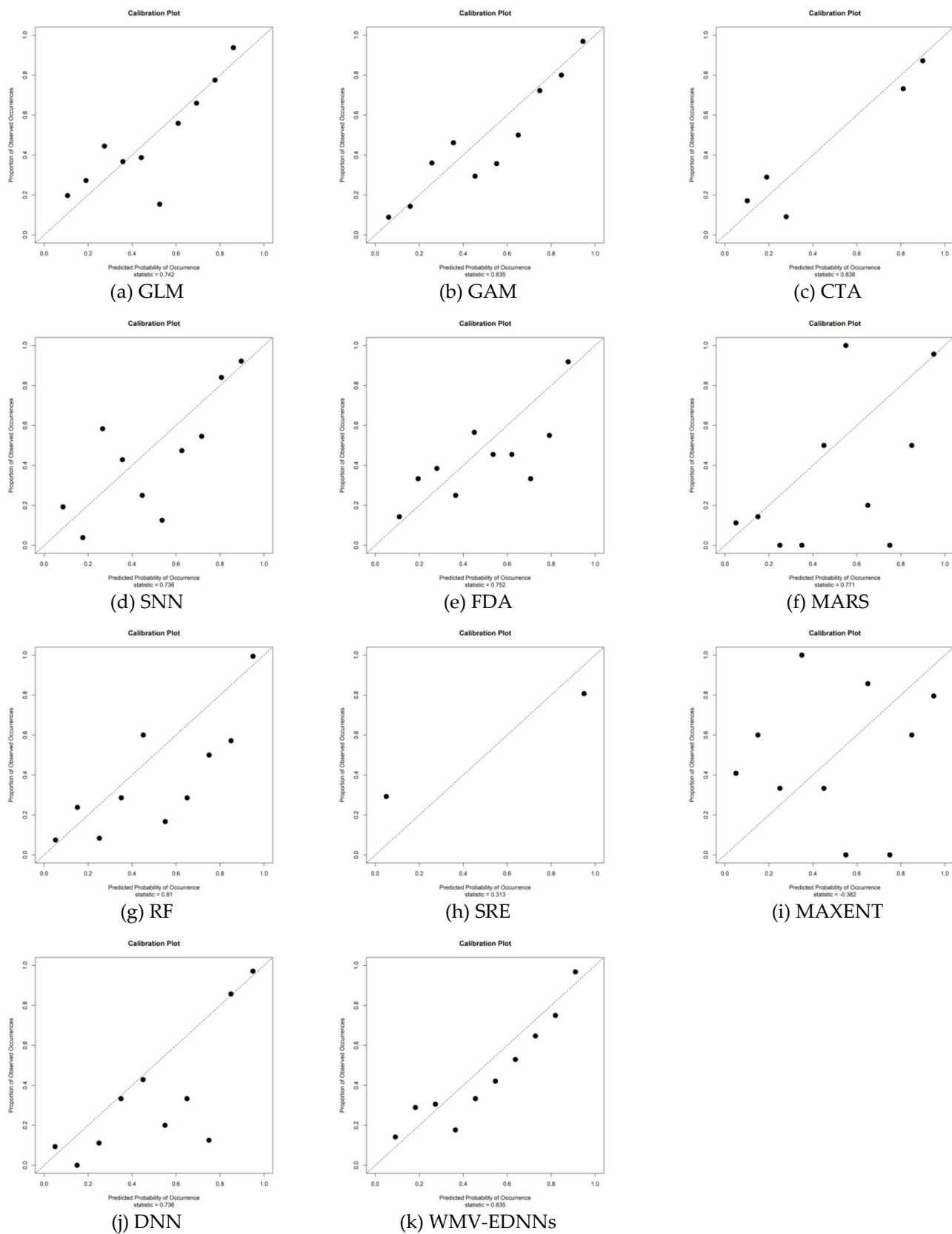


Figure S1. Calibration results for *Hynobius leechii* distribution.

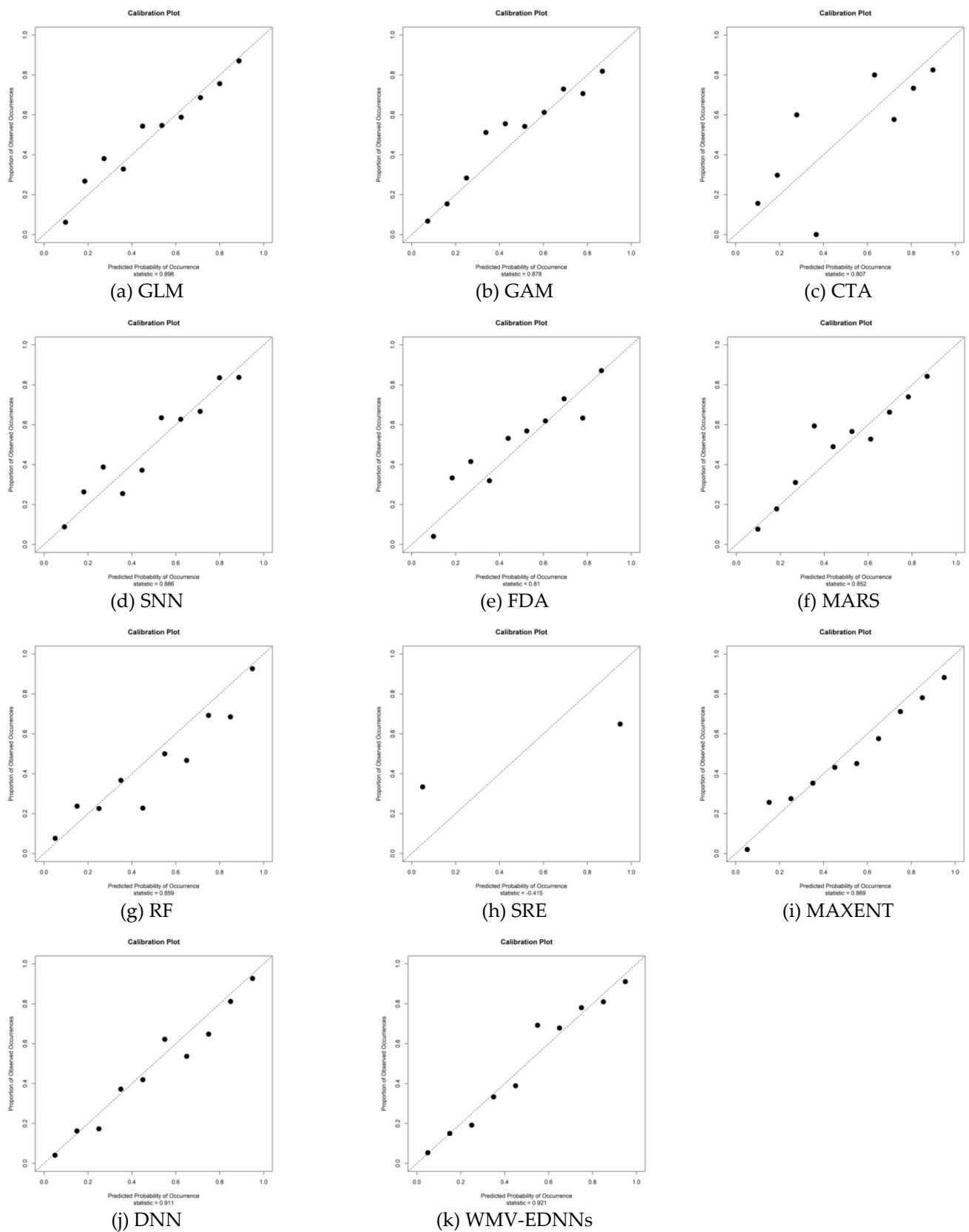


Figure S2. Calibration results for *Cyanopica cyanus* distribution.

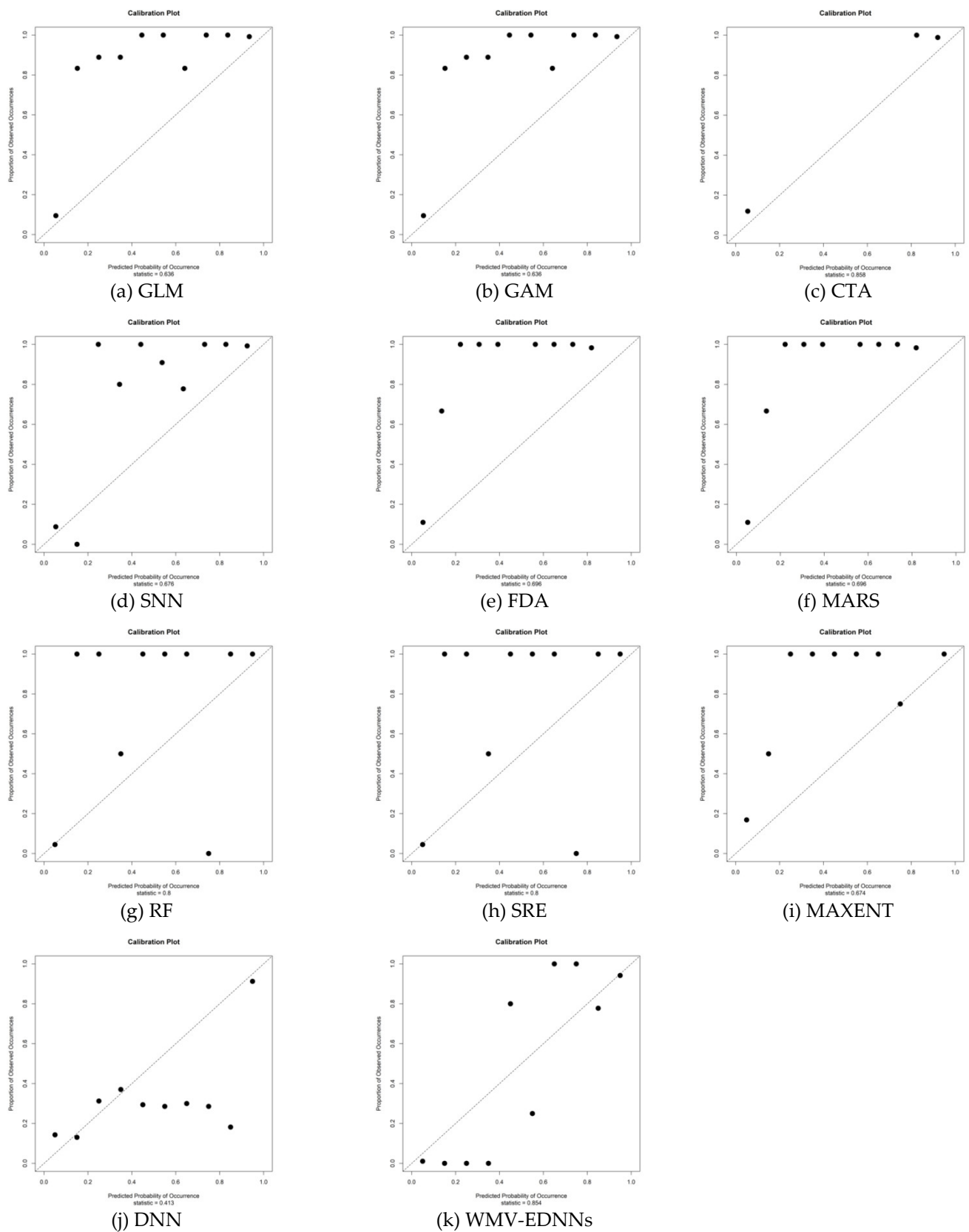


Figure S3. Calibration results for *Platalea minor* distribution.

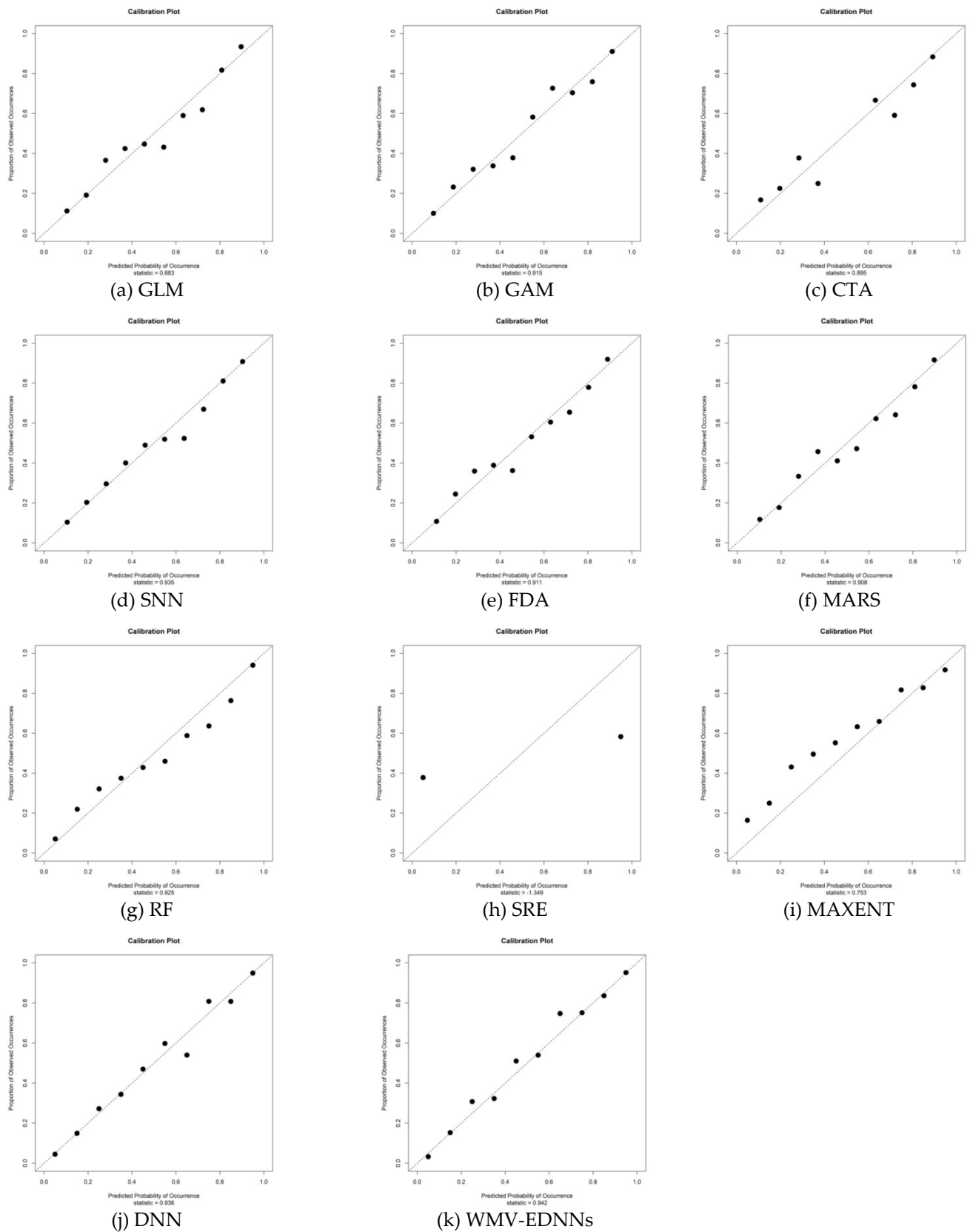


Figure S4. Calibration results for *Hypsipetes amaurotis* distribution.

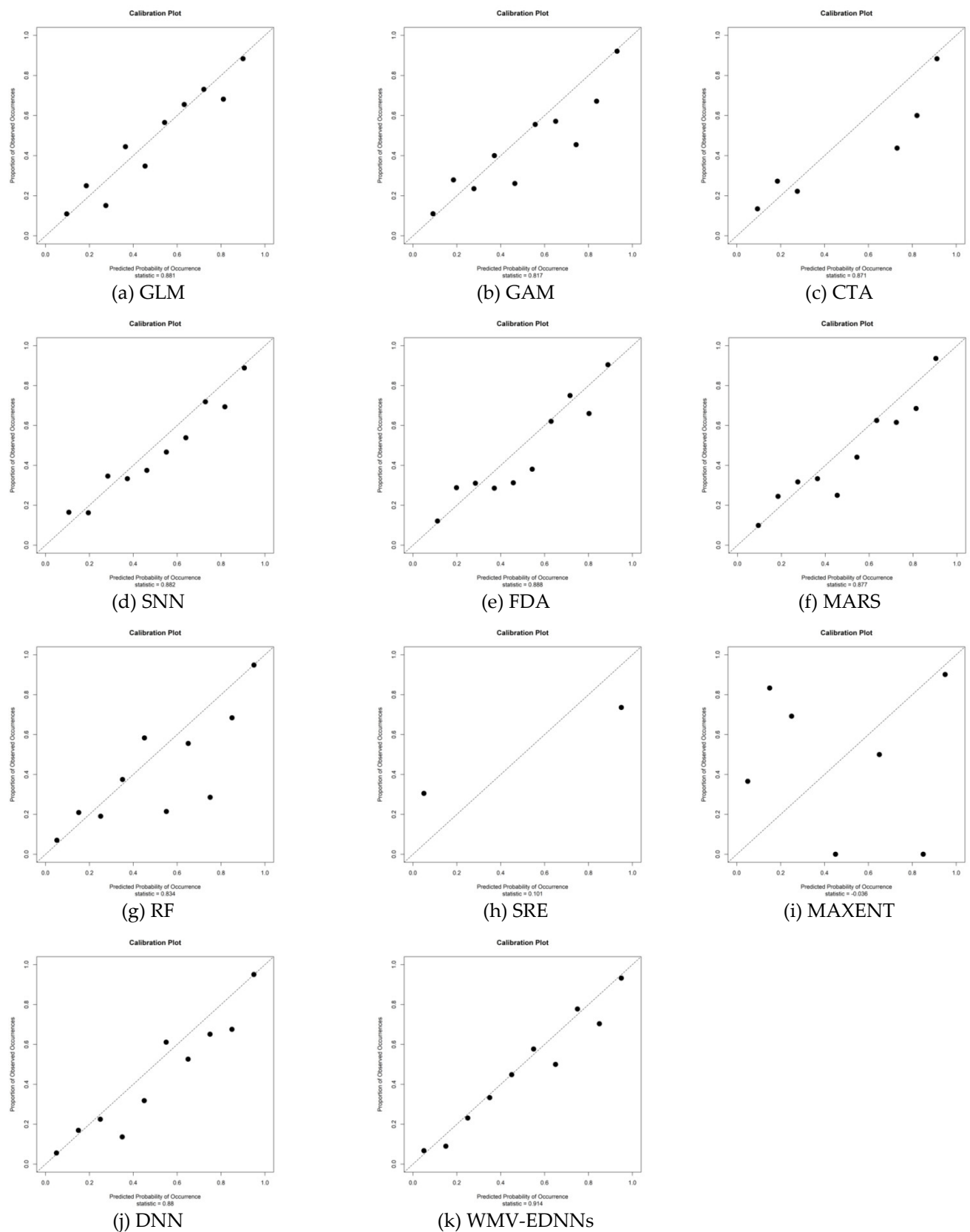


Figure S5. Calibration results for *Hyla japonica* distribution.