

Supplement

Table S0. Abbreviations

Abbreviation	Definition
ADCP	Acoustic Doppler Current Profiler
DGPS	Differential Global Positioning System
BT_Vel	Bottom-Track Velocity
CV (k-fold)	Cross-validation with k folds (here, k = 5; stratified on binned target)
SHAP	SHapley Additive exPlanations
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
LightGBM	Light Gradient Boosting Machine
CatBoost	Categorical Boosting
R ²	Coefficient of Determination (dimensionless)
MSE	Mean Squared Error ((m s ⁻¹) ²)
SNR	Signal-to-Noise Ratio
Mean_Speed	Mean water velocity

5

10

15

Table S1. Deep Learning Architecture and Input Layout

20 ANN: (batch, F); CNN: (batch, F, 1); recurrent models (LSTM/GRU/RNN): (batch, 1, F); CNN+LSTM: (batch, T, F/T) with $T \in \{1, 2\}$ (Conv1D uses kernel_size = 1, so both cases are valid).

Model	Architecture (base)	Input Shape	Optimizer / Learning Rate
ANN	Dense(64, ReLU) → Dense(32, ReLU) → Dense(16, ReLU) → Linear output	(F,)	Adam / 0.001
CNN	Conv1D(64, kernel_size = 3, ReLU) → Conv1D(32, kernel_size = 3, ReLU) → Flatten → Dense(64, ReLU) → Linear output.	(F, 1)	Adam / 0.001
LSTM	LSTM(64, return_sequences = True) → LSTM(32) → Dropout(0.2) → Dense(32, ReLU) → Linear output.	(1, F)	Adam / 0.001
RNN	SimpleRNN(64, return_sequences = True) → SimpleRNN(32) → Dense(16, ReLU) → Linear output	(1, F)	Adam / 0.0037
GRU	GRU(64, return_sequences = True) → GRU(32) → Dense(16, ReLU) → Linear output	(1, F)	Adam / 0.0037
LSTM+CNN	Conv1D(64, kernel_size = 1, ReLU) → Conv1D(32, kernel_size = 1, ReLU) → Flatten → Reshape((1, -1)) → LSTM(64, return_sequences = True) → LSTM(32) → Dropout(0.2) → Dense(32, ReLU) → Linear output	(T, F/T), $T \in \{1, 2\}$	Adam / 0.001

Abbreviations: F = 7 features; T = sequence length (ensembles per sample).

All models used MSE loss, ReLU activation, and early stopping (patience 5–7) with best weights restored.

25

30

35 **Table S2. Uncertainty analysis of cross-validation results for the laboratory dataset.**

Mean \pm standard deviation (SD) of R^2 and MSE (in $(\text{m s}^{-1})^2$) across 5 folds. MSE values are reported in $\times 10^{-6}$ unit.

Model	R^2 (mean \pm SD)	MSE (mean \pm SD) $\times 10^{-6}$
Random Forest	0.783 \pm 0.011	4.0 \pm 0.2
Gradient Boosting	0.757 \pm 0.019	4.0 \pm 0.1
LightGBM	0.743 \pm 0.014	4.0 \pm 0.6
CatBoost	0.741 \pm 0.020	5.0 \pm 0.7
XGBoost	0.731 \pm 0.025	5.0 \pm 0.8
LSTM+CNN	0.730 \pm 0.038	5.0 \pm 0.4
ANN	0.720 \pm 0.023	5.0 \pm 0.3
LSTM	0.718 \pm 0.049	5.0 \pm 0.4
CNN	0.690 \pm 0.063	5.0 \pm 0.7
RNN	0.675 \pm 0.028	6.0 \pm 0.5
GRU	0.656 \pm 0.106	6.0 \pm 1.7
Stacking Regressor	0.580 \pm 0.068	7.0 \pm 0.6

Table S3. Uncertainty analysis of cross-validation results for the field dataset.

Mean \pm standard deviation (SD) of R^2 and MSE (in $(\text{m s}^{-1})^2$) across 5 folds. MSE values are reported in $\times 10^{-4}$ units.

Model	R^2 (mean \pm SD)	MSE (mean \pm SD) $\times 10^{-4}$
Random Forest	0.603 \pm 0.044	2.28 \pm 0.32
CatBoost	0.584 \pm 0.046	2.39 \pm 0.33
LightGBM	0.573 \pm 0.041	2.45 \pm 0.29
XGBoost	0.567 \pm 0.050	2.48 \pm 0.36
Gradient Boosting	0.535 \pm 0.039	2.66 \pm 0.29
LSTM	0.529 \pm 0.033	2.57 \pm 0.04
LSTM+CNN	0.521 \pm 0.029	2.62 \pm 0.05
CNN	0.510 \pm 0.031	2.68 \pm 0.06
RNN	0.508 \pm 0.020	2.70 \pm 0.13
GRU	0.504 \pm 0.021	2.71 \pm 0.10
ANN	0.461 \pm 0.065	2.94 \pm 0.21
Stacking Regressor	0.464 \pm 0.044	3.07 \pm 0.35