Supplementary Materials

Temperature–Salinity Coupled Compensation for Conductance–Based Water Cut Measurement Using a Bayesian–Regularized Backpropagation Neural Network

This supplement provides reproducibility settings and detailed training-strategy, salinity-input, physical-plane holdout, condition-wise validation, and pointwise STM32 prediction results for the main manuscript.

Supplementary Methods

Supplementary Methods S1. Trilinear augmentation and normalization

The measured calibration tensor had dimensions 3 × 6 × 11, with T = [25, 45, 65] °C, S = [0, 5000, 10000, 20000, 50000, 80000] mg/L, and Yw = 0:0.1:1. MATLAB griddedInterpolant with linear interpolation along each axis was used to construct V = Ftri(T,S,Yw). The derived grid was generated with ndgrid(25:5:65, linspace(0,80000,20), linspace(0,1,50)), producing 9 × 20 × 50 = 9000 points entirely within the calibration bounds. The 40 measured evaluation inputs were not used in this construction.

All learned models used the same externally specified min–max transformation: t = (T−25)/40, s = S/80000, and v = (V−0.241)/4.234, corresponding to T ∈ [25,65] °C, S ∈ [0,80000] mg/L, and V ∈ [0.241,4.475] V. The input order was $[t,s,v]^T$; Yw remained a fraction in [0,1], and predictions were clamped to this interval before errors were reported.

Supplementary Methods S2. MPR nonlinear feature basis and ridge coefficient selection

MPR denotes the ridge-regularized nonlinear basis-expansion baseline used in the main manuscript. The MPR feature vector contains 27 terms in this order: 1; t, s, v; t^2, s^2, v^2, ts, tv, sv; t^3, s^3, v^3, t^2s, t^2v, s^2t, s^2v, v^2t, v^2s, tsv; v^4, v^5; 1/(v+0.05), exp(−3v), sqrt(max(s,0)), t/(v+0.05), and s/(v+0.05). All 27 coefficients, including the intercept, were subject to ridge regularization using an identity penalty matrix. Thirty logarithmically spaced λ candidates from 10^{-8} to 10^{1} were evaluated on an 80/20 split of the 9000 derived samples using validation RMSE. The split used MATLAB rng(8100); the selected λ was 10^{-8}, after which the model was refitted on all 9000 samples.

Supplementary Methods S3. SVR search settings

The ε-SVR used a Gaussian RBF kernel. The search grid was C ∈ {1, 10, 100}, ε ∈ {0.005, 0.01, 0.02, 0.05}, and kernel scale ℓ ∈ {0.1, 0.3, 1}. Parameters were selected by validation RMSE on one 80/20 split of the derived data generated with MATLAB rng(9100). The selected settings were C=100, ε=0.01, and ℓ=0.1. The final SVR was refitted on all 9000 derived samples using fitrsvm with Standardize=false because common normalization had already been applied.

Supplementary Methods S4. BR-BPNN architecture and training record

Candidate topologies were 3-5-1, 3-10-1, 3-8-4-1, and 3-10-5-1. Architecture screening used five random 80/20 partitions of the derived data with partition seeds 5001–5005 and initialization rng(10000+100*r+a), where r denotes the partition and a the candidate architecture. The final network was specified as fitnet([10 5],'trainbr'), initialized with rng(42), trained on all 9000 derived samples, and used divideFcn='dividetrain'. Both hidden layers

used tansig and the output layer used purelin. The normalized input order was $[t,s,v]^T$. The network has 101 trainable parameters and requires 85 weight multiplications per forward inference.

The recorded software environment was MATLAB R2024b Update 2 (24.2.0.2773142) with Deep Learning Toolbox 24.2. The effective trainbr settings were epochs=1000, goal=0, min_grad=1×10^{-7}, mu=0.005, mu_dec=0.1, mu_inc=10, mu_max=1×10^{10}, max_fail=0, and time=Inf. Training reached the preset 1000-epoch limit. The saved record reported γeff=91.23 and ssX=1.295×10^{4}. The final αBR and βBR values were not provided directly by the training record and were not reconstructed.

Supplementary Methods S5. Interpolation validation and direct trilinear inversion

At the 40 measured test conditions, forward interpolation yielded voltage MAE=0.068757 V, RMSE=0.081119 V, R^2=0.99481, and maximum absolute residual=0.17805 V. These metrics evaluate the forward voltage surface. For physical-plane holdouts, the selected plane was removed before rebuilding the response surface; its measured points were used only for evaluation. Each strategy used two training initializations per plane. Table S3 reports the mean plane-wise MAE. The 70% water-cut plane is the exception to the general benefit of interpolation. The 45 °C holdout has mean MAE 2.522 pp and a mean maximum absolute error of 29.147 pp, consistent with amplification of voltage differences in weakly sensitive inverse-response regions.

Direct trilinear inversion uses the full 198-point calibration tensor and the same linear forward interpolant Ftri(T,S,Yw) defined in Methods S1. It does not fit regression parameters to the 9000 derived samples. For each measured temperature–salinity pair, nodal voltages Vk are evaluated at water-cut fractions yk = 0.50, 0.60, 0.70, 0.80, 0.90, and 1.00. At a fixed water-cut node, this is bilinear interpolation over the enclosing temperature–salinity cell. The benchmark test temperatures and salinities lie inside the measured calibration bounds.

The six water-cut nodes are scanned in ascending order. For each consecutive pair, the inverse candidate is calculated as

$$y^* = y_k + \frac{V - V_k}{V_{k+1} - V_k}(y_{k+1} - y_k)$$

Here V is the measured mean sensor peak voltage, Vk = Ftri(T,S,yk), and Vk+1 = Ftri(T,S,yk+1), all in volts; yk and yk+1 are water-cut fractions. A segment is skipped when $|Vk+1-Vk| < 10^{-12}$ V. A candidate is admissible when $yk-10^{-12} \le y^* \le yk+1+10^{-12}$; it is then clamped to that segment before its forward residual |Ftri(T,S,y*)−V| is evaluated.

If several candidates are admissible, the one with the smallest forward residual is returned. An exact residual tie is resolved by the first candidate encountered in the ascending segment scan, consistent with MATLAB min. If no candidate is admissible, the algorithm returns the water-cut node whose nodal voltage is closest to the measured voltage; a nodal tie likewise selects the first node. The final estimate is clamped to [0.50,1.00]. Thus, a voltage outside the selected branch response is mapped to a nearest branch node rather than extrapolated along an inverse segment.

All 40 evaluated voltage–temperature–salinity inputs yielded one admissible candidate each; the flat-segment and no-candidate fallback rules were not activated for these inputs. Recalculation with the retained calibration and test voltages reproduced R^2=0.97026, RMSE=2.4388 pp, MAE=1.7974 pp, and maximum absolute error 5.5991 pp, matching main-text Table 1 to the reported precision.

Supplementary Tables

Table S1. Comparison of measured-only, repeated-sample, and interpolated-sample training on the same 40 measured test conditions (mean ± SD across three seeds).

Training strategy	Seeds	MAE (pp)	RMSE (pp)
A: Measured (198)	3	5.077 ± 1.117	8.326 ± 1.984
B: Repeated (9000)	3	8.120 ± 3.314	12.116 ± 6.471
C: Interpolated (9000)	3	1.566 ± 0.076	2.033 ± 0.094

Note: Errors are in water-cut percentage points (pp). Values are means ± standard deviations across three random initializations. This retrospective training-strategy comparison reuses the 40 original measured test conditions.

Table S2. Offline sensitivity of BR-BPNN to relative errors in externally supplied salinity.

Relative salinity error	MAE (pp)	RMSE (pp)	AEmax (pp)
−10%	1.7084	2.3610	6.0147
−5%	1.6066	2.1504	4.8737
−1%	1.7001	2.1425	4.7006
0%	1.7233	2.1652	4.8488
+1%	1.7458	2.1973	4.9915
+5%	1.8579	2.4111	5.5118
+10%	2.0116	2.8203	6.8393

Note: Salinity was perturbed numerically while measured voltage and temperature were held fixed. Mean absolute prediction shift is the mean across 40 test conditions of the absolute difference from the unperturbed prediction. For each perturbation magnitude, the mean shifts for the positive and negative perturbations were averaged; the resulting values for 1%, 5%, and 10% were 0.1093, 0.5453, and 1.0824 pp, respectively. Errors and shifts are in percentage points (pp).

Table S3. Mean absolute water-cut error (pp) on eight held-out physical calibration planes. Each result is averaged over two training initializations. See main-text Section 4.3.

Held-out measured plane	A: Measured	B: Repeated	C: Interpolated
45 °C	17.708	23.839	2.522
5,000 mg/L	12.367	8.448	2.902
10,000 mg/L	8.384	5.679	2.430
20,000 mg/L	6.454	7.009	2.663
50,000 mg/L	9.720	20.299	1.104
60% water cut	29.662	13.125	3.910
70% water cut	4.891	6.576	7.878
80% water cut	3.751	11.354	2.463

Note: A uses the remaining measured calibration samples, B repeats those samples to 9000 training entries, and C samples the rebuilt trilinear response surface. The omitted plane is excluded from its own response-surface construction. These folds reuse calibration measurements and complement the separate 40-condition blind test.

Table S4. Offline BR-BPNN error by temperature, salinity, and reference nominal water cut. Bias is predicted minus reference water cut; all error values are in pp. See main-text Section 4.4.1.

Factor	Condition	*n*	MAE (pp)	RMSE (pp)	Bias (pp)
T	35 °C	20	1.561	2.065	−0.659
	55 °C	20	1.885	2.261	+0.777
S	3,500 mg/L	10	1.825	2.264	+0.468
	6,500 mg/L	10	2.345	2.779	+0.340
	25,000 mg/L	10	1.230	1.681	+0.527
	65,000 mg/L	10	1.494	1.754	−1.100
Y_w	55%	8	1.162	1.384	+1.162
	65%	8	0.294	0.433	−0.192
	75%	8	2.815	3.024	−1.478
	85%	8	2.885	3.118	+0.750
	95%	8	1.461	1.573	+0.053

Note: Each temperature group contains 20 samples, each salinity group 10 samples, and each water-cut group 8 samples. Values describe the 40-condition test matrix; these are offline BR-BPNN results, distinct from the reacquired STM32 measurements.

Table S5. STM32-predicted nominal water cut (%) for the 40 test conditions. Column headings denote the reference nominal water cuts. See main-text Section 4.5 and Figure 8.

T (°C)	S (mg/L)	55%	65%	75%	85%	95%
35	3,500	56.20	67.20	69.90	90.24	94.70
35	6,500	52.20	61.00	77.80	82.80	94.60
35	25,000	60.10	66.00	73.80	82.50	95.40
35	65,000	54.10	68.60	71.40	85.90	94.50
55	3,500	53.50	63.70	79.00	84.00	95.50
55	6,500	57.50	65.80	74.40	86.30	94.30
55	25,000	49.76	62.50	76.50	87.50	95.70
55	65,000	55.60	65.30	74.78	84.20	95.22

Sample order is rowwise: five water-cut levels per temperature–salinity pair. Entries are predictions in percent; errors are calculated relative to the column headings and reported in percentage points.