

Supplementary Information for  
“Seven reanalyses, seven answers: ocean product  
choice can flip tropical-cyclone intensity  
conclusions”

Ze Zhang<sup>1</sup>, Letian Chen<sup>1</sup>, Bin Cheng<sup>1\*</sup>, Senliang Bao<sup>1</sup>,  
Huizan Wang<sup>1</sup>, Weimin Zhang<sup>1\*</sup>, Xiaohui Zhou<sup>1</sup>

<sup>1</sup>College of Meteorology and Oceanography, National University of  
Defense Technology, Changsha, China.

\*Corresponding author(s). E-mail(s): [chengbin13@nudt.edu.cn](mailto:chengbin13@nudt.edu.cn);  
[weiminzhang@nudt.edu.cn](mailto:weiminzhang@nudt.edu.cn);

**Abstract**

This supplement documents the exchange-coefficient choice and the sensitivity tests at greater length than the main text, together with the storm-characteristics and cooling-trend analyses, the atmospheric-profile-lead and D26 computations, Supplementary Figures S1–S6, and Supplementary Tables S1–S4 (the per-basin selection matrix, split over S1 and S2 to fit the page, the D26 bias table S3, and the trend-test multiplicity table S4). Figures S1–S6 and Tables S1–S4 are cited in the main text as “Supplementary Fig. 1–6” and “Supplementary Tables 1–4”. Passages that duplicated main-text results were removed in revision; the change-point, autocorrelation, multiplicity-control, 1°-coarsening, and monthly-averaging tests added in revision are documented here in full.

**Supplementary methods**

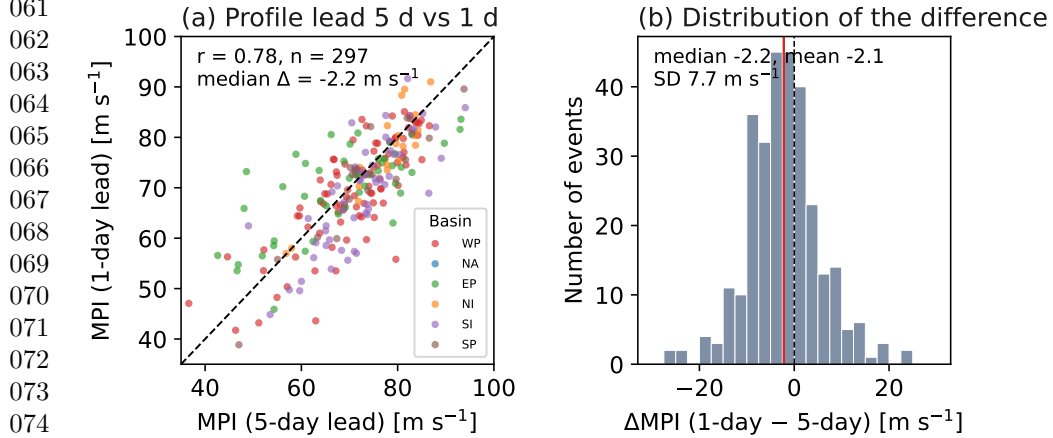
**Coupled and fixed-environment sensitivity bounds**

The event-level coupled empirical slope is  $7.0 \text{ m s}^{-1} \text{ }^{\circ}\text{C}^{-1}$  (GLORYS, 25–30 °C,  $n = 2,244$  events, convex as in [1]); when SST and the atmospheric profile co-vary, the literature value is  $\sim 4\text{--}6 \text{ m s}^{-1} \text{ }^{\circ}\text{C}^{-1}$  [2]. Together with the fixed-environment sensitivity

defined in the main text ( $10.8 \text{ m s}^{-1} \text{ }^{\circ}\text{C}^{-1}$ ), these bound the propagation of SST error into MPI.

## Atmospheric-profile lead

The pre-TC sampling convention pairs the atmospheric profile at the pre-TC instant  $t_{\text{pre}}$  with the five-day-earlier background SST—one profile, one ocean state, one instant. Against this operational choice, MPI computed with 1-day-lead profiles on the same background SST is highly correlated ( $r = 0.78$ ) and slightly lower (median difference  $-2.2 \text{ m s}^{-1}$ , 1-day-lead minus 5-day-lead; mean  $-2.1$ , SD  $7.7 \text{ m s}^{-1}$ ;  $n = 297$  events, GLORYS background SST; Fig. S1). Pairing a storm-time profile with the five-day-earlier SST collapses CAPE and yields  $\text{MPI} < 20 \text{ m s}^{-1}$  for 156 instants with  $\text{SST} > 26 \text{ }^{\circ}\text{C}$ , rising to  $32\text{--}72 \text{ m s}^{-1}$  once the pre-TC profile is used.



**Fig. S1** Sensitivity of MPI to the atmospheric-profile lead (GLORYS background SST,  $n = 297$  events). (a) MPI computed with 5-day-lead ERA5 profiles (the operational choice) versus 1-day-lead profiles on the same background SST; (b) distribution of the difference (1-day-lead minus 5-day-lead).

## D26 computation

Product D26 is computed from each product’s own temperature profile at the pre-TC instant (peak time minus five days), with the same bounded linear completion rule for profiles ending above 100 m as used for HYCOM in the main analysis (slope lower bound  $-0.10 \text{ }^{\circ}\text{C m}^{-1}$ , completion cap 200 m, isotherm-artifact guard); BOA D26 is read from the monthly  $1^{\circ}$  grid at the nearest point (Table S3).

## Exchange-coefficient ratio $C_k/C_d$

The exchange-coefficient ratio  $C_k/C_d$  is set to 0.9, the default of the Bister and Emanuel [1] algorithm, inside the 0.75–1.5 observational range of Emanuel [3] and consistent with the order-unity prediction of the similarity theory of Emanuel [4]; this

**Table S3** D26 bias against BOA-Argo by product and basin (m; positive = product D26 too deep). Global: all basins pooled;  $n$  = events with a valid D26 in both fields.

| Product     | $n$  | Global | WP    | NA   | EP    | NI   | SI    | SP    |
|-------------|------|--------|-------|------|-------|------|-------|-------|
| GLORYS      | 993  | -1.0   | +12.9 | -0.4 | -15.7 | -0.9 | -7.3  | +2.9  |
| GREP        | 1001 | -4.6   | +9.6  | -6.6 | -20.7 | +1.2 | -9.8  | +1.0  |
| C-GLORS     | 1002 | -5.5   | +12.1 | -5.9 | -24.3 | -3.9 | -10.5 | -2.4  |
| FOAM        | 995  | -3.0   | +9.9  | -4.3 | -17.3 | -2.5 | -5.3  | -1.7  |
| GLORYS-GREP | 989  | -3.0   | +10.5 | -3.9 | -18.5 | +0.4 | -8.7  | +1.5  |
| ORAS5       | 992  | -1.6   | +13.5 | -5.2 | -18.0 | +2.2 | -7.2  | +7.4  |
| HYCOM       | 1047 | +4.4   | +19.3 | -8.5 | -11.5 | +7.3 | +6.3  | +21.2 |

is the standard choice in PI-variability studies [5, 6]. Because  $V^2$  is linear in  $C_k/C_d$ , all products scale by the same factor and inter-product differences, bias propagation, and trends are insensitive to this value, though absolute magnitudes are not: a  $\pm 0.1$  perturbation is  $\pm 5.5\%$ , and the CBLAST high-wind measurement (mean  $\approx 0.4$ ; 7) would lower absolute MPI by about one third.

### Sensitivity tests

Sensitivity tests repeat key results (i) without the  $|\text{bias}| \leq 3^\circ\text{C}$  screen, (ii) excluding the 2015 El Niño year, (iii) with  $\pm 0.1$  perturbations to  $C_k/C_d$  ( $\Delta\text{DPI} \approx \pm 3.8 \text{ m s}^{-1}$ ), (iv) with the regime boxes shifted by  $\pm 5^\circ$ , and (v) with the DPI mixing length recomputed from a fixed environment wind (40, 50 and  $60 \text{ m s}^{-1}$ ) in place of the storm’s own intensity wind. Test (v) isolates a structural component of the ratio  $\alpha = V_{\text{max}}/\text{DPI}$ : because  $u_*$  in Eq. (3) follows the best-track intensity wind, DPI and the numerator of  $\alpha$  share  $V_{\text{max}}$ . Holding the wind at  $50 \text{ m s}^{-1}$  lowers the category-5 event-median  $\alpha$  against DPI from 0.93 to 0.87 (0.86–0.89 across the three winds;  $n = 152$  events, GLORYS) and GLORYS’s EP median from 1.03 to 0.92 (0.88–0.96;  $n = 229$ ), using the linear-profile relation for  $T_{dy}$  and the product’s dMPI/dSST sensitivity: the reading that category-5 storms nearly attain their DPI survives, while part of the EP excess above 1 is structural. Their outcomes: the  $|\text{bias}| \leq 3^\circ\text{C}$  screen moves the global event-level mean by  $0.01^\circ\text{C}$  (the global event-level GLORYS mean is  $+0.102^\circ\text{C}$  without and  $+0.092^\circ\text{C}$  with the screen); removing the 2015 El Niño leaves the bias tiers, the TCHP-cooling correlations, and the global DPI and mixing-correction trends intact ( $p = 0.018$  and  $0.040$ ) and weakens only the marginal signals reported in the Discussion of the main text; shifting the regime boxes by  $\pm 5^\circ$  leaves the warm-pool/CTIZ contrast; and perturbing  $C_k/C_d$  by  $\pm 0.1$  scales absolute MPI by  $\pm 5.5\%$  without touching any inter-product conclusion.

### Representativeness of the Argo reference: $1^\circ$ -coarsening and monthly-averaging tests

*Why these tests.* The category-5 TCHP deficits of the main text are measured against a monthly  $1^\circ$  gridded reference. The main-text decomposition separates the all-product

139 common deficit from the product-to-product spread, but it cannot, by construction,  
 140 attribute the common part: any error of the reference itself enters every product’s bias  
 141 identically. The tests below measure how much of the common category-5 deficit the  
 142 reference’s spatial and temporal smoothing can account for.

143 *Design.* For every event matched by Argo (the same event template as the main  
 144 analysis), the pre-TC temperature profile is extracted from the same window of the  
 145 native-resolution fields in three ways—at the nearest valid native column (the opera-  
 146 tional choice), as the  $1^\circ \times 1^\circ$  cell mean of all valid columns around the storm, and as  
 147 the calendar-month mean of that cell—and TCHP is computed from each profile with  
 148 the same algorithm, including the same bounded linear completion beyond 100 m.  
 149 The reference enters all comparisons identically, so the paired native-to-coarse shift of  
 150 the bias isolates the spatial-smoothing term and the native-to-monthly shift adds the  
 151 temporal one. The six products of the CMEMS archives are tested; HYCOM follows  
 152 its own full-depth pipeline and is excluded (its category-5 bias is not significant in the  
 153 main analysis). The nearest-column extraction differs slightly from the main analy-  
 154 sis’s caliber (validation against the main pipeline’s event TCHP gives  $r = 0.78$  with  
 155 a  $-0.6 \text{ kJ cm}^{-2}$  mean offset), so the bias levels below are not directly comparable to  
 156 Fig. 7 of the main text; the paired shifts are the quantities of interest. The monthly  
 157 arm (all valid daily files of the month, 28–31 days) reproduces the joint  $1^\circ$ -monthly  
 158 scale of the reference itself; because the monthly mean also retains the storm’s own  
 159 cold wake, which the sparse Argo sampling behind the reference captures only in part,  
 160 this arm measures the bias at the reference’s own scale rather than isolating a pure  
 161 timescale term. Across all matched events the three arms complete for 6,431, 6,434  
 162 and 6,380 event–product extractions respectively.

163 *Result.* Applied to the 528 category-5 event–product extractions (88 events  $\times$   
 164 6 products per arm, complete in all three arms), coarsening shifts the all-product  
 165 category-5 bias only from  $-2.9$  to  $-3.3 \text{ kJ cm}^{-2}$ , and the per-product ordering is pre-  
 166 served (C-GLORS remains the most negative at  $-6.3 \text{ kJ cm}^{-2}$ , GLORYS and ORAS5  
 167 remain near zero). The same coarsening narrows the category-5 product spread from  
 168  $7.2$  to  $5.5 \text{ kJ cm}^{-2}$ . Monthly averaging (the same  $n = 88$  category-5 events per product)  
 169 moves the all-product category-5 bias further, to  $-9.2 \text{ kJ cm}^{-2}$ , and again preserves  
 170 the ordering: C-GLORS is the most negative product in every arm ( $-11.0 \text{ kJ cm}^{-2}$   
 171 at the monthly scale) and GLORYS and ORAS5 the least negative ( $-7.1$  and  $-6.7$ ),  
 172 while the product spread narrows further, to  $4.3 \text{ kJ cm}^{-2}$ .

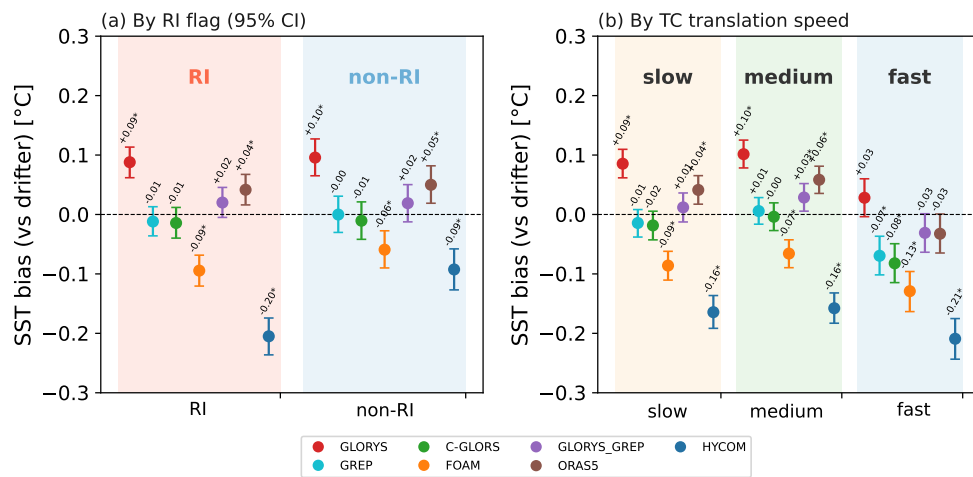
173 *Consequence.* Three conclusions follow. First, neither of the reference’s two limita-  
 174 tions can create the common deficit: the  $1^\circ$  grid removes at most  $\approx 0.4 \text{ kJ cm}^{-2}$  of the  
 175  $\approx 2.9 \text{ kJ cm}^{-2}$  common deficit, and the monthly timescale deepens it to  $-9.2 \text{ kJ cm}^{-2}$ ,  
 176 so the category-5 deficits reported in the main text are conservative—at the refer-  
 177 ence’s own scales they are larger, not smaller. Second, the category-5 reversal is not a  
 178 resolution- or timescale-mismatch artifact: it survives degrading the products to the  
 179 reference’s scale on both axes, with C-GLORS and GREP among the most negative  
 180 products in every arm and GLORYS and ORAS5 among the least negative. Third,  
 181 the product-to-product spread at category 5 narrows but persists at the reference’s  
 182 scale ( $4.3 \text{ kJ cm}^{-2}$ ), confirming that the differential is genuine information the coarse  
 183 monthly reference cannot hold.

184

## Supplementary analyses

### Bias stratified by storm characteristics

Stratifying the bias by storm characteristics shows that the structure is a property of the product, not of the storm (Fig. S2). Rapidly intensifying (RI;  $\geq 30$  kt per 24 h within an 18–30 h window, 8) and non-RI positions carry nearly identical biases in most products (differences of at most  $0.04^\circ\text{C}$ ,  $0.11^\circ\text{C}$  for HYCOM, whose RI sample is colder at  $-0.205^\circ\text{C}$ ). Translation speed modulates the magnitude in a consistent direction—fast-translating storms carry the coldest bias in all seven products, by  $0.04$ – $0.07^\circ\text{C}$  relative to slow storms—but never changes any product’s sign, except GLORYS, whose fast-storm bias ( $+0.028^\circ\text{C}$ ) is only marginally positive. A study that focuses on RI or on slow storms therefore inherits the same reanalysis bias as any other study, slightly shifted in magnitude; the bias cannot be removed by sample choice.



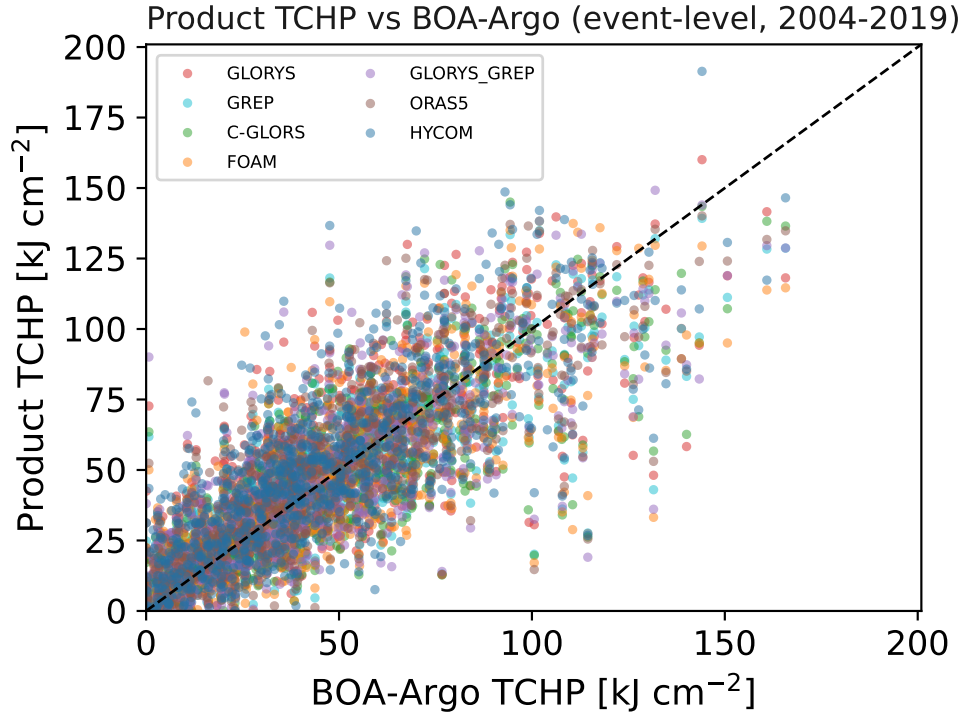
Shaded bands separate RI/non-RI and slow/medium/fast categories.  
Numbers show event-level mean bias ( $\pm$  95% block bootstrap CI); \* = significant.  
RI = rapidly intensifying ( $\geq 30$  kt per 24 h within 18–30 h).

**Fig. S2** SST bias stratified by storm characteristics (event-level means  $\pm$  95% confidence intervals). (a) Rapidly intensifying versus non-rapidly-intensifying positions; (b) translation-speed categories (slow  $< 5 \text{ m s}^{-1}$ , medium  $5$ – $8 \text{ m s}^{-1}$ , fast  $> 8 \text{ m s}^{-1}$ ). \* = CI excludes zero.

The companion TCHP cross-validation scatter is shown in Fig. S3.

### Cooling trends split by basin

Cooling trends split by basin in the same way, their spatial structure set by thermal preconditions (Fig. S4). GLORYS cooling deepens in WP ( $+0.024^\circ\text{C decade}^{-1}$ ,  $p = 0.020$ ) and NI ( $+0.061$ ,  $p = 0.002$ ) and weakens in EP ( $-0.025$ ,  $p = 0.046$ ); the basin trend correlates with basin-mean TCHP at  $r = 0.79$  (with mean MLD  $0.30$ , with



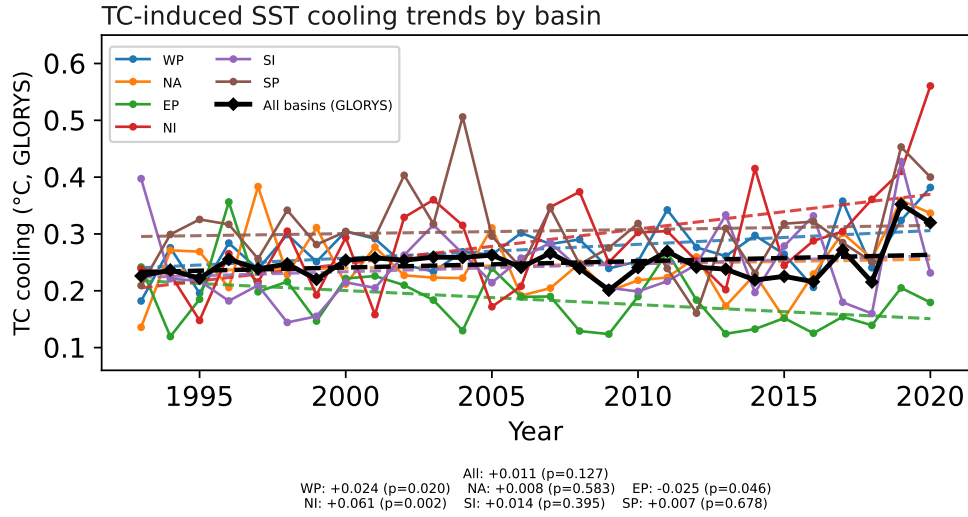
**Fig. S3** Product TCHP against BOA-Argo (event-level, 2004–2019), requiring positive TCHP in both fields;  $n = 1,028$ – $1,044$  events per product.

mean intensity  $-0.54$ ), so basins with a deep warm layer cool more and the cold-tongue corridor cools less. The absolute magnitude, however, is product-dependent: HYCOM’s global cooling trend is  $+0.03$   $^{\circ}\text{C}$   $\text{decade}^{-1}$  ( $p = 0.089$ ) against  $+0.001$   $^{\circ}\text{C}$   $\text{decade}^{-1}$  for GLORYS (n.s.)—reported cooling trends must name their product.

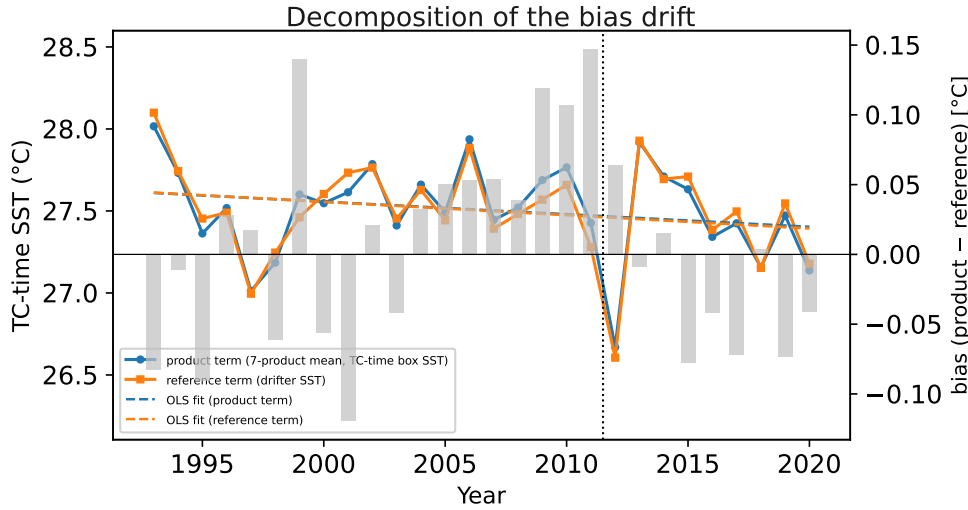
Fig. S5 shows the decomposition of the pooled SST-bias drift into a product term and a reference term used in the main text’s non-stationarity analysis.

### MPI trend decomposition: SST term versus the atmospheric profile term

The MPI trend decomposes into an SST term (the fixed-environment sensitivity times the product’s SST trend) and a residual atmospheric-profile term. Across products the SST term spans  $-0.1$  to  $+1.0$   $\text{m s}^{-1}$   $\text{decade}^{-1}$  (FOAM  $+1.0$ , HYCOM  $-0.1$ ), while the profile term is nearly identical at  $-0.5$   $\text{m s}^{-1}$   $\text{decade}^{-1}$  (range  $-0.47$  to  $-0.53$ ), because all products share a single atmospheric reanalysis; it is therefore not an independent observational constraint, and a multi-reanalysis test with MERRA-2 or JRA-55 remains future work. The basin-resolved DPI trends and their 2015 sensitivity are discussed in the main text.



**Fig. S4** TC-induced SST cooling trends by basin (GLORYS, all-fix annual means; dashed: ordinary least-squares fits).



**Fig. S5** Decomposition of the pooled SST-bias drift into product and reference terms. Annual means of the seven-product TC-time box SST (product term) and the drifter SST (reference term), with the bias as the small residual (right axis); dashed lines are OLS fits; the vertical line marks the 2012 split.

## Homogeneity of the HYCOM intensity series

The HYCOM DPI decline reported in the main text (the only significant reanalysis intensity trend) and the mirror-image rise of its mixing correction are steps, not gradual trends. A Pettitt change-point test [9] on the 1993–2020 annual means locates a

single break in 2010 for the global series ( $p = 0.004$ ): the mean is  $64.8 \text{ m s}^{-1}$  through 2010 and  $62.5 \text{ m s}^{-1}$  from 2012 onward—the 2011 annual mean already sits at the post-step level, so the exact break year within 2010–2012 is ambiguous—and no sub-period shows a significant trend of its own (1993–2012:  $-0.2 \text{ m s}^{-1} \text{ decade}^{-1}$ ,  $p = 0.73$ ; 2013–2020:  $-0.9$ ,  $p = 0.69$ ); the South Indian sector repeats the pattern (change point 2010,  $p = 0.039$ ). The step coincides with the end of the GLBu0.08 reanalysis-era experiments and the product’s transition to the GOFS 3.1 analysis in late 2012; the missing 1993 records, filled from the product’s full archive, add a further inhomogeneity at the start of the series. The ordinary-least-squares trend of the mixing correction ( $+0.34 \text{ m s}^{-1} \text{ decade}^{-1}$ ,  $p = 0.032$ ) is likewise marginal: its residual lag-1 autocorrelation is 0.34, and an AR(1)-adjusted test with Quenouille effective degrees of freedom [10] gives  $p = 0.137$ . The global DPI trend itself remains significant under the same adjustment ( $\rho = 0.10$ ,  $p = 0.025$ ). We therefore read the HYCOM decline as a production-era inhomogeneity marker rather than evidence of a physical decline, as stated in the main text.

### Multiple-comparison control of trend tests

The trend narrative of the main text rests on one claim—that a single reanalysis intensity trend is significant—so the trend tests reported in the main text are placed under familywise control. Two families are defined by what they feed: the intensity family (MPI, DPI and MPI – DPI for the seven products, the drifter-anchored MPI, the HYCOM basin DPI and NA MPI trends, and the observed category-5 LMI trend; 28 tests) and the environment family (per-product SST-bias and TCHP trends with the Argo TCHP trend, the basin cooling trends and the HYCOM global cooling trend, the WP GLORYS SST trend and the category-4 Argo TCHP trend; 24 tests). Benjamini–Hochberg false-discovery rates are computed within each family (Table S4).

No test survives at  $q < 0.05$  in either family (0/28 and 0/24), so every marginal signal in the main text—the FOAM and HYCOM SST-bias drifts, the WP GLORYS warming, the category-5 LMI rise, and the basin cooling trends—is nominal. The change-point test behind the HYCOM DPI step is a different procedure, and it is robust to the same concern: its  $p = 0.004$  remains significant under a Bonferroni correction across the seven products ( $p \times 7 = 0.028$ ), which is why the main text reads the decline as a production-era step rather than as a significant physical trend.

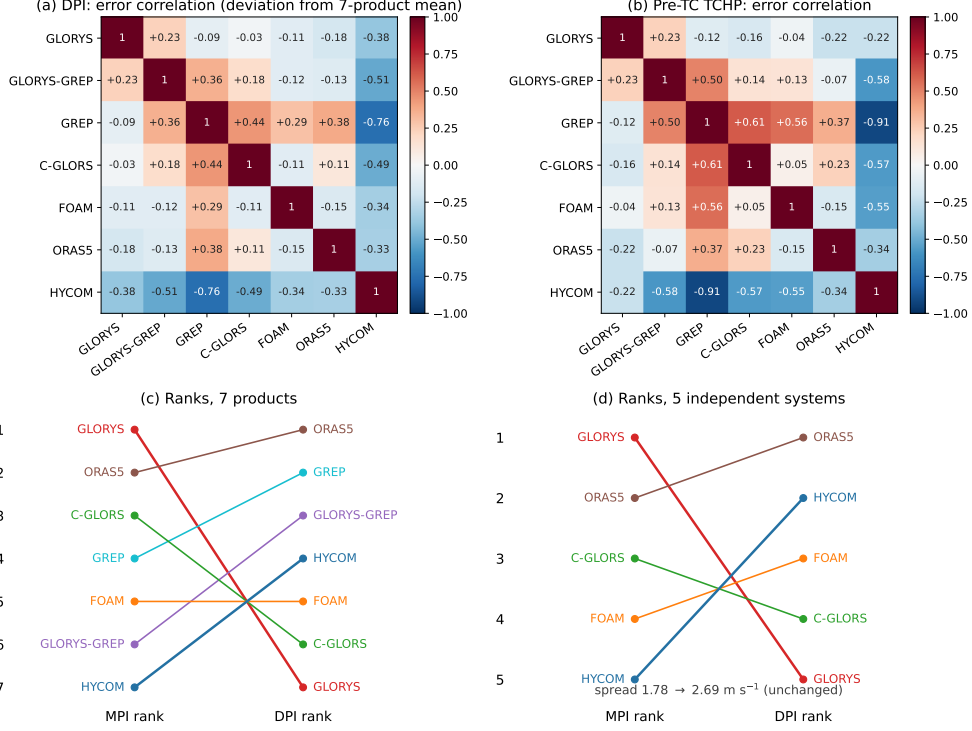
### Interdependence of the seven products

The seven products are not seven independent estimates: counted by underlying system they collapse to five, and the shared systems leave a detectable signature in the error structure (Fig. S6). On the 2,433 events matched by all seven products, Fig. S6a,b correlates each product’s event-level deviation from the 7-product ensemble mean—the shared storm environment removed—for DPI and pre-TC TCHP. The two Mercator products covary positively (DPI  $+0.23$ , TCHP  $+0.23$ , SST  $+0.34$ ), GREP covaries with each of its members ( $+0.29$  to  $+0.61$ , by ensemble construction), and HYCOM, the only system outside the CMEMS family, varies in opposite phase to all six others in every layer ( $-0.22$  to  $-0.91$ ); because the deviations sum to zero event

**Table S4** Trend tests of the main text with Benjamini–Hochberg adjusted  $p$  values within the intensity and environment families. No test survives  $q < 0.05$  in either family; the closest call is the NI basin cooling trend ( $q = 0.051$ ), and the bold rows mark the manuscripts only significant intensity trend at the nominal level (HYCOM DPI).

| Family      | Test                      | $p$           | BH $q$        |
|-------------|---------------------------|---------------|---------------|
| Intensity   | MPI C-GLORS               | 0.9368        | 0.9865        |
| Intensity   | DPI C-GLORS               | 0.7818        | 0.9865        |
| Intensity   | MPI-DPI C-GLORS           | 0.9145        | 0.9865        |
| Intensity   | MPI FOAM                  | 0.2875        | 0.8513        |
| Intensity   | DPI FOAM                  | 0.1733        | 0.7757        |
| Intensity   | MPI-DPI FOAM              | 0.7048        | 0.9865        |
| Intensity   | MPI GLORYS                | 0.4174        | 0.9505        |
| Intensity   | DPI GLORYS                | 0.4413        | 0.9505        |
| Intensity   | MPI-DPI GLORYS            | 0.3915        | 0.9505        |
| Intensity   | MPI GLORYS-GREP           | 0.8362        | 0.9865        |
| Intensity   | DPI GLORYS-GREP           | 0.8075        | 0.9865        |
| Intensity   | MPI-DPI GLORYS-GREP       | 0.6938        | 0.9865        |
| Intensity   | MPI GREP                  | 0.9671        | 0.9865        |
| Intensity   | DPI GREP                  | 0.7522        | 0.9865        |
| Intensity   | MPI-DPI GREP              | 0.9205        | 0.9865        |
| Intensity   | MPI HYCOM                 | 0.0655        | 0.3666        |
| Intensity   | <b>DPI HYCOM</b>          | <b>0.0128</b> | <b>0.2380</b> |
| Intensity   | MPI-DPI HYCOM             | 0.0323        | 0.3011        |
| Intensity   | MPI ORAS5                 | 0.2216        | 0.7757        |
| Intensity   | DPI ORAS5                 | 0.3040        | 0.8513        |
| Intensity   | MPI-DPI ORAS5             | 0.9114        | 0.9865        |
| Intensity   | MPI drifter-anchored      | 0.9865        | 0.9865        |
| Intensity   | DPI HYCOM NI              | 0.0642        | 0.3666        |
| Intensity   | DPI HYCOM SI              | 0.2216        | 0.7757        |
| Intensity   | DPI HYCOM NA              | 0.5150        | 0.9613        |
| Intensity   | MPI HYCOM NA              | 0.5000        | 0.9613        |
| Intensity   | MPI GLORYS NA             | 0.9500        | 0.9865        |
| Intensity   | LMI category 5 (observed) | 0.0170        | 0.2380        |
| Environment | SST bias C-GLORS          | 0.1504        | 0.3008        |
| Environment | SST bias FOAM             | 0.0175        | 0.1195        |
| Environment | SST bias GLORYS           | 0.9562        | 0.9722        |
| Environment | SST bias GLORYS-GREP      | 0.2739        | 0.4382        |
| Environment | SST bias GREP             | 0.3324        | 0.4693        |
| Environment | SST bias HYCOM            | 0.0333        | 0.1332        |
| Environment | SST bias ORAS5            | 0.1457        | 0.3008        |
| Environment | TCHP C-GLORS              | 0.1214        | 0.2913        |
| Environment | TCHP FOAM                 | 0.0743        | 0.2228        |
| Environment | TCHP GLORYS               | 0.1995        | 0.3419        |
| Environment | TCHP GLORYS-GREP          | 0.3743        | 0.4984        |
| Environment | TCHP GREP                 | 0.1931        | 0.3419        |
| Environment | TCHP HYCOM                | 0.9722        | 0.9722        |
| Environment | TCHP ORAS5                | 0.3319        | 0.4693        |
| Environment | TCHP Argo                 | 0.7868        | 0.8584        |
| Environment | cooling EP                | 0.0461        | 0.1581        |
| Environment | cooling NA                | 0.5827        | 0.6992        |
| Environment | <b>cooling NI</b>         | <b>0.0021</b> | <b>0.0507</b> |
| Environment | cooling SI                | 0.3946        | 0.4984        |
| Environment | cooling SP                | 0.6780        | 0.7749        |
| Environment | cooling WP                | 0.0199        | 0.1195        |
| Environment | SST WP GLORYS             | 0.0140        | 0.1195        |
| Environment | TCHP category-4 Argo      | 0.0330        | 0.1332        |
| Environment | cooling HYCOM global      | 0.0890        | 0.2373        |

by event, part of the negative side of the matrix is forced, but the positive same-system blocks and HYCOM's uniform opposition are not. Nothing in the paper rests on the doubled members: restricted to the five independent systems, the MPI→DPI reordering is intact (HYCOM rises from last to second, GLORYS falls from first to last; Fig. S6c,d) and the cross-product spread is unchanged ( $1.78 \rightarrow 2.69 \text{ m s}^{-1}$ ).



**Fig. S6** Interdependence of the seven products (2,433 common events). (a) DPI and (b) pre-TC TCHP: Pearson correlation, across events, of each product's deviation from the 7-product ensemble mean (the shared environment removed). (c) MPI→DPI rank slope for the seven products. (d) The same for the five independent systems; the cross-product spread is unchanged.

Tables S1 and S2 tabulate the per-basin, per-product selection matrix behind the selection rules of the main text.

**Table S1** Per-basin, per-product selection matrix, part 1 (WP, NA, EP). Event-level means of MPI and DPI (positions where both are defined), the fixed-environment sensitivity dMPI/dSST, pre-TC TCHP, TC-induced cooling (GLORYS-based, as in the main text), and the ratio  $\alpha = V_{\max}/\text{MPI}$  (event medians) for the seven products in each basin (MPI, DPI and dMPI/dSST in  $\text{m s}^{-1}$ , the latter per  $^{\circ}\text{C}$ ; TCHP in  $\text{kJ cm}^{-2}$ ; cooling and SST bias in  $^{\circ}\text{C}$ ), together with the drifter-referenced SST bias (from Table 1 of the main text; \* = 95% confidence interval excludes zero). Part 2 (NI, SI, SP) is Table S2.

| Basin | Product     | $n$ | MPI  | DPI  | dMPI/dSST | TCHP | Cooling | $\alpha$ (MPI) | SST bias |
|-------|-------------|-----|------|------|-----------|------|---------|----------------|----------|
| WP    | GLORYS      | 714 | 74.2 | 63.9 | 9.9       | 67.3 | 0.270   | 0.62           | +0.115*  |
| WP    | GREP        | 714 | 73.0 | 65.1 | 10.1      | 62.5 | 0.270   | 0.63           | +0.005   |
| WP    | C-GLORS     | 714 | 73.1 | 64.6 | 10.1      | 62.2 | 0.270   | 0.63           | -0.012   |
| WP    | FOAM        | 714 | 72.7 | 64.6 | 10.1      | 62.0 | 0.270   | 0.63           | -0.070*  |
| WP    | GLORYS-GREP | 714 | 72.8 | 65.1 | 10.1      | 63.9 | 0.270   | 0.63           | +0.039*  |
| WP    | ORAS5       | 714 | 73.7 | 66.2 | 10.0      | 66.9 | 0.270   | 0.62           | +0.061*  |
| WP    | HYCOM       | 714 | 71.6 | 64.1 | 10.2      | 70.1 | 0.270   | 0.64           | -0.247*  |
| NA    | GLORYS      | 446 | 75.0 | 63.9 | 10.3      | 38.1 | 0.258   | 0.54           | -0.032   |
| NA    | GREP        | 446 | 74.4 | 65.9 | 10.4      | 33.9 | 0.258   | 0.55           | -0.114*  |
| NA    | C-GLORS     | 446 | 74.2 | 64.8 | 10.4      | 32.6 | 0.258   | 0.55           | -0.150*  |
| NA    | FOAM        | 446 | 74.8 | 66.0 | 10.3      | 33.8 | 0.258   | 0.54           | -0.092*  |
| NA    | GLORYS-GREP | 446 | 73.9 | 65.9 | 10.4      | 36.1 | 0.258   | 0.55           | -0.112*  |
| NA    | ORAS5       | 446 | 74.5 | 66.7 | 10.4      | 37.5 | 0.258   | 0.54           | -0.098*  |
| NA    | HYCOM       | 445 | 71.9 | 64.1 | 9.4       | 37.4 | 0.258   | 0.57           | -0.196*  |
| EP    | GLORYS      | 426 | 68.5 | 54.6 | 11.6      | 31.9 | 0.183   | 0.64           | +0.080*  |
| EP    | GREP        | 426 | 67.4 | 56.1 | 11.8      | 27.5 | 0.183   | 0.65           | -0.011   |
| EP    | C-GLORS     | 426 | 67.9 | 55.7 | 11.7      | 26.5 | 0.183   | 0.64           | +0.013   |
| EP    | FOAM        | 426 | 67.7 | 56.6 | 11.8      | 28.1 | 0.183   | 0.64           | -0.086*  |
| EP    | GLORYS-GREP | 426 | 66.9 | 56.0 | 11.8      | 28.9 | 0.183   | 0.65           | +0.007   |
| EP    | ORAS5       | 426 | 67.5 | 56.7 | 11.8      | 30.1 | 0.183   | 0.66           | +0.032   |
| EP    | HYCOM       | 425 | 69.2 | 58.7 | 11.0      | 29.7 | 0.183   | 0.62           | -0.081*  |

**Table S2** Per-basin, per-product selection matrix, part 2 (NI, SI, SP). Same quantities as Table S1, for the remaining three basins.

| Basin | Product     | $n$ | MPI  | DPI  | dMPI/dSST | TCHP | Cooling | $\alpha$ (MPI) | SST bias |
|-------|-------------|-----|------|------|-----------|------|---------|----------------|----------|
| NI    | GLORYS      | 149 | 78.7 | 69.1 | 10.1      | 62.9 | 0.282   | 0.49           | +0.015   |
| NI    | GREP        | 148 | 78.4 | 71.9 | 10.1      | 58.4 | 0.282   | 0.50           | -0.035   |
| NI    | C-GLORS     | 148 | 78.3 | 70.9 | 10.1      | 54.8 | 0.282   | 0.50           | -0.078*  |
| NI    | FOAM        | 147 | 77.9 | 70.9 | 10.2      | 58.2 | 0.282   | 0.51           | -0.089*  |
| NI    | GLORYS-GREP | 148 | 78.0 | 72.4 | 10.1      | 61.5 | 0.282   | 0.50           | -0.008   |
| NI    | ORAS5       | 148 | 79.2 | 73.4 | 10.1      | 63.0 | 0.282   | 0.50           | +0.038   |
| NI    | HYCOM       | 149 | 77.2 | 70.7 | 10.3      | 64.6 | 0.282   | 0.50           | -0.299*  |
| SI    | GLORYS      | 409 | 72.2 | 60.9 | 11.2      | 37.8 | 0.243   | 0.62           | +0.198*  |
| SI    | GREP        | 407 | 71.0 | 62.5 | 11.3      | 33.7 | 0.243   | 0.62           | +0.096*  |
| SI    | C-GLORS     | 407 | 71.6 | 62.2 | 11.2      | 33.4 | 0.243   | 0.62           | +0.126*  |
| SI    | FOAM        | 407 | 69.6 | 61.0 | 11.5      | 32.2 | 0.243   | 0.63           | -0.054   |
| SI    | GLORYS-GREP | 407 | 70.8 | 62.4 | 11.3      | 35.0 | 0.243   | 0.63           | +0.126*  |
| SI    | ORAS5       | 407 | 72.0 | 63.8 | 11.2      | 37.9 | 0.243   | 0.62           | +0.184*  |
| SI    | HYCOM       | 409 | 70.7 | 63.1 | 11.4      | 42.4 | 0.243   | 0.62           | -0.026   |
| SP    | GLORYS      | 296 | 76.7 | 66.8 | 10.9      | 54.5 | 0.301   | 0.59           | +0.114*  |
| SP    | GREP        | 296 | 75.5 | 68.8 | 11.0      | 49.6 | 0.301   | 0.60           | -0.001   |
| SP    | C-GLORS     | 296 | 75.7 | 68.2 | 11.0      | 48.6 | 0.301   | 0.59           | -0.014   |
| SP    | FOAM        | 296 | 75.3 | 68.2 | 11.0      | 48.6 | 0.301   | 0.61           | -0.054   |
| SP    | GLORYS-GREP | 296 | 75.3 | 68.6 | 11.0      | 49.2 | 0.301   | 0.60           | +0.016   |
| SP    | ORAS5       | 296 | 76.0 | 70.0 | 10.9      | 54.3 | 0.301   | 0.59           | +0.050   |
| SP    | HYCOM       | 296 | 74.5 | 68.5 | 10.8      | 69.1 | 0.301   | 0.59           | -0.005   |

## References

- [1] Bister, M. & Emanuel, K. A. Low frequency variability of tropical cyclone potential intensity. 1. Interannual to interdecadal variability. *J. Geophys. Res.* **107**, 4801 (2002).
- [2] Kossin, J. P. Validating atmospheric reanalysis data using tropical cyclones as thermometers. *Bull. Amer. Meteor. Soc.* **96**, 1089–1096 (2015).
- [3] Emanuel, K. A. Sensitivity of tropical cyclones to surface exchange coefficients and a revised steady-state model incorporating eye dynamics. *J. Atmos. Sci.* **52**, 3969–3976 (1995).
- [4] Emanuel, K. A. A similarity hypothesis for air–sea exchange at extreme wind speeds. *J. Atmos. Sci.* **60**, 1420–1428 (2003).
- [5] Wing, A. A., Emanuel, K. & Solomon, S. On the factors affecting trends and variability in tropical cyclone potential intensity. *Geophys. Res. Lett.* **42**, 8669–8677 (2015).
- [6] Gilford, D. M. pyPI (v1.3): Tropical cyclone potential intensity calculations in Python. *Geosci. Model Dev.* **14**, 2351–2369 (2021).
- [7] Bell, M. M., Montgomery, M. T. & Emanuel, K. A. Air–sea enthalpy and momentum exchange at major hurricane wind speeds observed during CBLAST. *J. Atmos. Sci.* **69**, 3197–3222 (2012).
- [8] Kaplan, J. & DeMaria, M. Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic basin. *Wea. Forecasting* **18**, 1093–1108 (2003).
- [9] Pettitt, A. N. A non-parametric approach to the change-point problem. *J. Roy. Statist. Soc. Ser. C (Appl. Statist.)* **28**, 126–135 (1979).
- [10] Quenouille, M. H. *Associated Measurements* (Butterworths, London, 1952).