

OceanPredict
Advancing the science of ocean prediction

Intercomparison and Validation Task Team (IV-TT)

Chunxue Yang, Stefania Ciliberti, Babette Tchonang, Fabrice Hernandez, Gregory Smith

OPST Meeting | Montreal, 15-17 April 2026

IV-TT Task Team

Evolution of the Task Team

Former Co-Chairs: G. Smith and F. Hernandez

- New Co-Chairs since Mar 2026: C. Yang, S. Ciliberti and B. Tchonang
- B. Tchonang joins as Early Career Scientist
- F. Hernandez to support/help if they need the new team

No	First name	Last name	Affiliation	Country	Role
1	Maria	Aristizabal	NOAA/EMC	USA	Member
2	Laurent	Bertino	NERSC	Norway	Member
3	Stefania	Ciliberti	NewSystems	Spain	New member/co-chair
4	Andrea	Cipollone	CMCC	Italy	Member
5	Emanuela	Clementi	CMCC	Italy	Member
6	Jim	Cummings	NRL	USA	Member
7	Prasanth	Divakaran	BoM	Australlia	Member
8	Marie	Drevillon	MOI	France	Member
9	Yosuke	Fujii	MRI-JMA	Japan	Member
10	Fabrice	Hernandez	IRD/LEGOS	France	Co-chair
11	Patrick	Hogan	NOAA	USA	Member
12	Robert	King	Met Office	UK	Member
13	Jae-Il	Kwon	KIOST	Korea	Member
14	Yvonnick	La Clainche	DFO	Canada	Member
15	Matt	Martin	Met Office	UK	Member
16	Avichal	Mehra	NOAA-NCEP	USA	Member
17	Peter	Oke	CSIRO	Australlia	Member
18	Fangli	Qiao	FIO	China	Member
19	Sulagna	Ray	NOAA	USA	Member
20	Charly	Regnier	MOI	France	Member
21	Pavel	Sakov	CSIRO	Australlia	Member
22	Greg	Smith	ECCC	Canada	Co-chair
23	Clemente	Tanajura	UFBA	Brazil	Member
24	Babette	Tchonang	NOAA	USA	Member
25	LiyIng	Wan	NMEFC	China	Member
26	Chunxue	Yang	CNR	Italy	New member/co-chair

Recent activities led by former co-chairs

A new scientific paper was submitted to BAMS thanks to contribution of all IV-TT members:

Near Real-time Evaluation of Operational Oceanographic Predictions: Over 12 Years of Intercomparison by Global Ocean Forecasting Centers

Main topics:

- Introduction of the IV-TT in the last 12 years
- Lessons learnt:
 - **Multiple international forecasting systems have been compared** using the same observations (T, S, SL, etc.)
 - **Class-4 intercomparison:**
 - *Opportunity:*
 - Feasibility and robustness of the Class-4 approach to **forecast verification** was made for global scales.
 - **Errors are defined as the average statistics of differences** (mean and standard deviation) between operational system daily mean fields and observations.
 - *Limits:*
 - **Lack of independent observations.**
 - **Limited number** of assessed variables.
 - **Sparse observations** in Arctic, coastal, and deep ocean regions

Legacy:

- Class-4 metrics provide a standardized framework for best practices
- Pre-operational framework secured with ECCC cloud hosting Class4 and standardization of metrics
- Possibility to move to Class 1 metrics to support user- and process- oriented metrics

Challenges:

- Move from research “best effort” to fully operational evaluation system (link with OPOS)
- Expand datasets verification (e.g., sea ice, currents, biogeochemistry)
- Adapt or improve for AI-based forecasting systems (→ AI-TT)
- Improve global observing networks (→ OS Eval-TT for linkage with GOOS et al.)

Variable	Data type	Source of data	Initial production center
Potential temperature & salinity	In situ Argo profiles	Copernicus Marine Data Store: INSITU_GLO_PHYBGCWAV_DISCRETE_MYNRT_013_030	UK Metoffice
Sea level anomaly	Satellite altimetry	CLS Aviso Level 3 data. SLA data used include all altimeters which are assimilated into UK Met Office operational system. Currently that is Cryosat-2, SARAL-AltiKa, Jason-3, Sentinel-3A, Sentinel-3B and Sentinel-6A. Previously also included Jason-1 and Jason-2.	UK Metoffice
Sea surface temperature	In situ surface drifter data	From US FNMOC via the USGODAE server: http://www.usgodae.org/cgi-bin/datalist.pl?dset=fnmoc_obs_sfcobs&summary=Go	UK Metoffice
Sea ice concentration	AMSR-2 data	JAXA (http://suzaku.eorc.jaxa.jp/GCOM_W/)	ECCC
15-m velocities	CMEMS GDP drifter trajectories/velocities	Copernicus Marine Data Store: INSITU_GLO_UV_REP_OBSERVATIONS_013_048	MOi

Table 1: Observation datasets used in the Class-4 intercomparison and their respective sources. The production center responsible for producing the reference file is indicated in the right-most column.

- Setup of new server at ECCC to support Class-4 calculation.
- Defining policies for sharing and using Class-4 metrics.
- Discussing future approaches for improving robustness of Class-4 intercomparison.

Smith et al., 2026

IV-TT Meetings

June 2025

Agenda

- Status of Class4 intercomparison (US GODAE, transition to ECCC server) **(5 mins)**
- Proposal for new standard datasets to use for class4 intercomparison (SLA, SST, profiles) **(20mins)**
- Plan to improve robustness of Class4 intercomparison **(5mins)**
- Policy with respect to use and sharing of class4 data **(5mins)**
- 5-yr strategic plan (links to other TT and UN Decade initiatives) **(5mins)**
- Preparation of publication on Class 4 intercomparison **(5mins)**
- Online seminar series **(5mins)**
- Opportunity for new co-chairs **(5mins)**

March 2026

Agenda

- Opening and News from IV-TT (Greg and Fabrice, **5mins**)
- Propose new matrix to integrate to Class4 intercomparison or establish a new Class (e.g., Class5 to valid AI based forecasts and connection with OceanBench, AI-TT, COSS-TT) **(15mins)**
- Current status of Class4 (Greg, **10mins**)
- 5-yr and 10-yr strategic plan (links to other TT and UN Decade initiatives) to be discussed during OPST meeting in April 2026 (Greg, **10mins**)
- In person meeting for OSPT meeting in Montreal **(5mins)**
- IV-TT In-person meeting 2026 in China **(5mins)**
- Online seminar series **(5mins)**
- Place to share documents **(5mins)**

Bridging toward the future

- Move from research to fully operational evaluation system
- Expand datasets verification (e.g., sea ice, currents, biogeochemistry, etc.)
- Extend verification to regional system
- Develop new metrics for AI-based forecasting systems and connection with Ocean Bench
- Improve global observing networks (link with GOOS et al.)
- Link to other activities, e.g MER-EP, Ocean Prediction DCC

Link, liaise with:

OPOS-TT

OSEval-TT, MEAP-TT

COSS-TT

AI-TT

OSEval-TT

DTO, ForeSea

From ***Smith et al. 2026*** there is the need to:

- Present the strengths and limitations of different operational system at regional levels
- Provide an international standard for validating the forecast system
- Non-stationary observing system could result in forecasts error changes.
- Sample issues in observations (e.g., sea ice)
- MDT (model and satellite) and Argo data (coastal, sea ice zone)

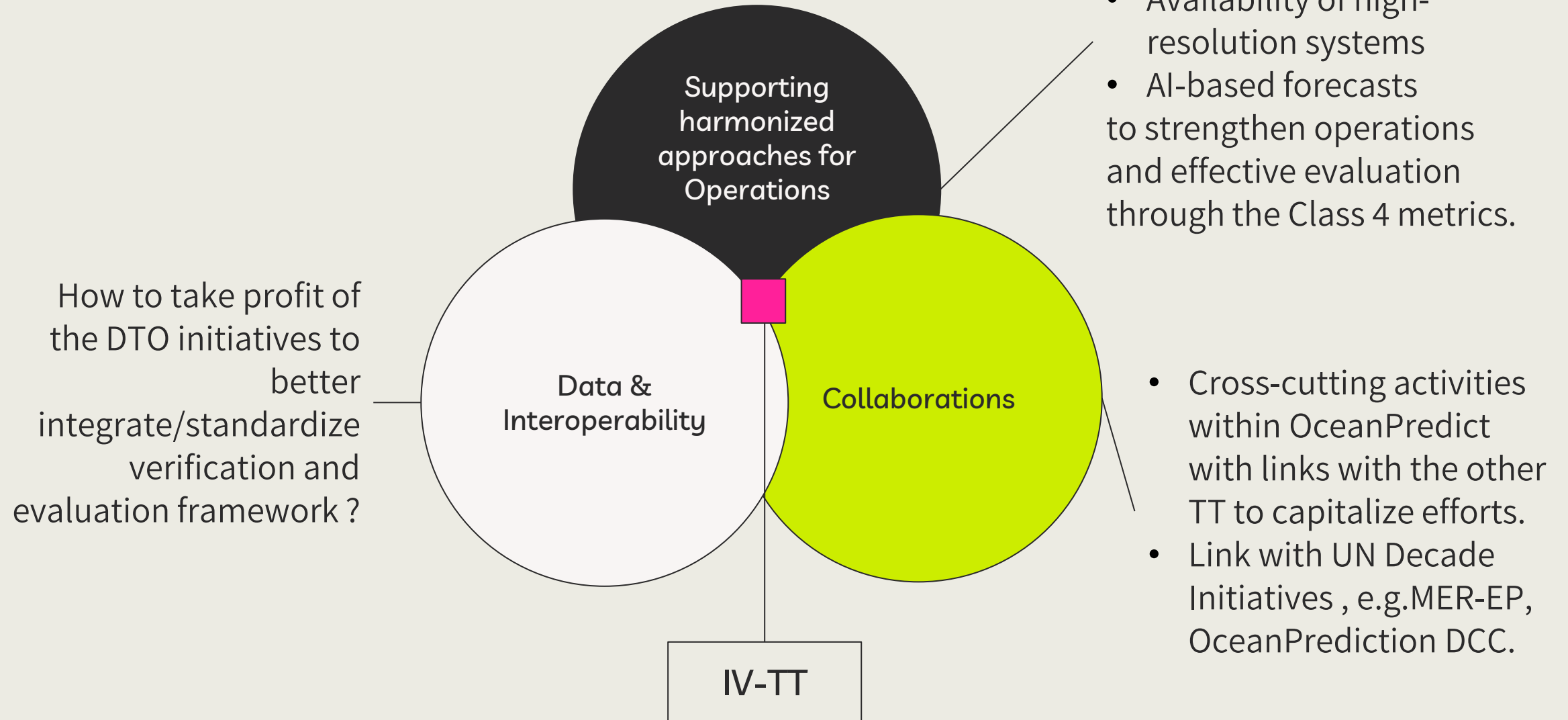


How can the next phase help in this regard?

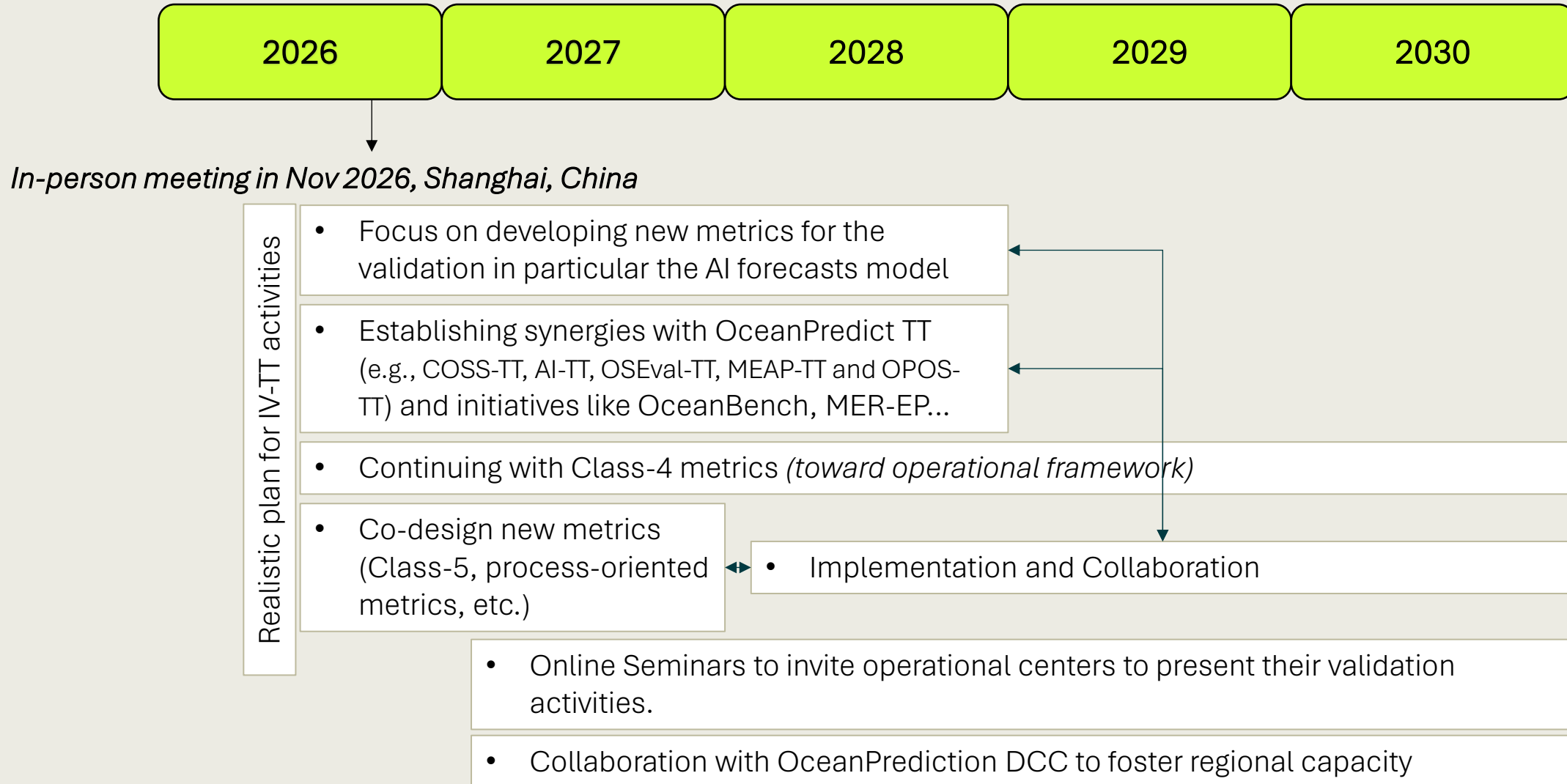
- Assessing improved forecast accuracy through ensemble methods
- Better model physics and data assimilation thanks to systems' evolution
- Systems are evolving with:
 - Higher resolution models
 - Coupled ocean-atmosphere systems
 - AI-based approaches
 - Cloud based organization (DTO)

implying **need of common framework for performing intercomparison and validation.**

Challenges for future IV-TT workplan



Initial IV-TT workplan



Contribution to Ocean Predict 5 Years Plan

2. Verification

- Develop and consolidate new metrics for validation (evolution of Class-4, process-oriented metrics, new variables...)

AI-TT

COSS-TT

7. Operationality

- Provide validation metrics to operational systems

OPOS

11. Monitoring and assessments

- Provide validation metrics for forecast system assessments

IV-TT Core

8. International collaborations

- Connect with other initiatives such as MER-EP and Ocean Prediction DCC (expansion of the ORL Validation criteria...) with outlook on EDITO, ForeSea, etc...



THANKS!