INTERNATIONAL RESEARCH CENTER OF BIG DATA FOR SUSTAINABLE DEVELOPMENT GOALS
可持续发展大数据国际研究中心

Marine data mining and digital twin supporting marine SDGs

Prof. XUE Cunjin

Sep.3 2025

Outlines

- Challenges and demands from Marine SDGs
- Marine data-added mining technologies
- Digital twin of the ocean promotes 'what-if' scenarios simulations
- Marine information system for marine SDGs
- Case studies of monitoring and evaluating marine SDGs

Who am I ?

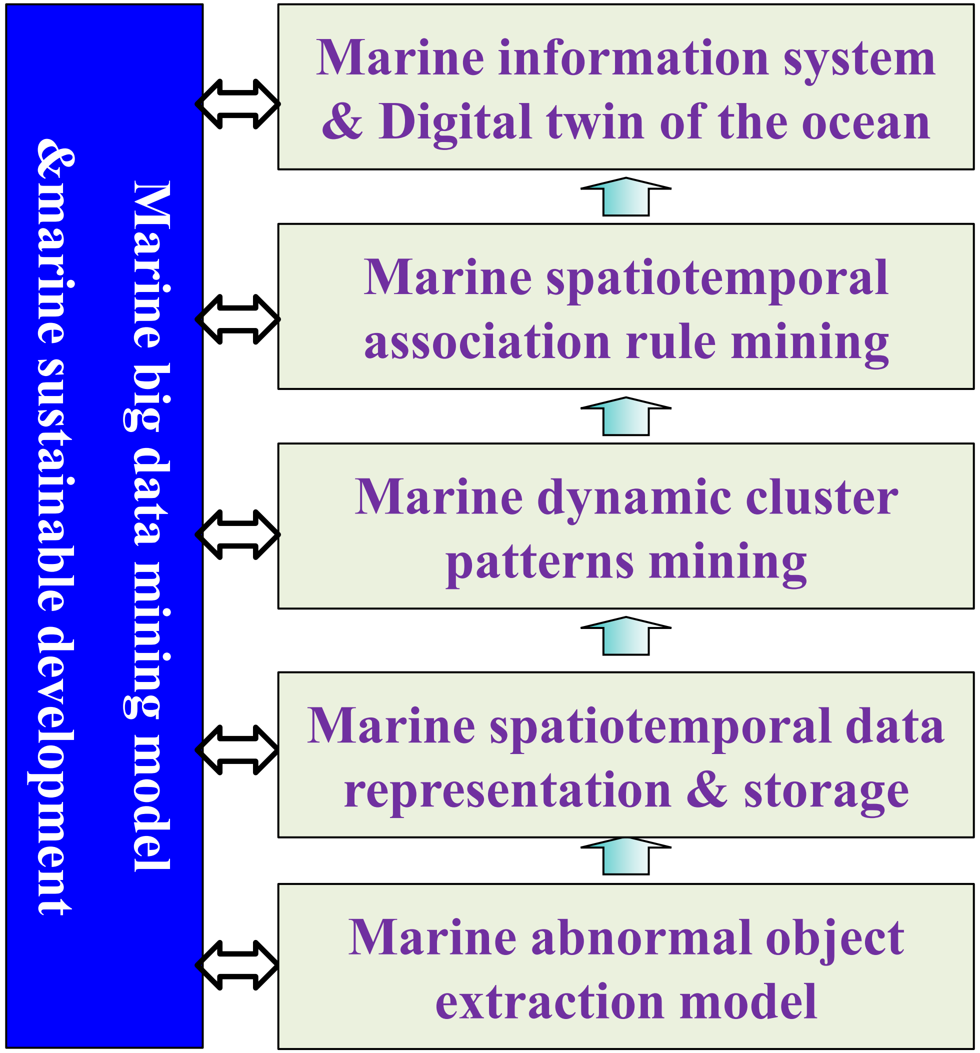
Email: xuecj@aircas.ac.cn
https://people.ucas.edu.cn/~goldensnow

Wet Chat

XUE Cunjin, Male, Dr. Prof. International Research Center of Big Data For Sustainable Development Goals (CBAS) and Aerospace Information Research Institute, Chinese Academy of Sciences (CAS)

- A director of Digital Ocean and Air Research Division.
- An engaged backbone professor of Chinese Academy of Sciences (2023)
- A member of Youth Promotion Association of Chinese Academy of Sciences (2013)
- A member of Academic Construction Committee of CBAS (2024)
- A member of Digital Ocean Committee, National Committee of China of the International Digital Earth Society
- A member of Digital Coastal Committee, National Committee of China of the International Digital Earth Society
- A member Virtual Reality Committee, National Committee of China of the International Digital Earth Society

Main researches



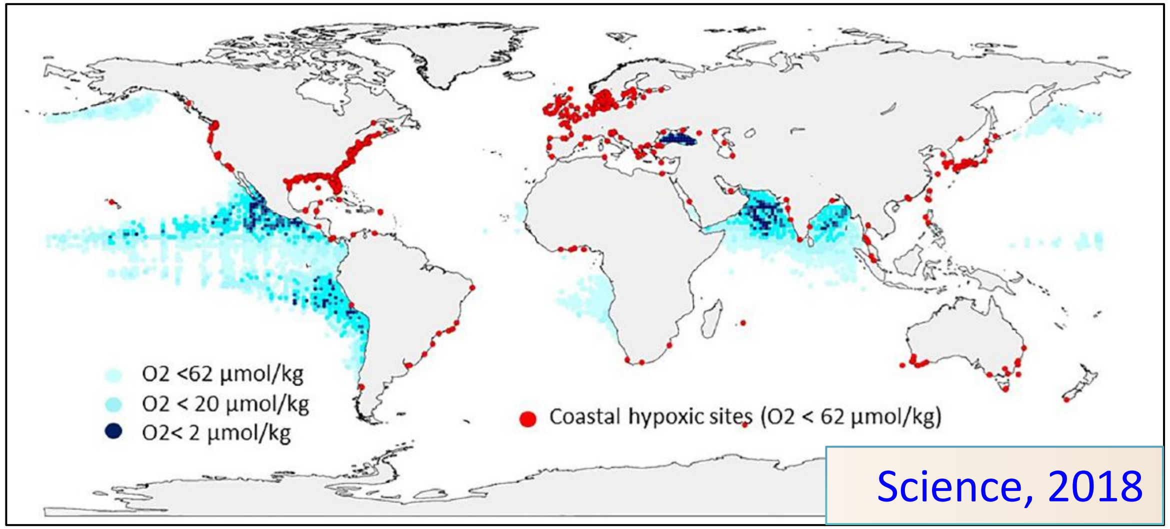
Outlines

- **Challenges and demands from Marine SDGs**
- **Marine data-added mining technologies**
- **Digital twin of the ocean promotes 'what-if' scenarios simulations**
- **Marine information system for marine SDGs**
- **Case studies of monitoring and evaluating marine SDGs**

一、Challenges from SDG14 and UNDOS

Ocean has being under threats

- Over 40% of the ocean's surface are affected by land-based and sea-based activities
- Live coral cover has nearly halved in the last 150 years
- Only 19% of the ocean floor are mapped, vast swathes of the deep ocean and Arctic and Polar regions are unknown
- Ocean 'dead zones' is being created
- Decline of structure, function and benefit of marine ecosystem



Science, 2018

Ocean science & technologies are largely competent for diagnosing problems

- Stereo monitoring capacity: Unmanned survey vessel, Satellite RS (SDGSAT-1), Space RS, Buoy
- Earth big data technology
- Ocean science development

一、 Challenges from SDG14 and UNDOS

SDG14: Conserve and sustainably use the oceans, seas and marine resources for sustainable development

14.1:Knowledge and solutions to reduce pollution on land and at sea

14.2 Knowledge and solutions for management of ecosystems faced with multiple stressors

14.3 Knowledge and solutions to reduce effects of ocean acidification

14.4 Knowledge and solutions for sustainable fisheries

14.5 Knowledge and solutions for area-based management tools

14.6 Knowledge and solutions for fishing management tools

14.7 Knowledge and solutions for a sustainable ocean economy

14.a Increased scientific knowledge, research capacity and transfer of marine technology

14.b Knowledge and solutions for increased access to markets for small-scale fishers

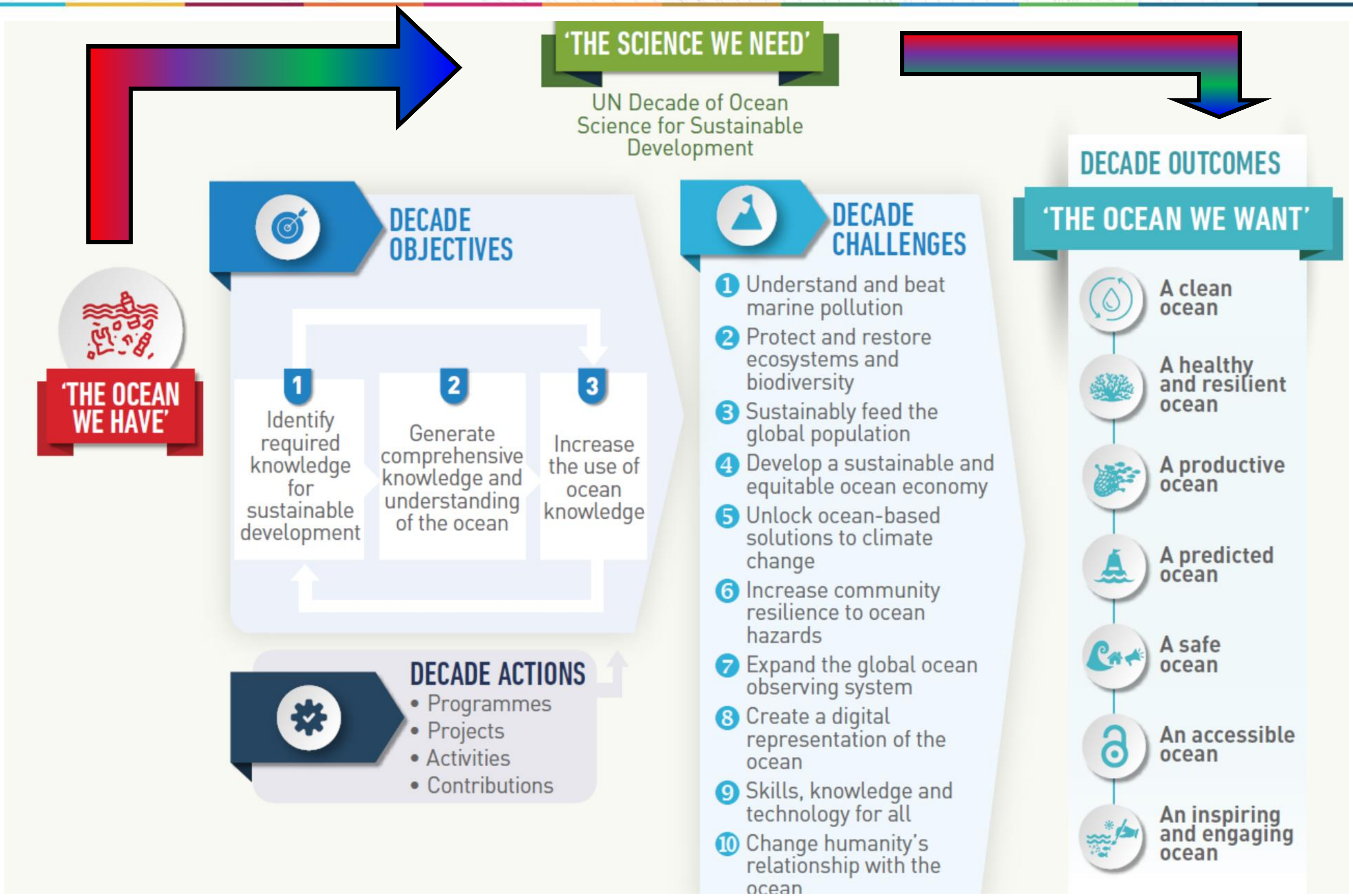
14.c Significant contribution to application of UNCLOS for conservation and sustainable use of the ocean

Action is not advancing at the speed or scale required to meet SDG14， failure to achieve targets 14.2, 14.4, 14.5 and 14.6 that matured in 2020

一、 Challenges from SDG14 and UNDOS

Vision: The science we need for the ocean we want.

Mission: To catalyse transformative ocean science solutions for sustainable development, connecting people and our ocean



一、Challenges from SDG14 and UNDOS

Objectives:

Objective 1: Identify required knowledge for sustainable development and increase the capacity of ocean science to deliver needed ocean data and information

Objective 2: Build capacity and generate comprehensive knowledge and understanding of the ocean, including human interactions and interactions with the atmosphere, cryosphere and the land-sea interface.

Objective 3: Increase the use of ocean knowledge and understanding, and develop capacity to contribute to sustainable development solutions.

一、Challenges from SDG14 and UNDOS

Outcomes: Seven oceans

Outcome 1: A clean ocean where sources of pollution are identified and reduced or removed.

Outcome 2: A healthy and resilient ocean where marine ecosystems are understood, protected, restored and managed.

Outcome 3: A productive ocean supporting sustainable food supply and a sustainable ocean economy.

Outcome 4: A predicted ocean where society understands and can respond to changing ocean conditions.

Outcome 5: A safe ocean where life and livelihoods are protected from ocean-related hazards.

Outcome 6: An accessible ocean with open and equitable access to data, information and technology and innovation.

Outcome 7: An inspiring and engaging ocean where society understands and values the ocean in relation to human well-being and sustainable development.

一、 Challenges from SDG14 and UNDOS

Knowledge and Solutions Challenges

Challenge 1: Understand and map land- and sea-based sources of pollutants and contaminants and their potential impacts on human health and ocean ecosystems, and develop solutions to remove or mitigate them.

Challenge 2: Understand the effects of multiple stressors on ocean ecosystems and develop solutions to monitor, protect, manage and restore ecosystems and their biodiversity under changing environmental, social and climate conditions.

Challenge 3: Generate knowledge, support innovation and develop solutions to optimize the role of the ocean in sustainably feeding the world’s population under changing environmental, social and climate conditions.

Challenge 4: Generate knowledge, support innovation and develop solutions for equitable and sustainable development of the ocean economy under changing environmental, social and climate conditions.

Challenge 5: Enhance understanding of the ocean-climate nexus and generate knowledge and solutions to mitigate, adapt and build resilience to the effects of climate change across all geographies and at all scales, and to improve services including predictions for the ocean, climate and weather.

一、Challenges from SDG14 and UNDOS

Essential Infrastructure Challenges

Challenge 6: Enhance multi-hazard early warning services for all geophysical, ecological, biological, weather-, climate- and anthropogenic-related ocean and coastal hazards, and mainstream community preparedness and resilience.

Challenge 7: Ensure a sustainable ocean observing system across all ocean basins that delivers accessible, timely and actionable data and information to all users.

Challenge 8: Through multi-stakeholder collaboration, develop a comprehensive digital representation of the ocean, including a dynamic ocean map, which provides free and open access for exploring, discovering and visualizing past, current and future ocean conditions in a manner relevant to diverse stakeholders.

Foundational Challenges

Challenge 9: Ensure comprehensive capacity development and equitable access to data, information, knowledge and technology across all aspects of ocean science and for all stakeholders.

Challenge 10: Ensure that the multiple values and services of the ocean for human well-being, culture and sustainable development are widely understood, and identify and overcome barriers to behaviour change required for a step change in humanity’s relationship with the ocean.

一、Challenges from SDG14 and UNDOS

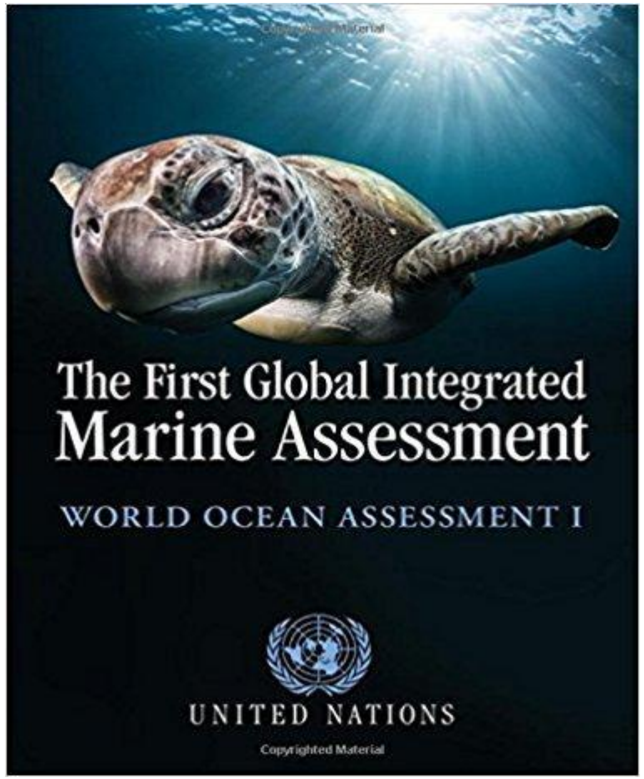
- Understand and beat marine pollution
- Protect and restore ecosystems and biodiversity
- Sustainably feed the global population
- Develop a sustainable and equitable ocean economy
- Unlock ocean-based solutions to climate change
- Increase community resilience to ocean hazards
- Expand the Global Ocean Observing System
- Create a digital representation of the ocean
- Deliver data, knowledge and technology to all
- Change humanity's relationship with the ocean

Challenge 8:
develop a comprehensive digital representation of the ocean,

Challenge 9:
comprehensive capacity development and equitable access to data, information, knowledge and technology across all aspects of ocean science ,

一、 Challenges from SDG14 and UNDOS

The ocean we have



The ocean we want

- A Clean ocean
- A healthy and resilient ocean
- A productive ocean
- A predicted ocean
- A safe ocean
- An accessible ocean
- An inspiring and engaging ocean

SDG14

SDG Goals

SDG1
SDG2
SDG3
SDG4
SDG5
SDG7
SDG8
SDG10
SDG11
SDG13

Facing these challenges, what we should do?

Outlines

- Challenges and demands from Marine SDGs
- Marine data-added mining technologies
- Digital twin of the ocean promotes 'what-if' scenarios simulations
- Marine information system for marine SDGs
- Case studies of monitoring and evaluating marine SDGs

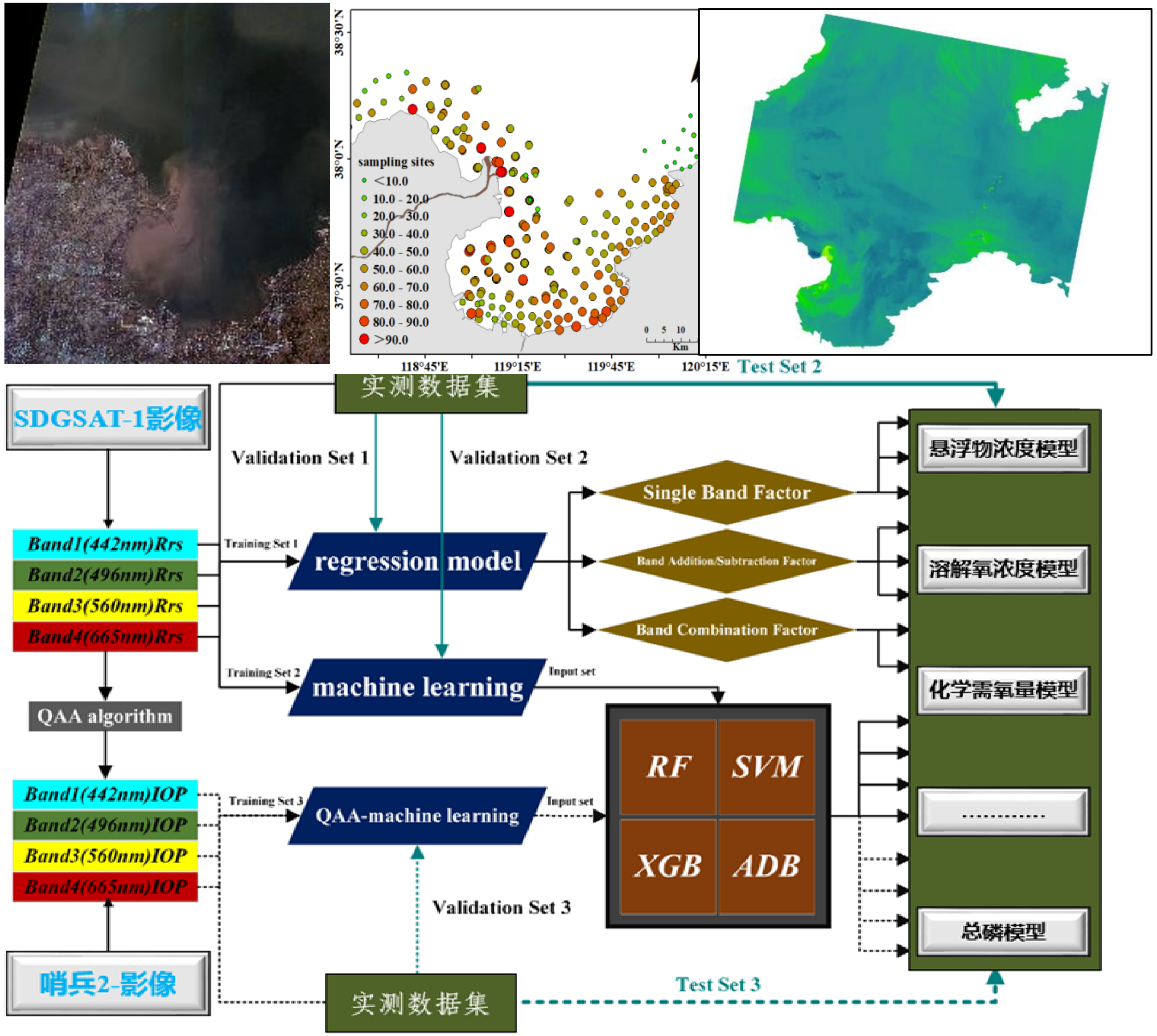
二、 Marine data-added mining technologies

- Remote sensing images-based retrieval models for marine environmental parameters
- Argo profiles-based reconstructed models for marine dissolved oxygen concentrations
- Process-oriented data mining models for marine dynamics

2.1 Remote sensing images-based retrieval models

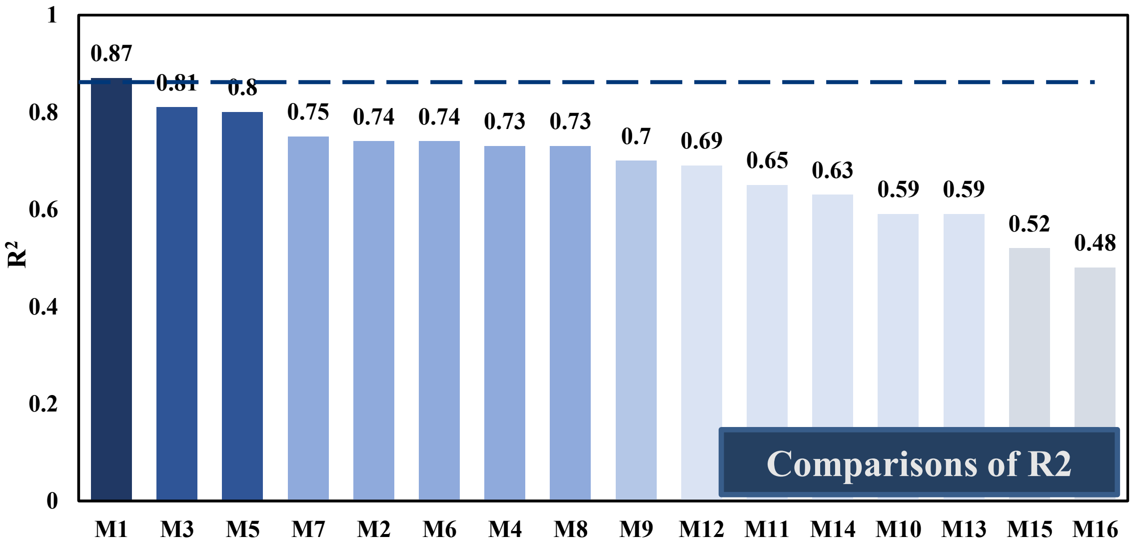
Coupling traditional regression models (QAA) and machine learning models (RF, XGB, SVM), a water quality monitoring model is developed in coastal region

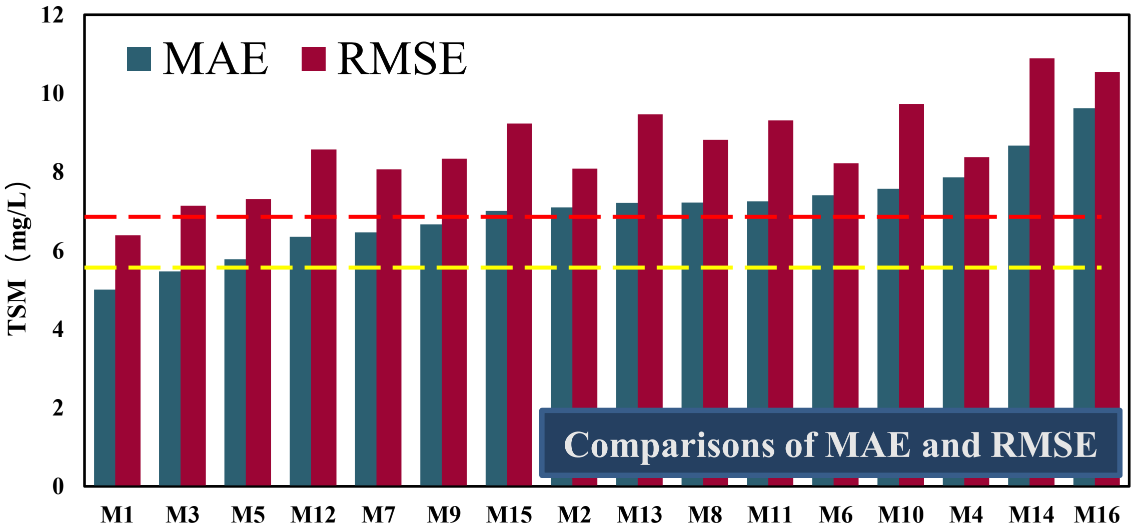
- Total suspended solids concentration
- Dissolved oxygen concentration
- Total phosphorus
- Total nitrogen
- Chemical oxygen demand
- ……



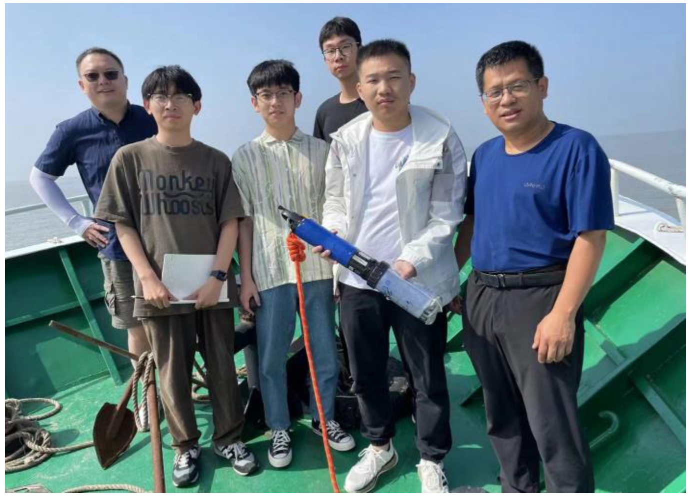
2.1 Remote sensing images-based retrieval models

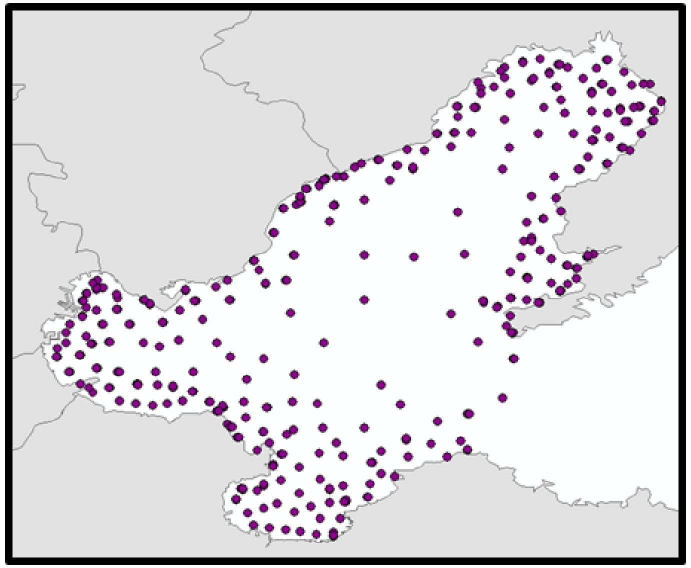
Name	Model	R^2	MAE	RMSE
M1	QAA-RF	0.87	5.01	6.39
M2	RF	0.74	7.10	8.08
M3	QAA-SVR	0.81	5.47	7.14
M4	SVR	0.73	7.86	8.38
M5	QAA-XGB	0.80	5.78	7.31
M6	XGB	0.74	7.41	8.22
M7	QAA-ADB	0.75	6.46	8.07
M8	ADB	0.73	7.22	8.82
M9	$C_{TSM} = 985.24 \times B1 - 104.38$	0.70	6.67	8.34
M10	$C_{TSM} = 839.25 \times B2 - 106.06$	0.59	7.57	9.73
M11	$C_{TSM} = 598.71 \times B3 - 77.708$	0.65	7.25	9.31
M12	$C_{TSM} = 391.14 \times (B1 + B3) - 93.676$	0.69	6.35	8.57
M13	$C_{TSM} = 355.28 \times (B2 + B3) - 91.502$	0.59	7.21	9.47
M14	$C_{TSM} = 311.04 \times (B3 + B4) - 72.811$	0.63	8.67	10.89
M15	$C_{TSM} = 598.71 \times \frac{B3+B4}{B4} - 676.42$	0.52	7.01	9.23
M16	$C_{TSM} = 449.21 \times \frac{B1+B2+B3}{B1+B2} - 689.49$	0.48	9.62	10.54





2.1 Remote sensing images-based retrieval models

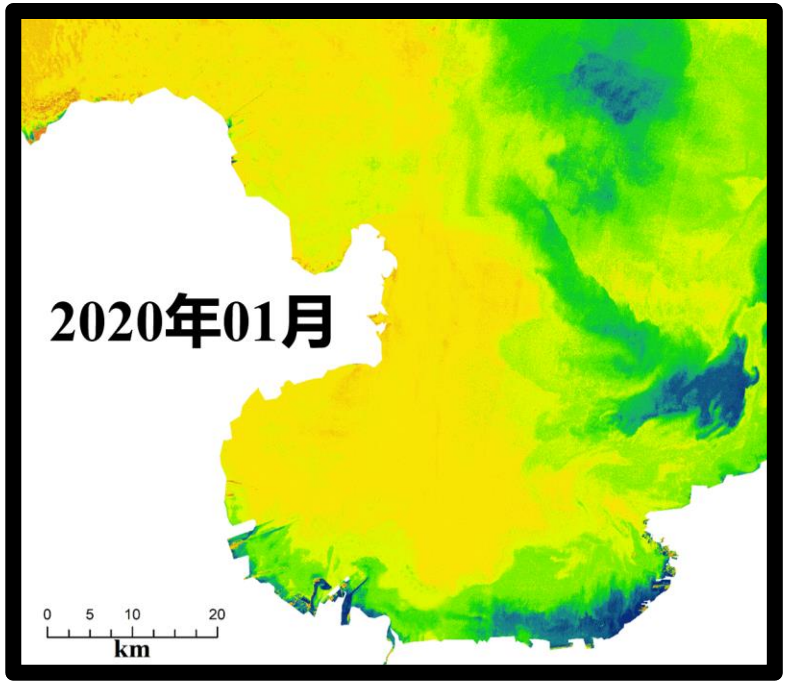




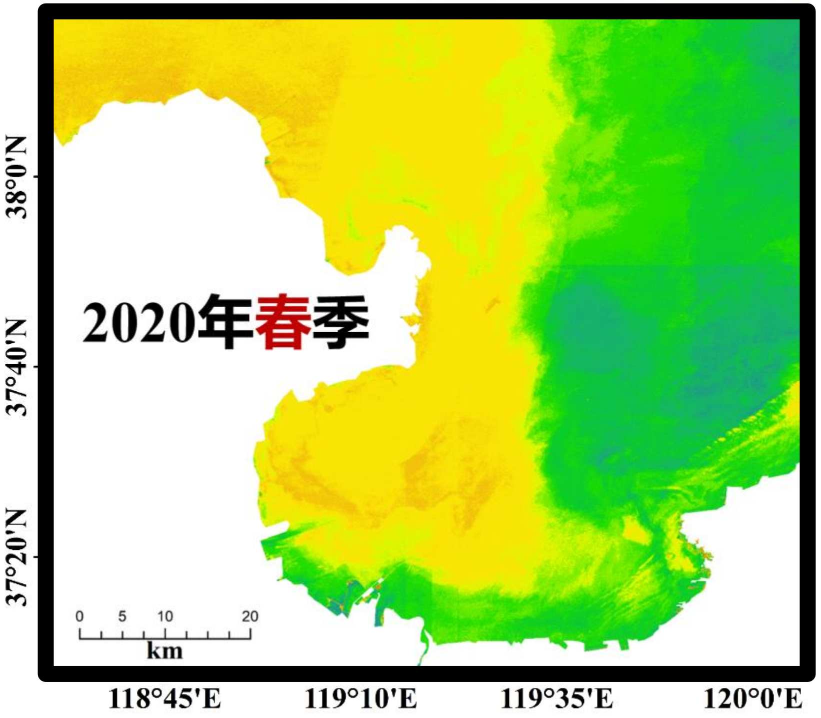
	Model	R2	MAE	RMSE
TSM	QAA-RF	0.87	5.01	6.39
DOC	QAA-RF	0.86	0.58	0.79
COD	QAA-RF	0.86	0.16	0.20
AP	QAA-RF	0.84	0.23	0.25
SO	QAA-RF	0.81	0.35	0.52
IN	QAA-RF	0.88	0.43	0.62

2.1 Remote sensing images-based retrieval models

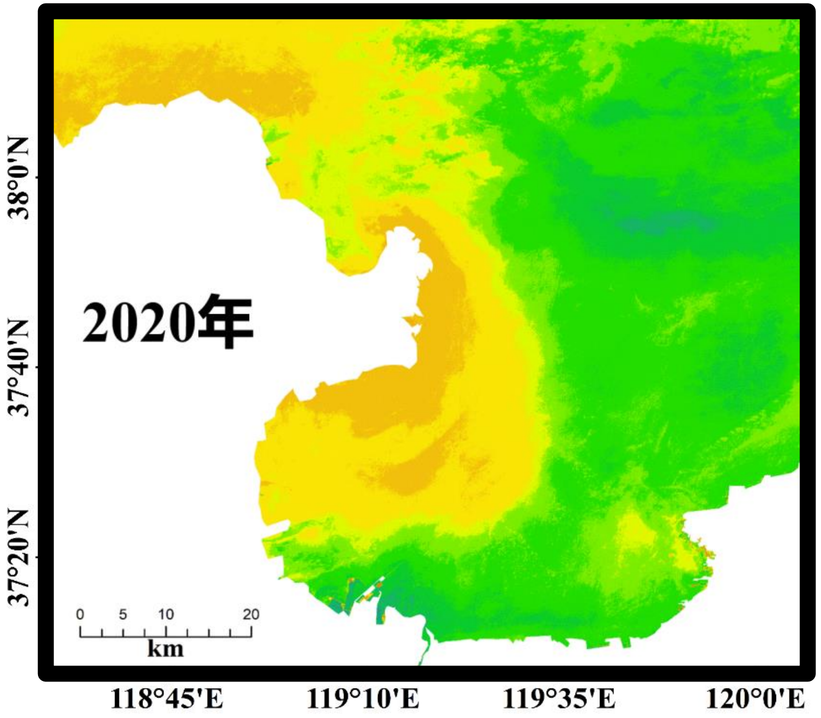
Monthly



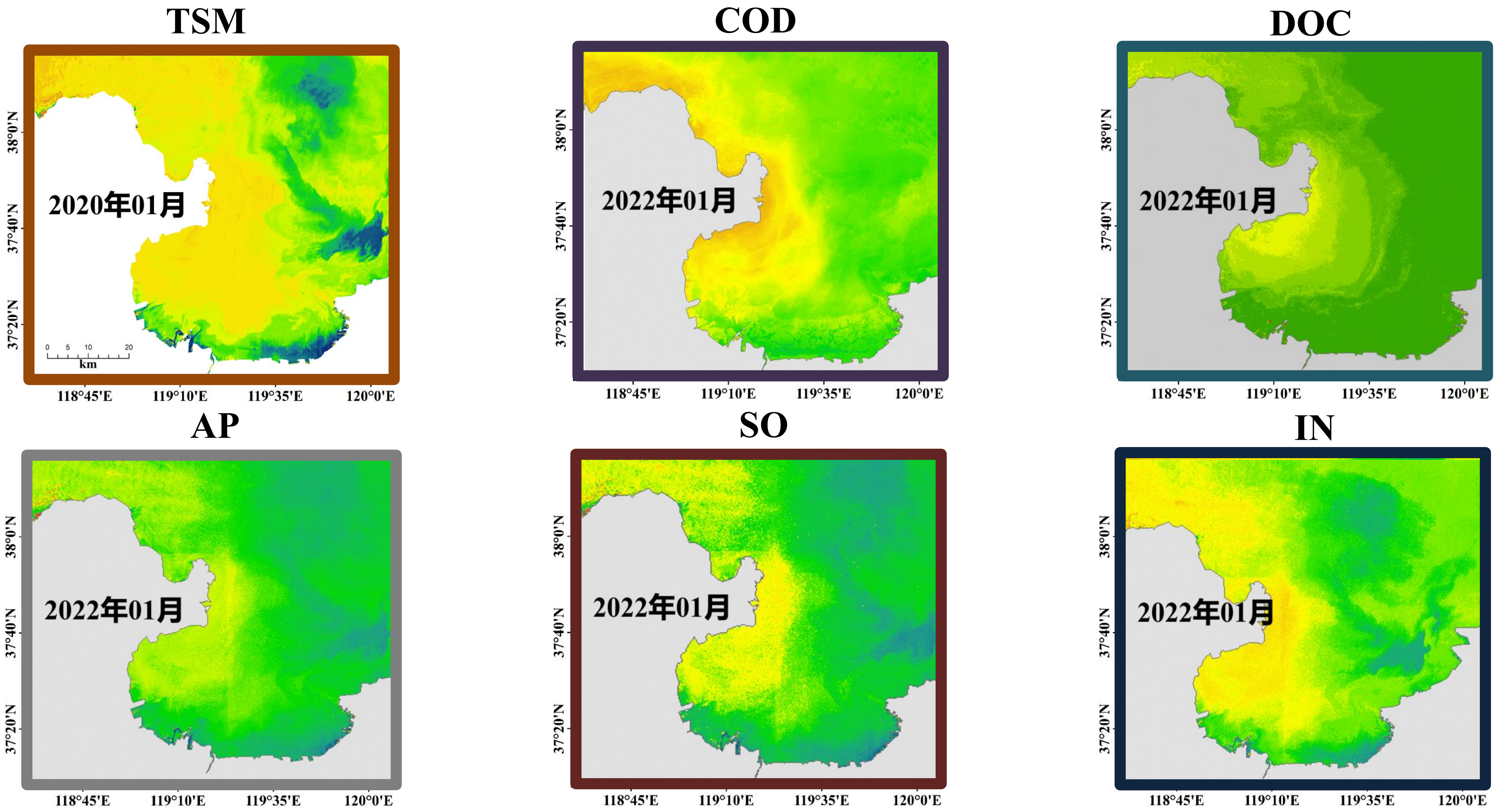
Seasonal



Annual



2.1 Remote sensing images-based retrieval models



2.2 Argo profiles-based reconstructed models

Oxygen is critical to the health of the ocean. It structures aquatic ecosystems and is a fundamental requirement for marine life from the intertidal zone to the greatest depths of the ocean

A key parameter of marine biogeochemical cycles and global climate change.

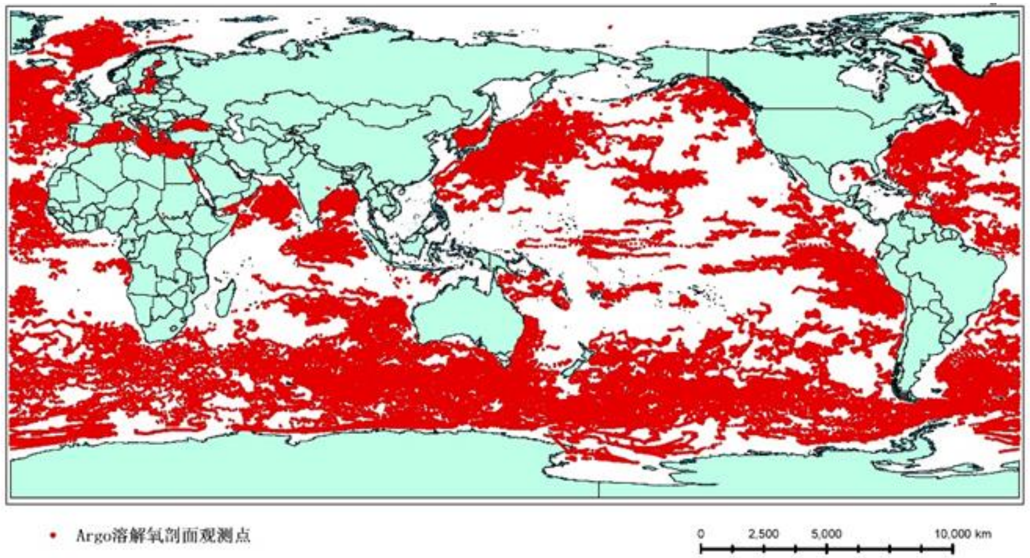
Ocean deoxygenation at global is well-known

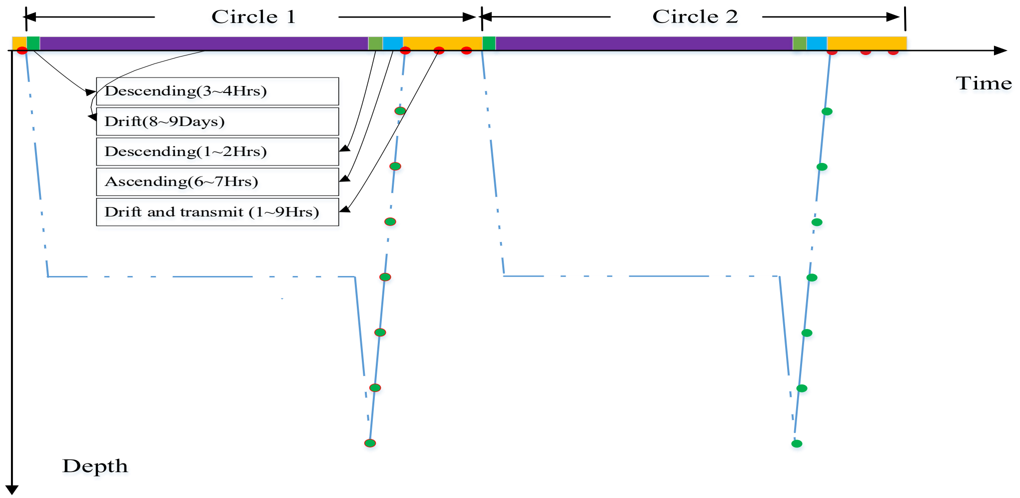
- Since the 1960s, the area of low oxygen water in the open ocean has increased by 4.5 million km2, and over 500 low oxygen sites have been identified in estuaries and other coastal water bodies.
- Deoxygenation is predicted to increase in the coming years.

Question: Do we have much more data to clarify the status of marine dissolved oxygen and deoxygenation, and their spatial distribution at global ocean ?

2.2 Argo profiles-based reconstructed models

Spatial distribution of BGC-Argo profiles





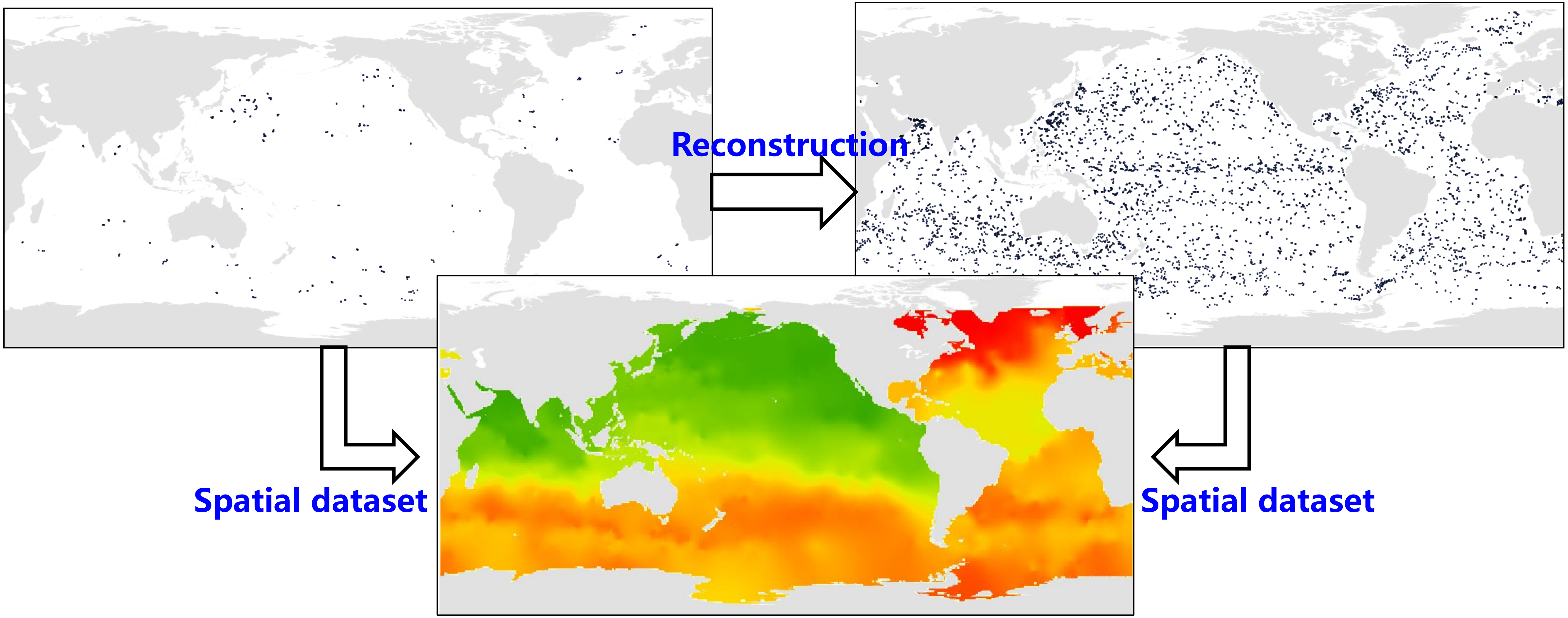
Argo is a major component of the Global Ocean Observing System , which provides hydrographic and biogeochemical data from ocean surface to medium-depth. Up date to Dec. 2024, there are more than 280， 000 BGC-Argo profiles in global ocean, and the data are public shared.

- **Depth resolution: <2000dbar**
- **Time resolution: 5-10 days**
- **Spatial distribution: Global exception of North/South Polar region**
- **Capabilities of continuous updating data acquisition**

Up to now, few such spatial dataset.

2.2 Argo profiles-based reconstructed models

Based on machining learning and geographical statistics, we develop a value-added spatiotemporal approach for exploring ocean dissolved oxygen from BGC-Argo profiles, named VAEDO. VAEDO addresses two issues. One is to reconstruct the Argo-S profiles based on T-S profiles; the second develops their spatial datasets with higher temporal and spatial resolutions.



2.2 Argo profiles-based reconstructed models

Argo temperature-salinity-oxygen profiles between 2010-2021

Four out of five of these profiles to train the model

One out of five of these profiles to test the model

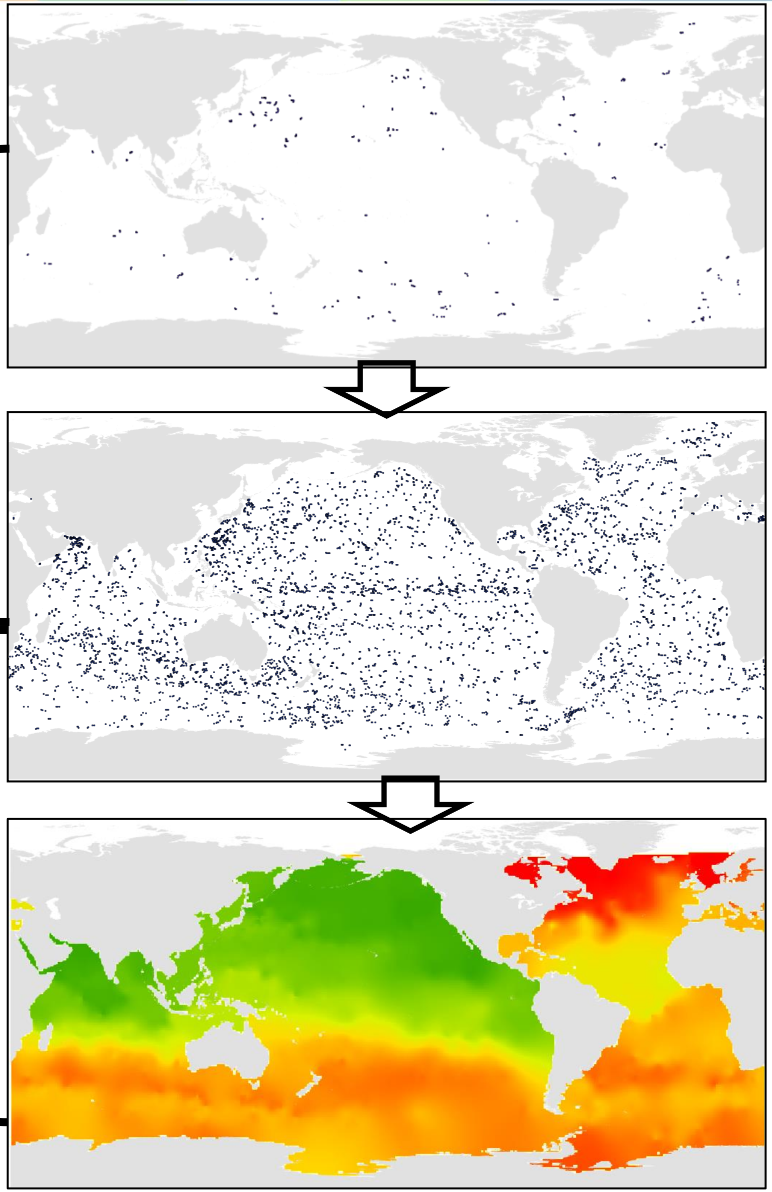
Reconstructing model for oxygen profiles based on temperature-salinity profiles

Argo oxygen profiles between 2010-2021

Reconstructed oxygen profiles between 2005-2021

Space-quality Interpolation model for ocean oxygen profiles

Global ocean oxygen spatial datasets



2.2 Argo profiles-based reconstructed models

Main steps

I. Data pretreatment

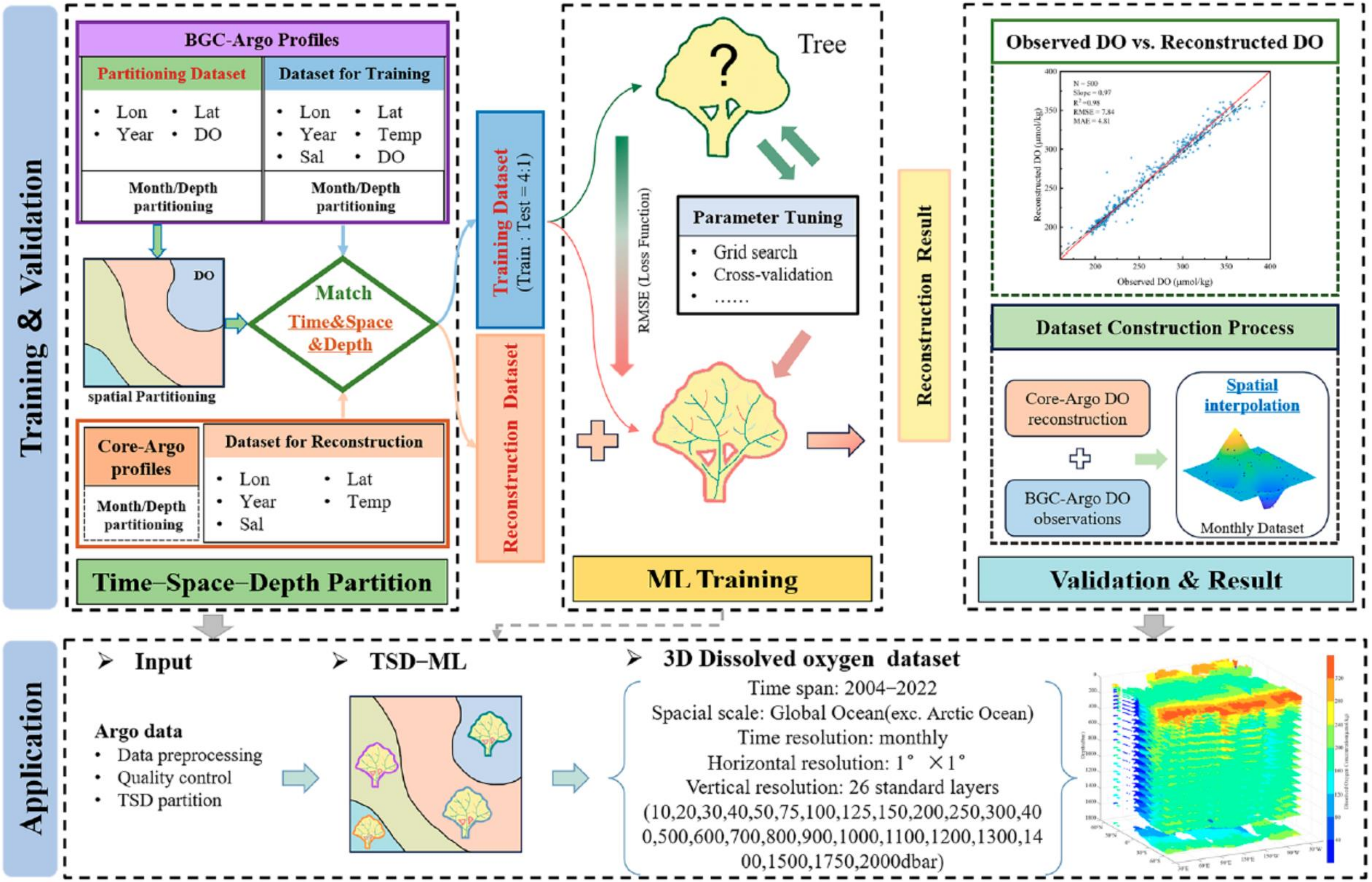
II. Time–Space–Depth Partition

III. Model training

IV. Validations

Highlights

- Time–Space–Depth Partition-based model
- Irregular spatial distribution to spatial gridded dataset

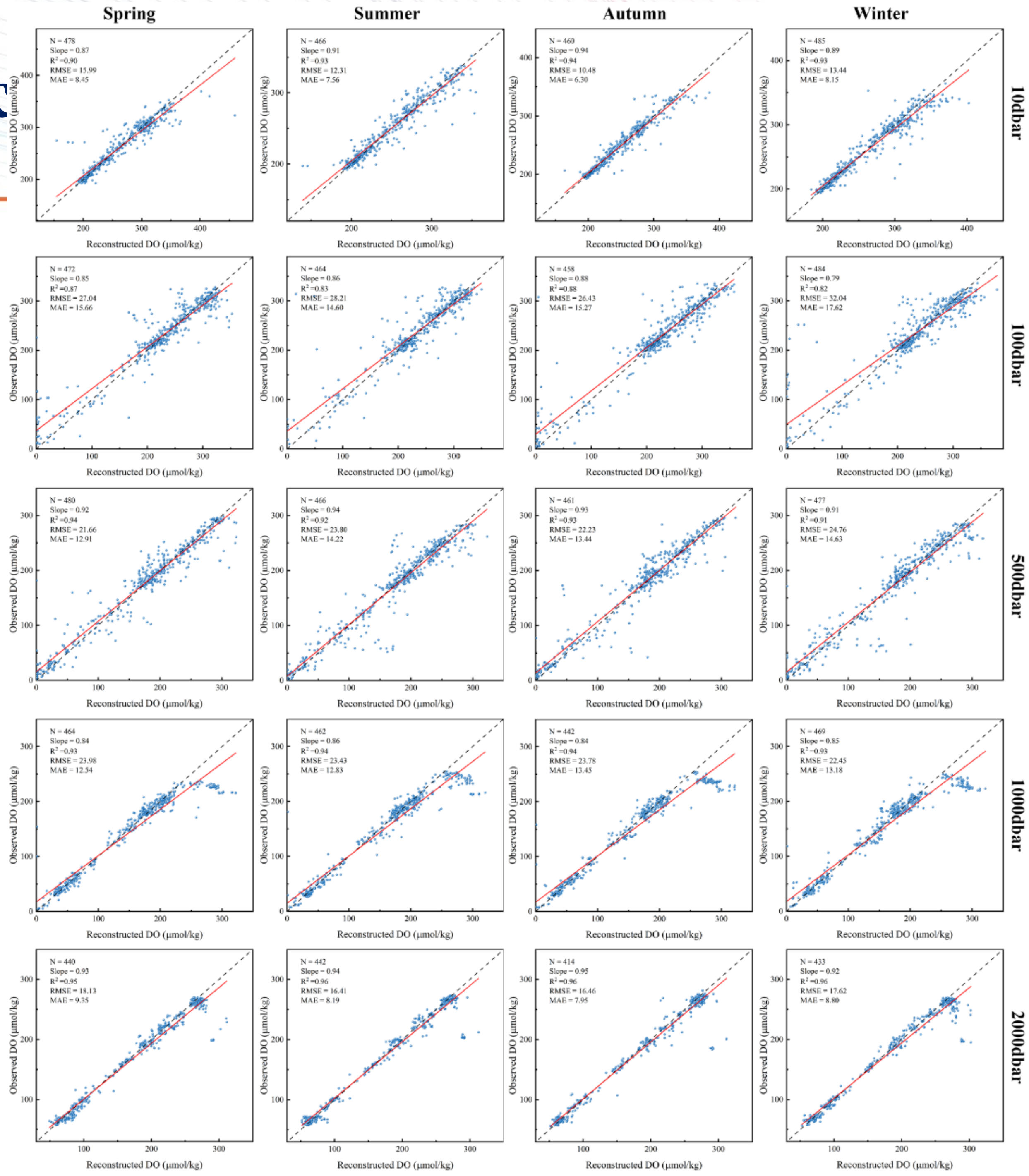


2.2 Argo profiles-based reconstruc

Model performances (Comparisons with real values)	
Depth	10 dbar、100 dbar、500 dbar、1000 dbar、2000 dbar
Month	Apr.(Spring), Jul.(Summer), Oct.(Autumn), and Jan.(Winter)

a. A strong linear relationship between a reconstructed value and a real value

b. RMSE<28 μmol/kg, MAE<18 μmol/kg, R^2 >0.8



2.2 Argo profiles-based recon

Model performances (Comparisons with WOA18)

Depth	10 dbar、100 dbar、500 dbar、1000 dbar、2000 dbar
Month	Apr.(Spring), Jul.(Summer), Oct.(Autumn), and Jan.(Winter)

- At depths of 10dbar and 2000dbar，the RE lower than 10% between them is more than 85%；
- At depths of 100dbar, 500dbar and 1000dbar，they have a low consistency with the RE lower than 10% is between 70% to 80%.

深度层	季节	<1%	1%-5%	5%-10%	>10%
10dbar	Spring	34.52	49.37	11.09	5.02
	Summer	32.62	50.00	12.45	4.94
	Autumn	36.09	52.61	8.26	3.04
	Winter	31.75	51.75	13.40	3.09
100dbar	Spring	17.37	44.07	17.16	21.40
	Summer	19.18	44.61	16.81	19.40
	Autumn	15.28	42.14	18.34	24.24
	Winter	16.70	40.50	19.21	23.59
500dbar	Spring	15.83	34.17	16.04	33.96
	Summer	13.30	36.27	19.53	30.90
	Autumn	13.67	34.71	19.52	32.10
	Winter	12.16	34.38	19.92	33.54
1000dbar	Spring	11.42	38.79	21.34	28.45
	Summer	14.07	32.47	22.51	30.95
	Autumn	12.93	32.43	22.00	32.65
	Winter	12.15	36.25	19.62	31.98
2000dbar	Spring	16.82	51.59	18.64	12.95
	Summer	23.30	47.74	15.38	13.57
	Autumn	19.08	55.56	16.67	8.70
	Winter	21.71	49.19	18.94	10.16

2.2 Argo profiles-based reconstructed models

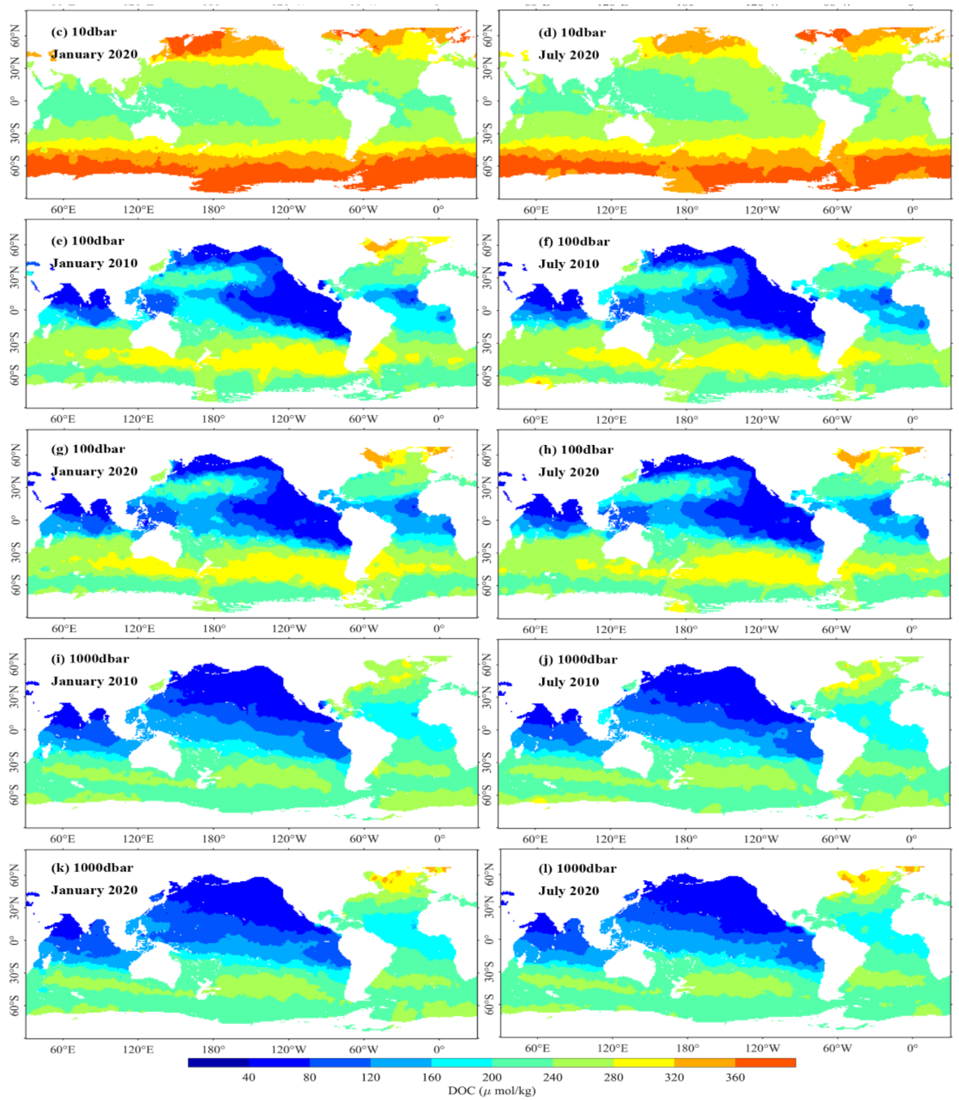
3) G4D-DOC dataset

Time	Space	Depth	Resolution	Projection	Format	Size
Jan. 2005-Dec.2022	Global	26 layers above 2000dbar	1°×1°	WGS1984	HDF、GeoTiff	1.33 GB

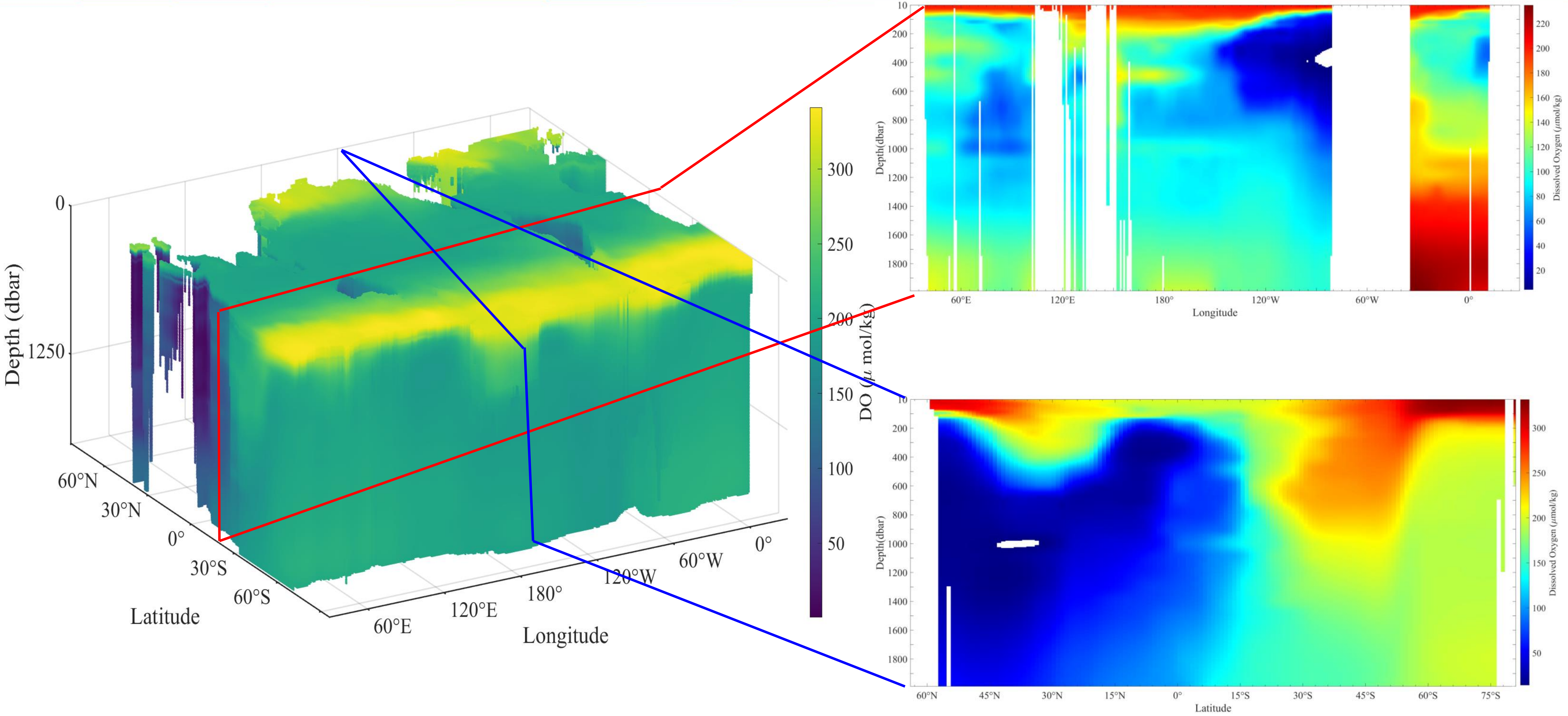
Public shared
https://doi.org/10.5281/zenodo.10147890

G4D-DOC sub-datasets

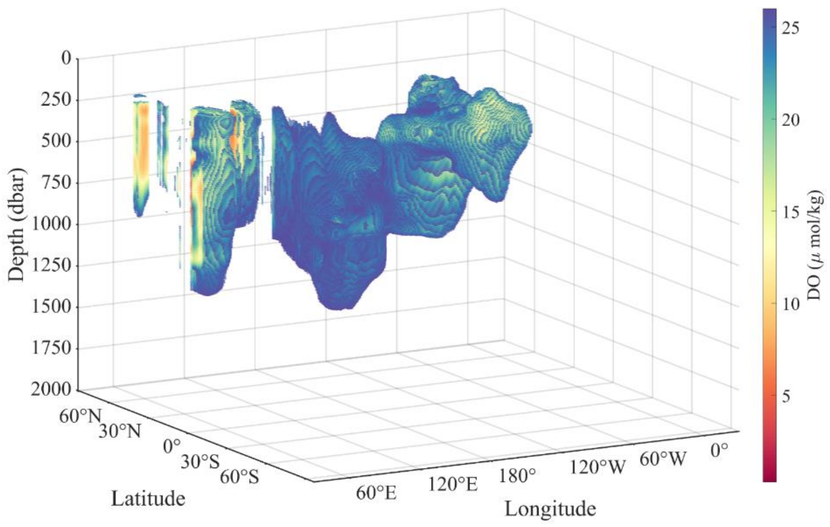
	G4D-MDOC	G4D-CMDOC	G4D-SDOC	G4D-CSDOC	G4D-ADOC	G4D-MYADOC
Temporal resolution	Monthly	Climatology monthly	Seasonal	Climatology Seasonal	Annual	Averaged Annual



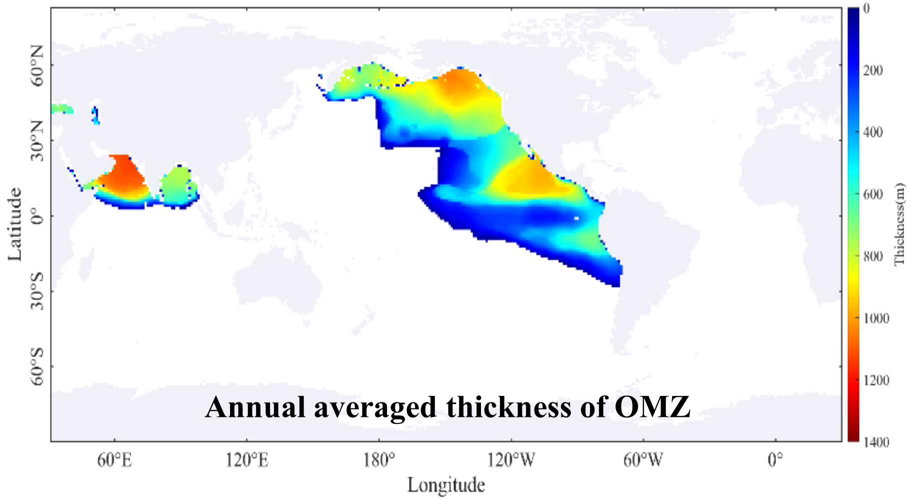
2.2 Argo profiles-based reconstructed models



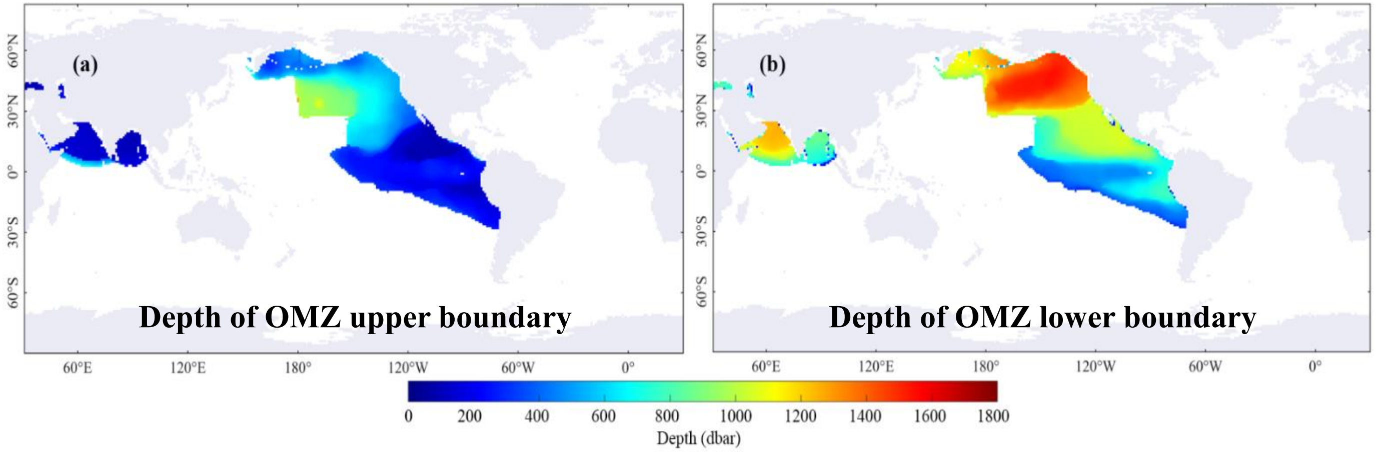
2.2 Argo profiles-based reconstructed models



Oxygen minimum zone (OMZ) distribution in space and depth



Annual averaged thickness of OMZ



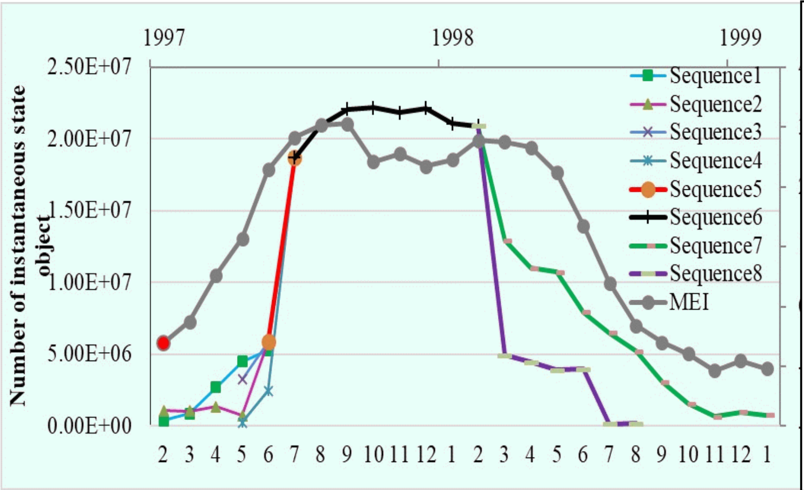
Depth of OMZ upper boundary

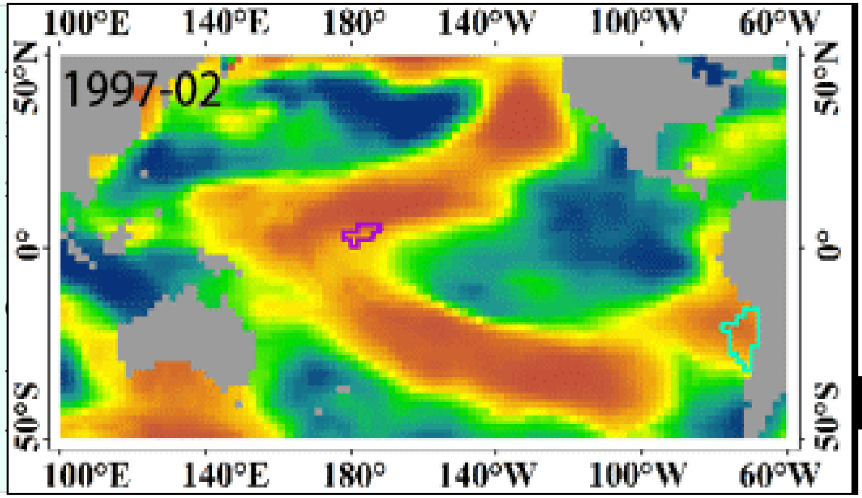
Depth of OMZ lower boundary

2.3 Process-oriented data mining models for marine dynamics

Marine abnormal variation means a variation compared to the long-term mean status, e.g. annual, seasonal, monthly, which has a property from production through development to destination.

Marine evolution behavior is the foundation of geographic prediction and early warning research as well, which closely relates to extreme climate events, and has significant impacts on global socio-economic development. Marine big data provides a data foundation for studying such evolution behaviors, but lacks analyzing methods.





Challenges

Research objects:

Object Vs

Object + behaviors

Scientific problems:

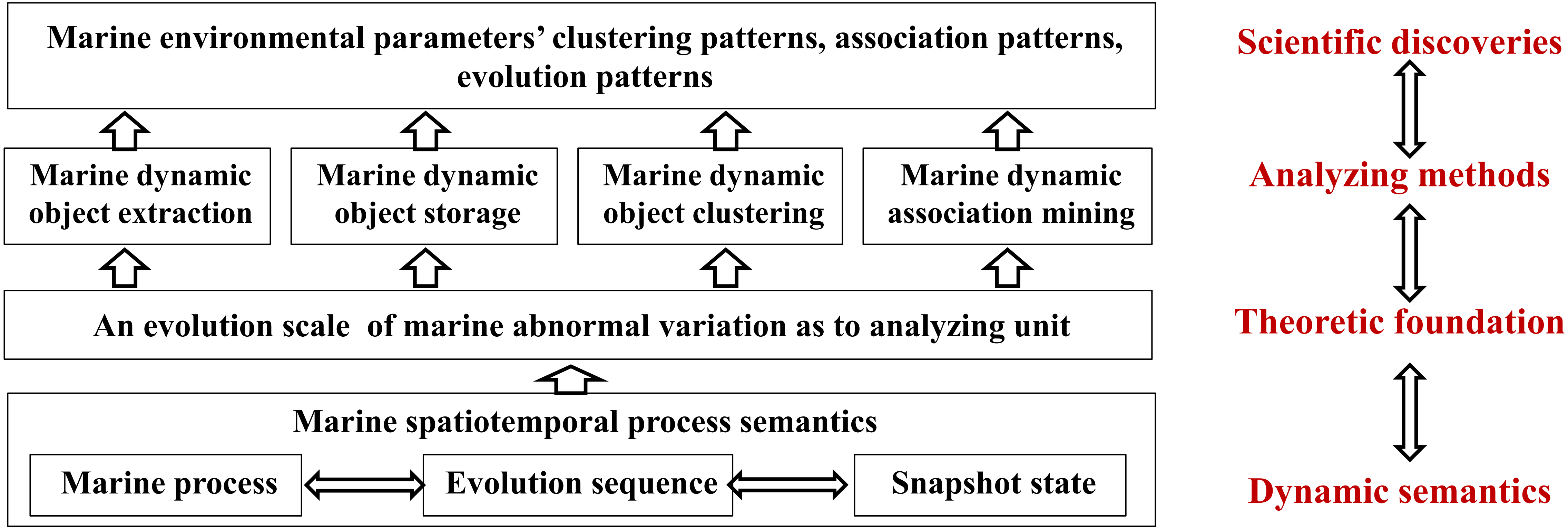
Spatiotemporal patterns Vs

Patterns + mechanisms

A novel algorithm for mining marine dynamic information

2.3 Process-oriented data mining models for marine dynamics

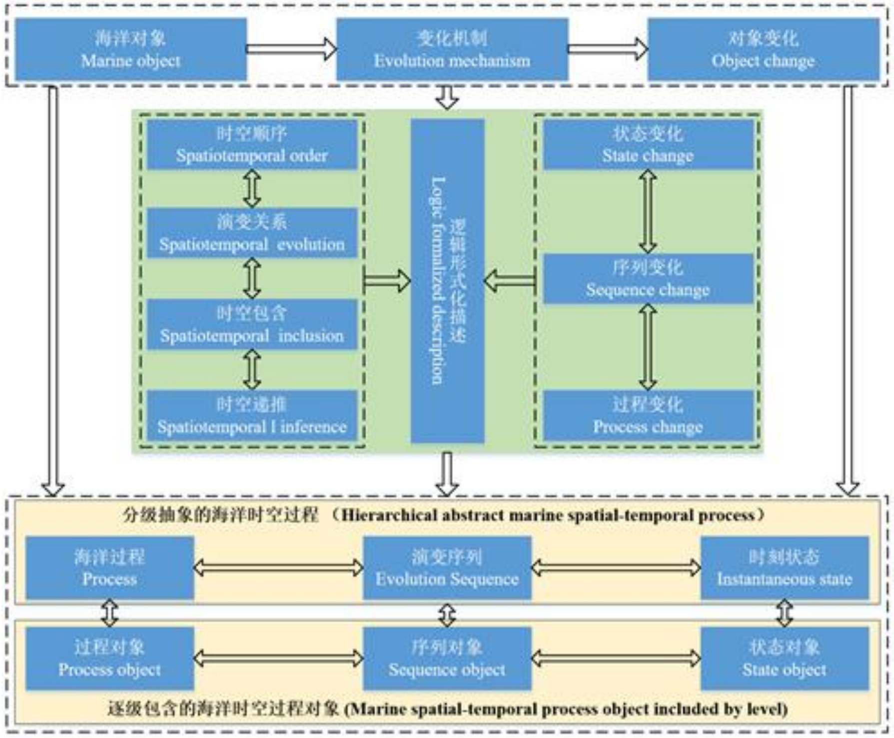
Core idea: Taking an evolution scale of marine abnormal variation as to analyzing unit to design marine spatiotemporal dynamic mining approach, including process object recognition and extraction, process object organization and management and process mining methods.

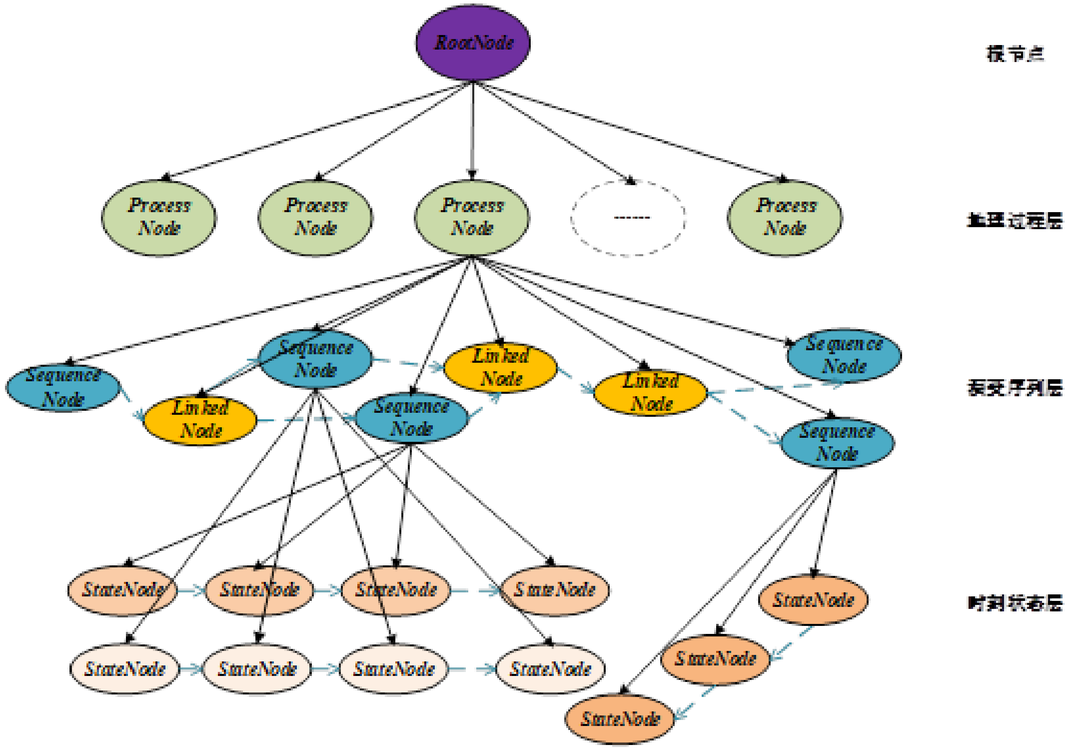


2.3 Process-oriented data mining models for marine dynamics

Problem: Lack of capabilities to represent and store evolutionary behaviors

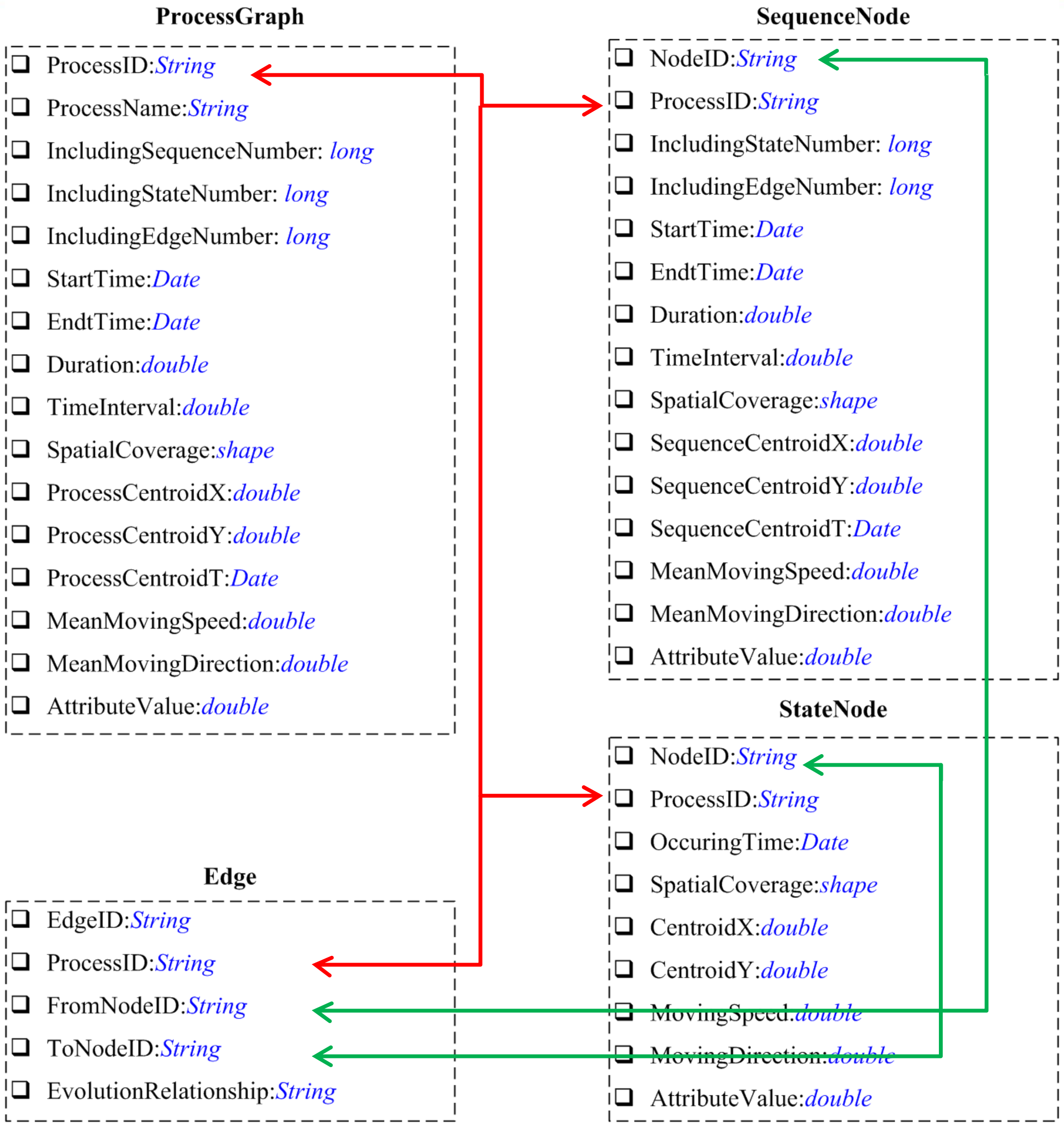
Solutions: A hierarchical and abstract bidirectional representing architecture with "marine process-evolution sequence-snapshot state" and " process object-sequence object-snapshot object" is designed, and process-oriented graph model is established, which realizes the optimal time scale to represent and store dynamic information.



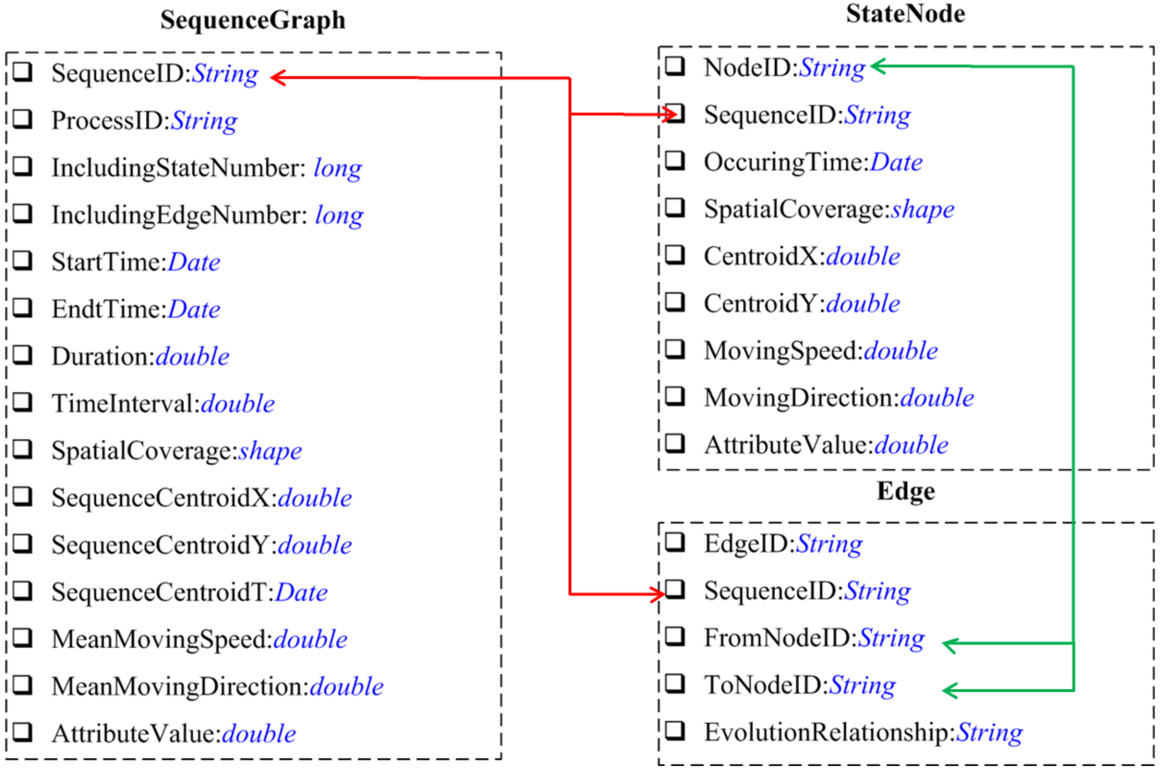


2.3 Process-oriented data mining models for marine dynamics

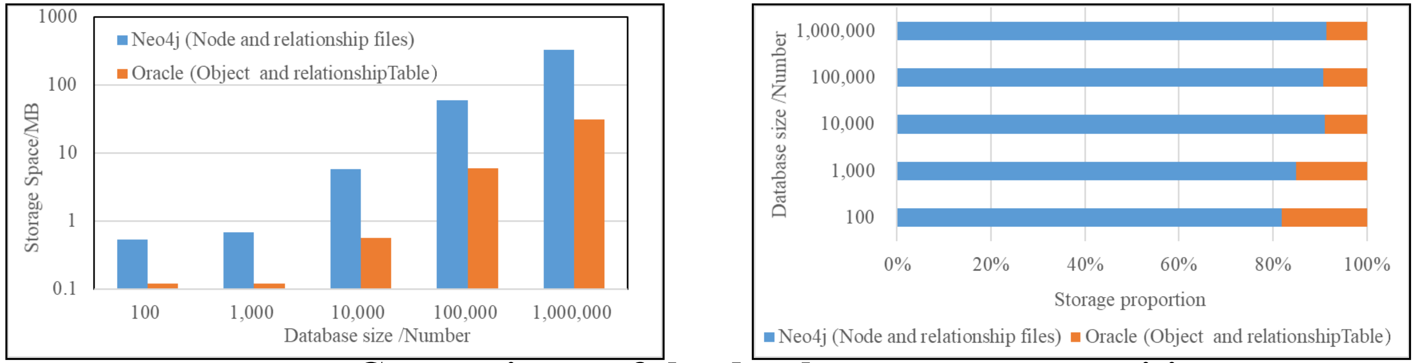
Process graph model



Sequence graph model



2.3 Process-oriented data mining models for marine dynamics



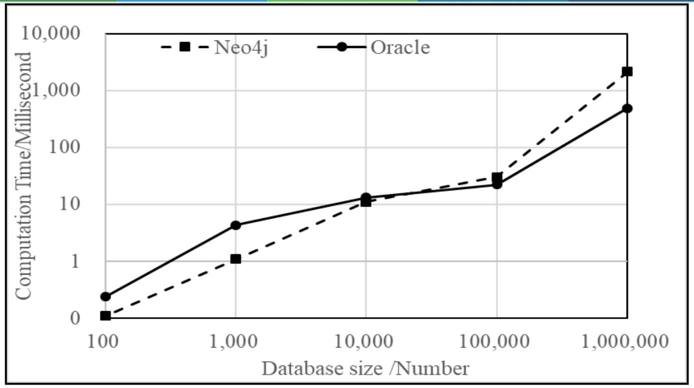
Comparisons of the database storage capacities

Regarding to database storage capacity, Oracle spatial performs better than Neo4j. While with an increase of the database size, the advantage will reduce.

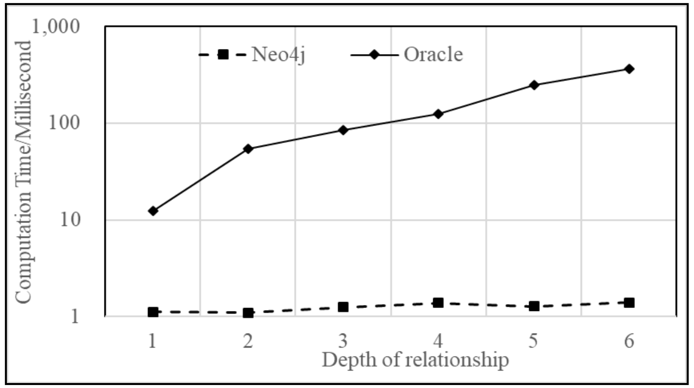
Regarding to spatial object query, Neo4j has a slight advantage due to its small database size. With an increase of the database size, Oracle spatial performs better than Neo4j.

Regarding to Evolutionary behavior query, Neo4j performs much better than the Oracle spatial database, and the performance advantage is enhanced with an increase of the depth of the relationship or database size.

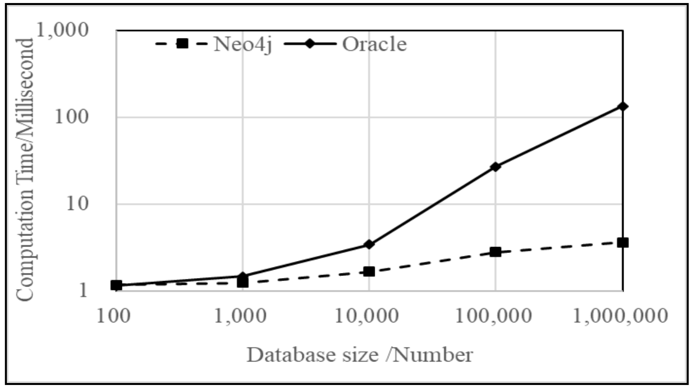
Process object



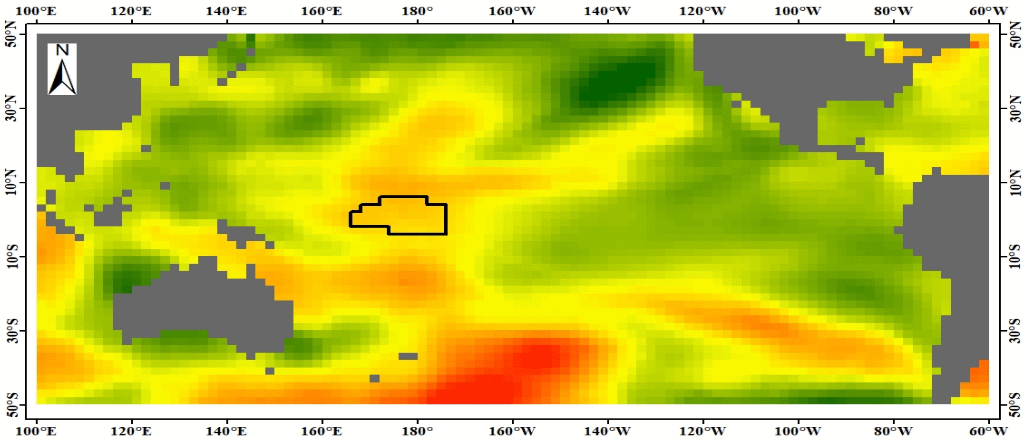
Snapshot object



Evolutionary behavior

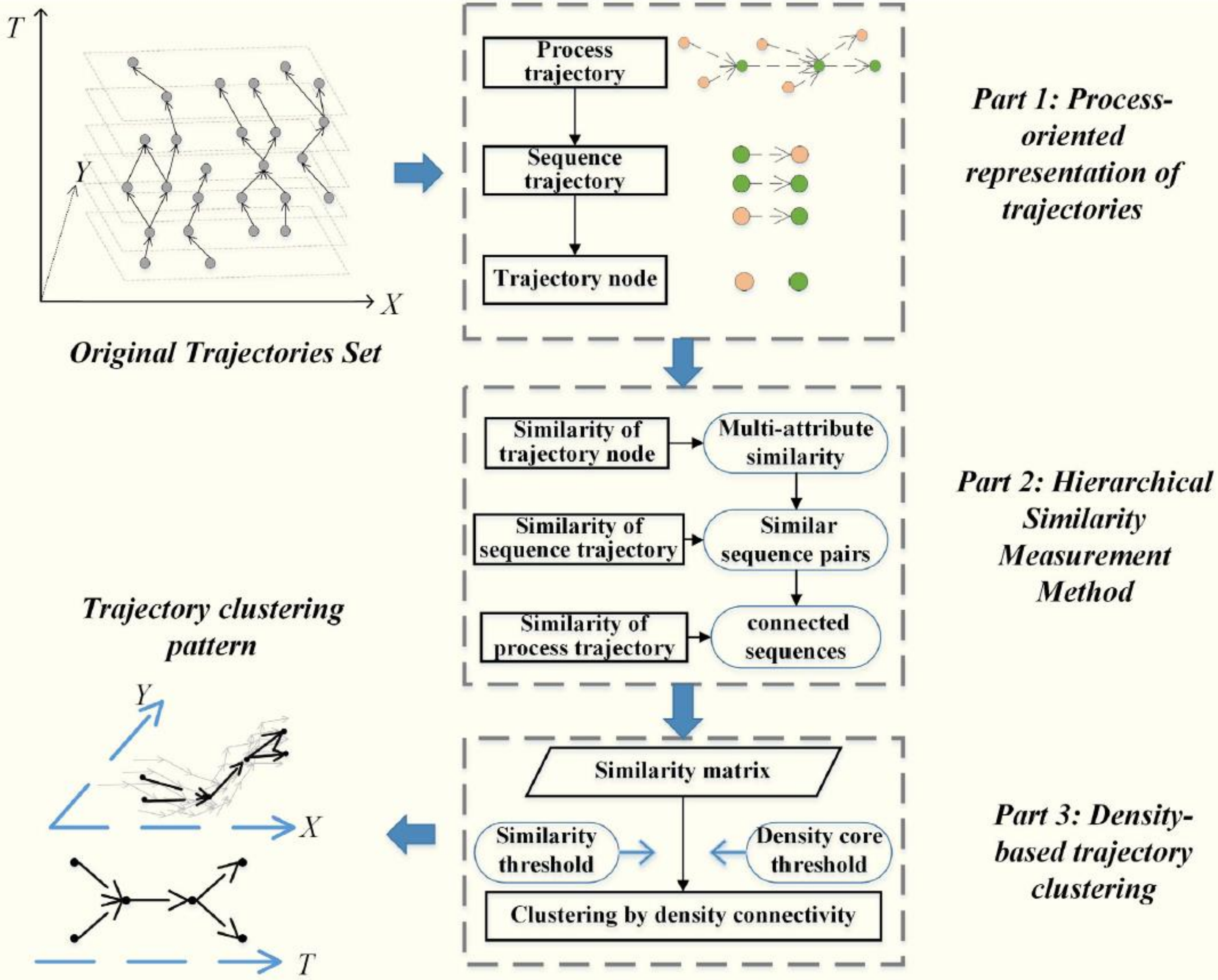


2.3 Process-oriented data mining models for marine dynamics



Problem: lacks of the ability to depict evolutionary behaviors, the spatiotemporal mining algorithms have challenges of analyzing dynamic information.

Solutions: Taking the evolution as an unit to redefine basic concepts of clustering, and design the analyzing method for objects and evolution behaviors as well。



2.3 Process-oriented data mining models for marine dynamics

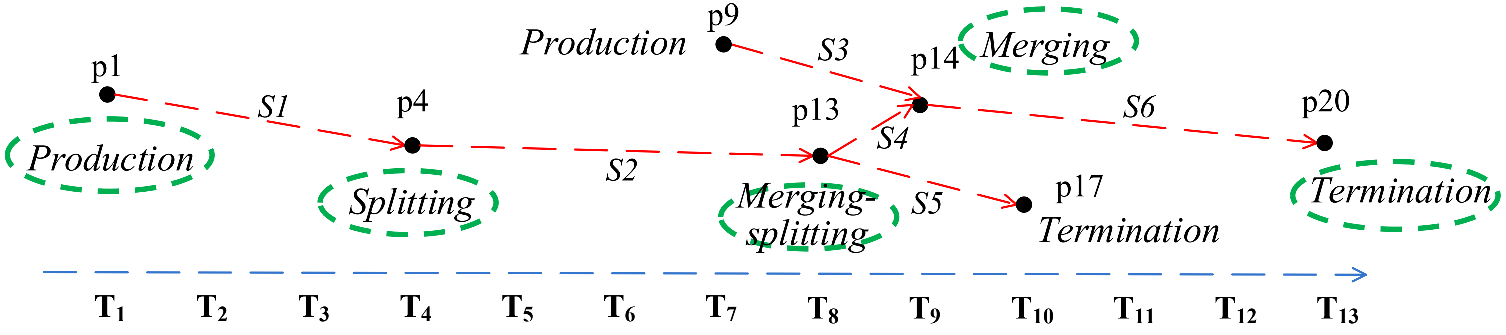
Marine process semantics:

marine process-evolution sequence-instantaneous state"

Representation of trajectory

process trajectory- sequence trajectory - trajectory node

$$\begin{cases} MSTP = \{ES_1, ES_2, \dots, ES_n\} & n \geq 1 \\ ES_i = \{S_i^S, S_i^E\} & 1 \leq i \leq n \\ S_i^{S/E} = \left(\{ESf_i^{S/E}\}, \{EMf_i^{S/E}\}, \{EAf_i^{S/E}\}, T_i^{S/E}, StateType_i^{S/E}\right) \end{cases}$$



Tra (S1, S2, S3, S4, S5, S6) $\rightarrow$ *S1 (p1, p4)* $\rightarrow$ *p1* (x_1, y_1, T_1, *attributes*, ***production***)

p4(x_4, y_4, T_4, *attributes*, ***splitting***)

2.3 Process-oriented data mining models for marine dynamics

Process-oriented similarity measurement

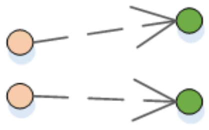
Level-1

space similarity : Euclidean distance ;
time similarity: relative /absolute distance;
structure similarity: with the same node type

Similarity of trajectory node

Level-2

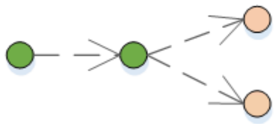
starting and ending trajectory nodes are similar at the same time



similar pairs

Similar sequence trajectory pairs

Level-3



MaxNum: number of connected sequences

$$Similarity\ (Tra_m, Tra_n) = MaxNum\ /\ (M + N - MaxNum)$$

Similarity of complex trajectory

MaxNum: The connected sequence node number

M and N: The sequence number of $Traj_m$ and $Traj_n$

2.3 Process-oriented data mining models for marine dynamics

Simulated dataset

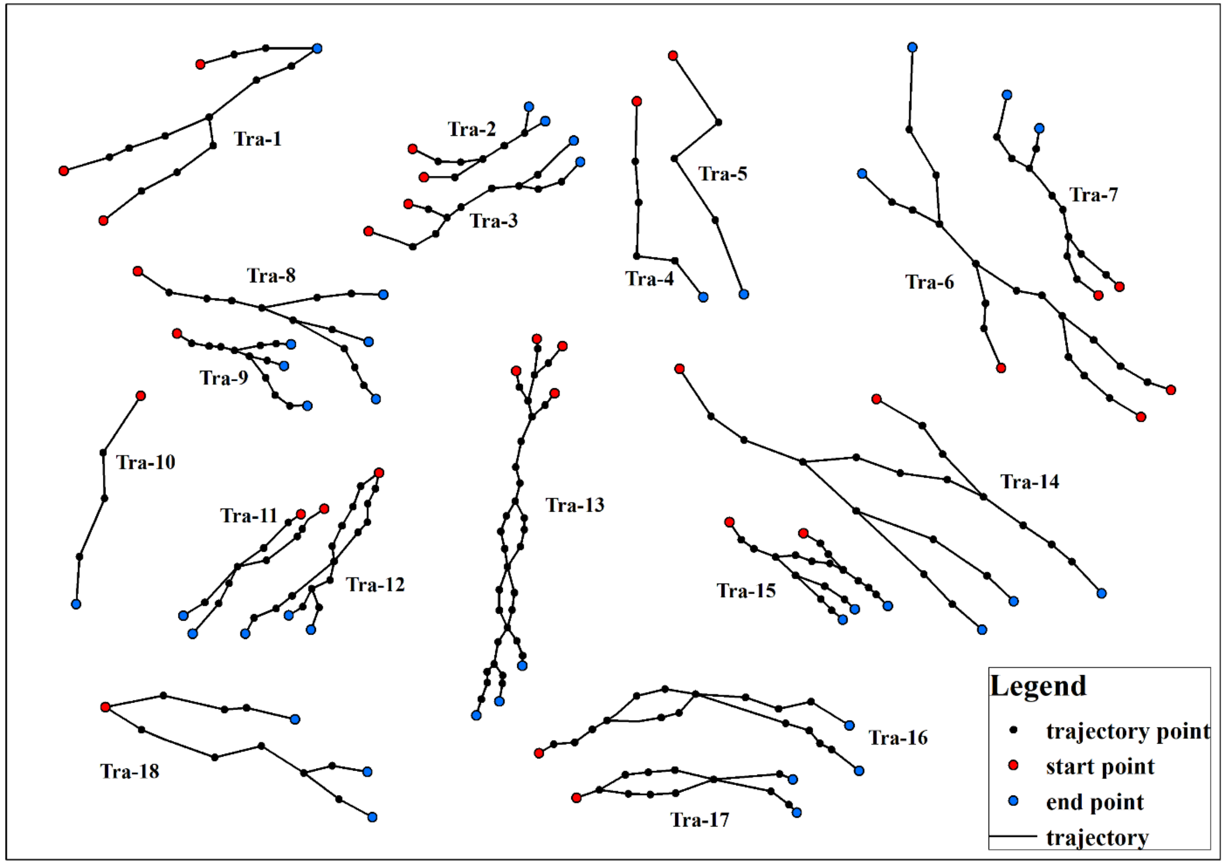
- 18 trajectories with
- 285 trajectory points including
- 4 kinds of structures and
- a most complex trajectory (Tra-13).

Cluster1: Tra-1/-2/-3/-6/-7

Cluster2: Tra-4/-5/-10

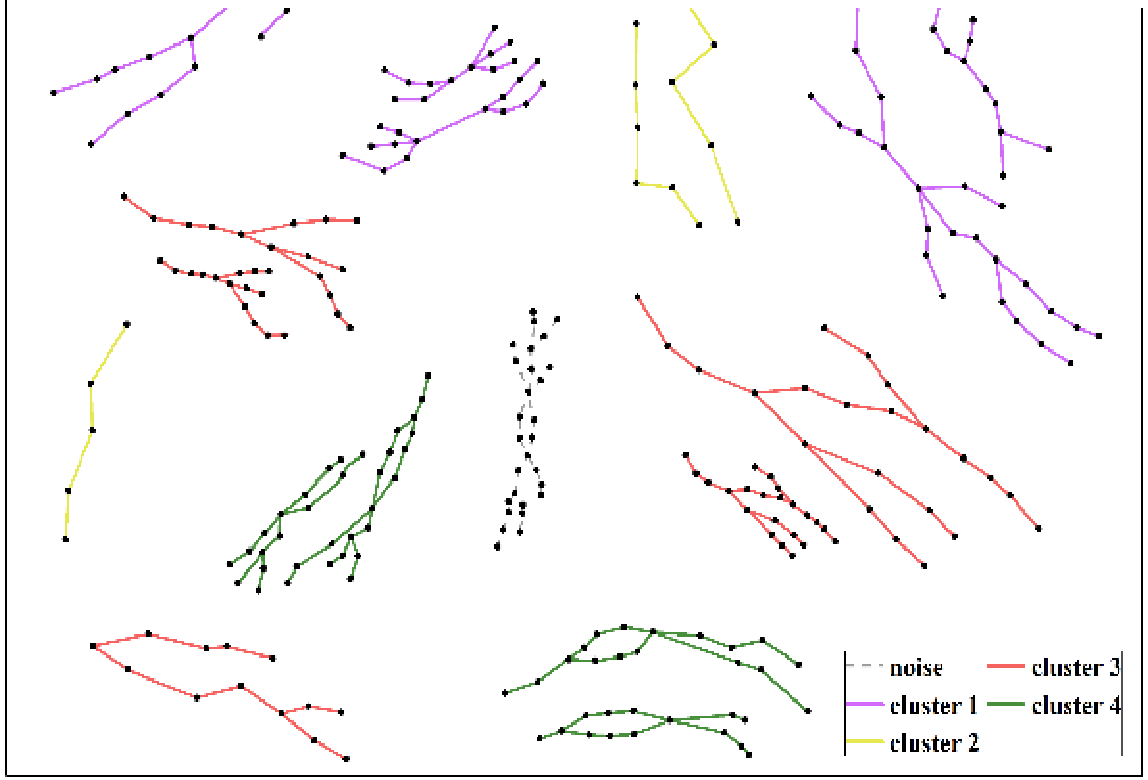
Cluster3: Tra-8/-9/-14/-15/-18

Cluster4: Tra-11/-12/-16/-17

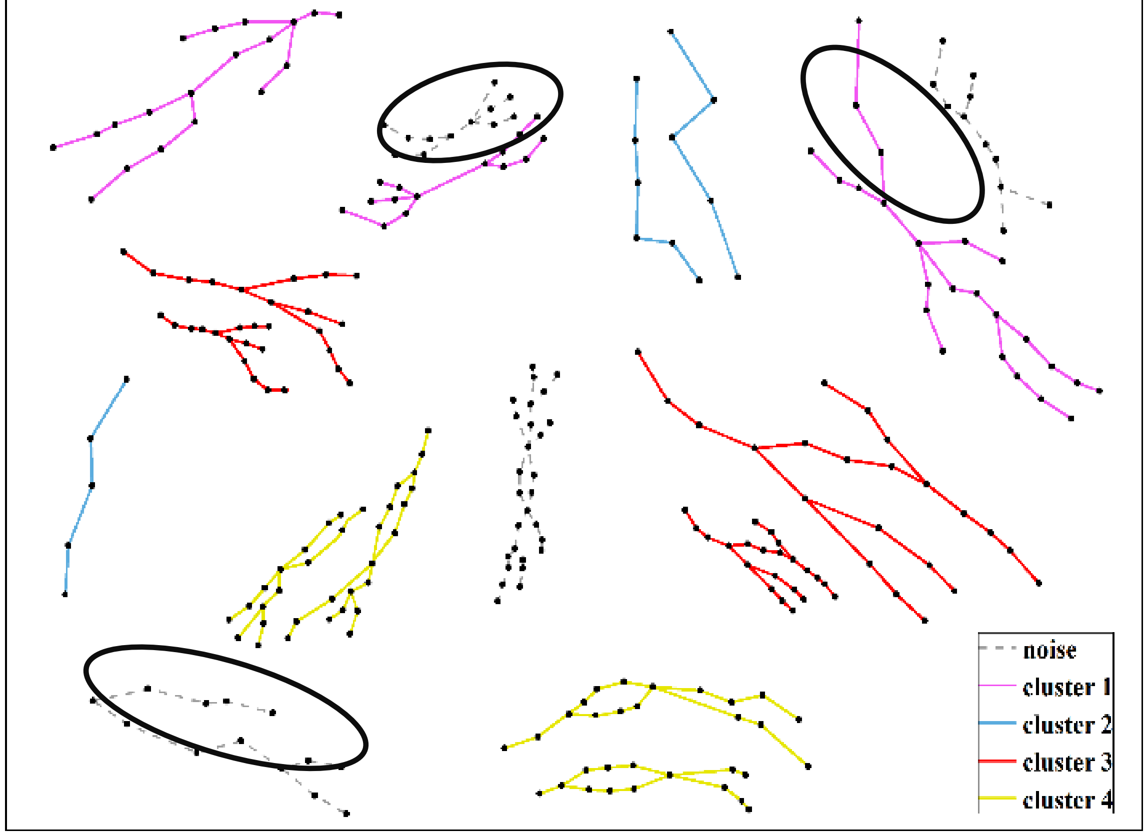


2.3 Process-oriented data mining models for marine dynamics

Our proposed approach： time complicity $O(N^2 * M^4)$， PCVI=1.82



VF2： time complicity $O(N! * N)$ ， PCVI=1.71



Regard to clustering capability: Our approach performs better than VF2, especially in conditions where some evolutionary structures are similar

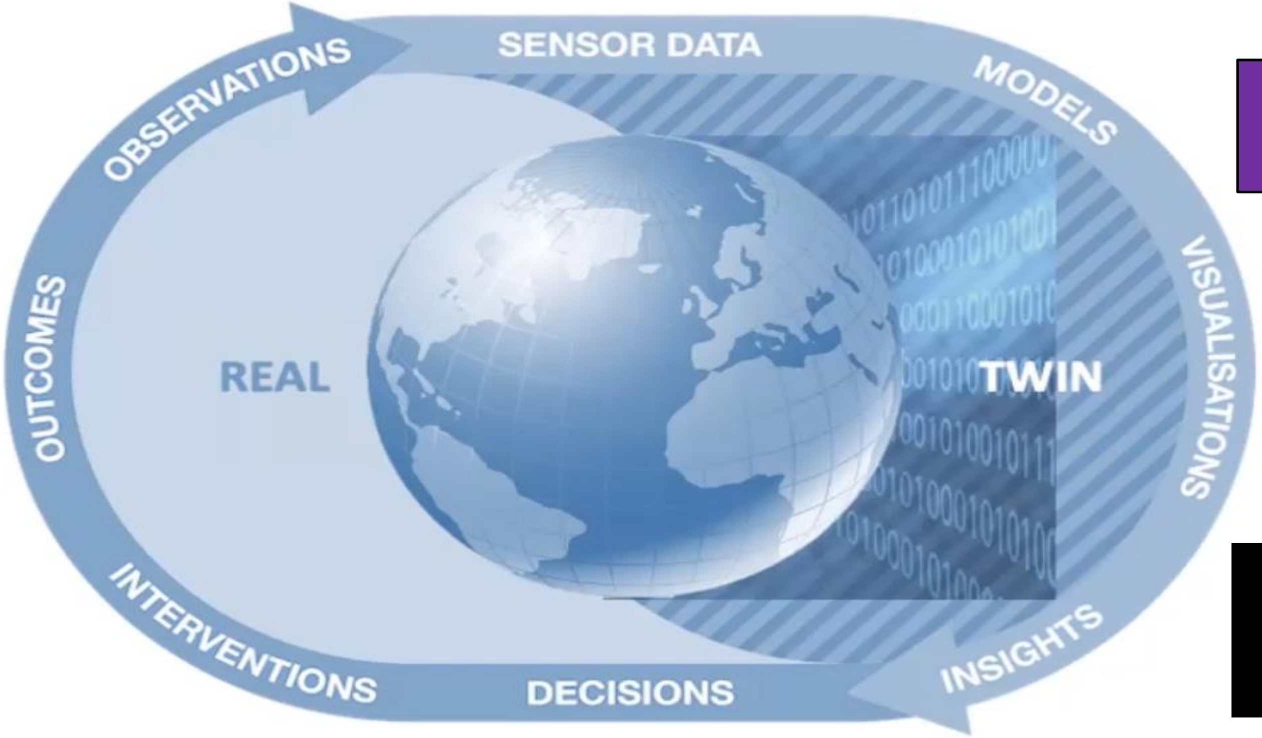
Regrading to calculating capability： Our approach performs better than VF2

Outlines

- Challenges and demands from Marine SDGs
- Marine data-added mining technologies
- Digital twin of the ocean promotes 'what-if' scenarios simulations
- Marine information system for marine SDGs
- Case studies of monitoring and evaluating marine SDGs

What is the Digital Twins of Ocean

A Digital Twin of the Ocean (DTO) is a virtual representation of the real ocean that has a two-way connection with it. Observations from the real ocean, in combination with models, data science and AI, are used to create a digital twin that adapts as the real world changes. Manipulating the twin to address 'what if' scenarios can provide information for decision-making and highlight regions of the real ocean in need of better or different observations（DITTO, July 2023）。



Decision-making

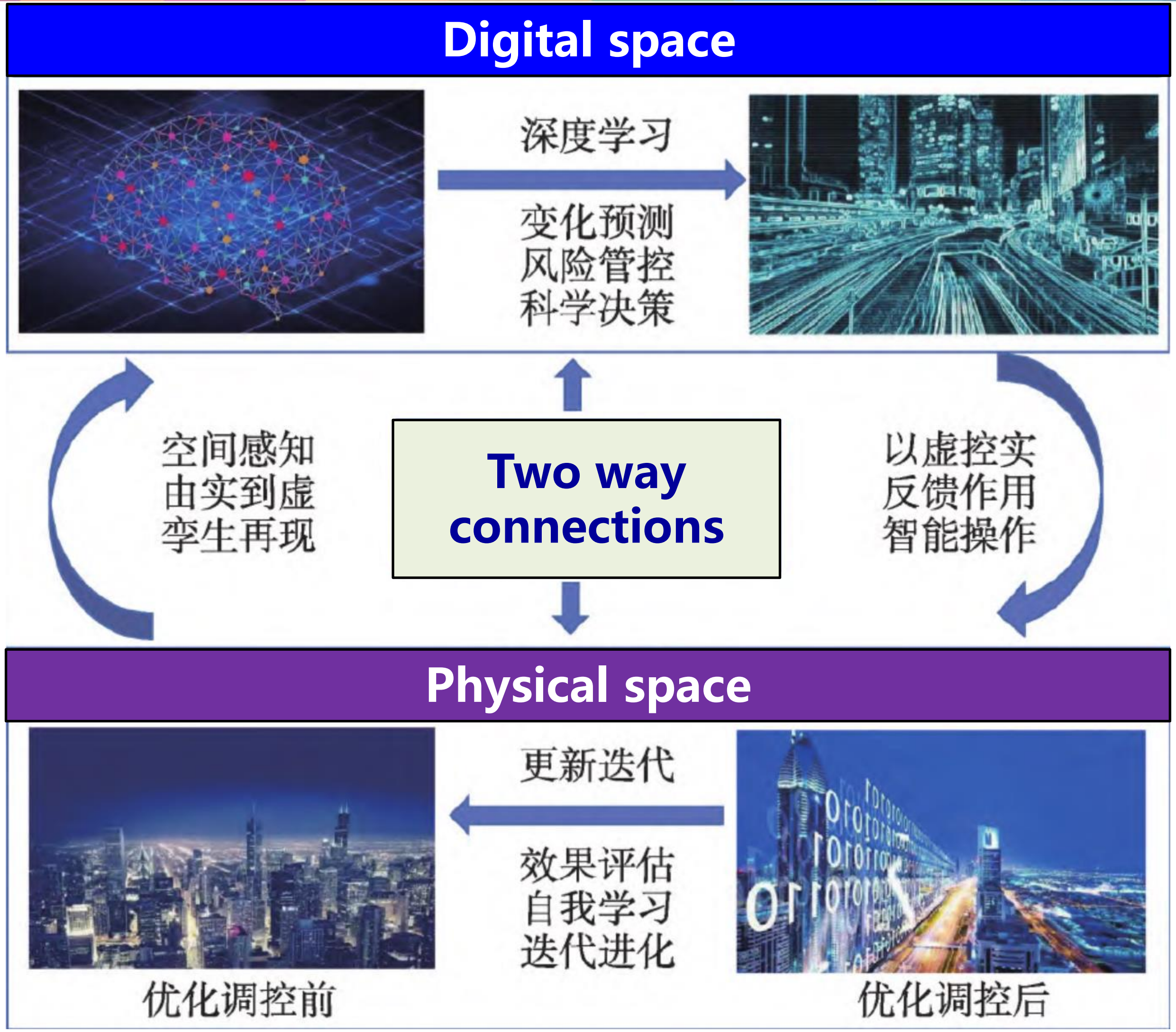
Interventions on real ocean

'what if' scenarios

A virtual representation of the real ocean

What is the Digital Twins of Ocean Framework

Digital twin of physical space： With the support of information collection and transmission technologies, the elements, relationships, processes, and patterns of natural entity (physical space) are mapped into a virtual space, thereby constructing a digital space, in which intelligent control such as simulation, simulation, reconstruction, regulation, and optimization can be done for natural entity space （Shuangcheng LI，2022）



What is the Digital Twins of Ocean Framework

Triple representing model of Digital Twins of the Ocean： Physical space, Digital space, and the interrelationships among them

$$\Omega = (P, D, R)$$

P: Physical space of the ocean: elements, spatiotemporal patterns, and their variations in the natural world.

D: Digital space of the ocean: objects, functions and relationships based on models in the digital world

R: Interrelationships: the mapping relationship from physical to digital space, the optimization and decision-making from digital to physical space, and the correlation relationship between physical and digital space

Outlines

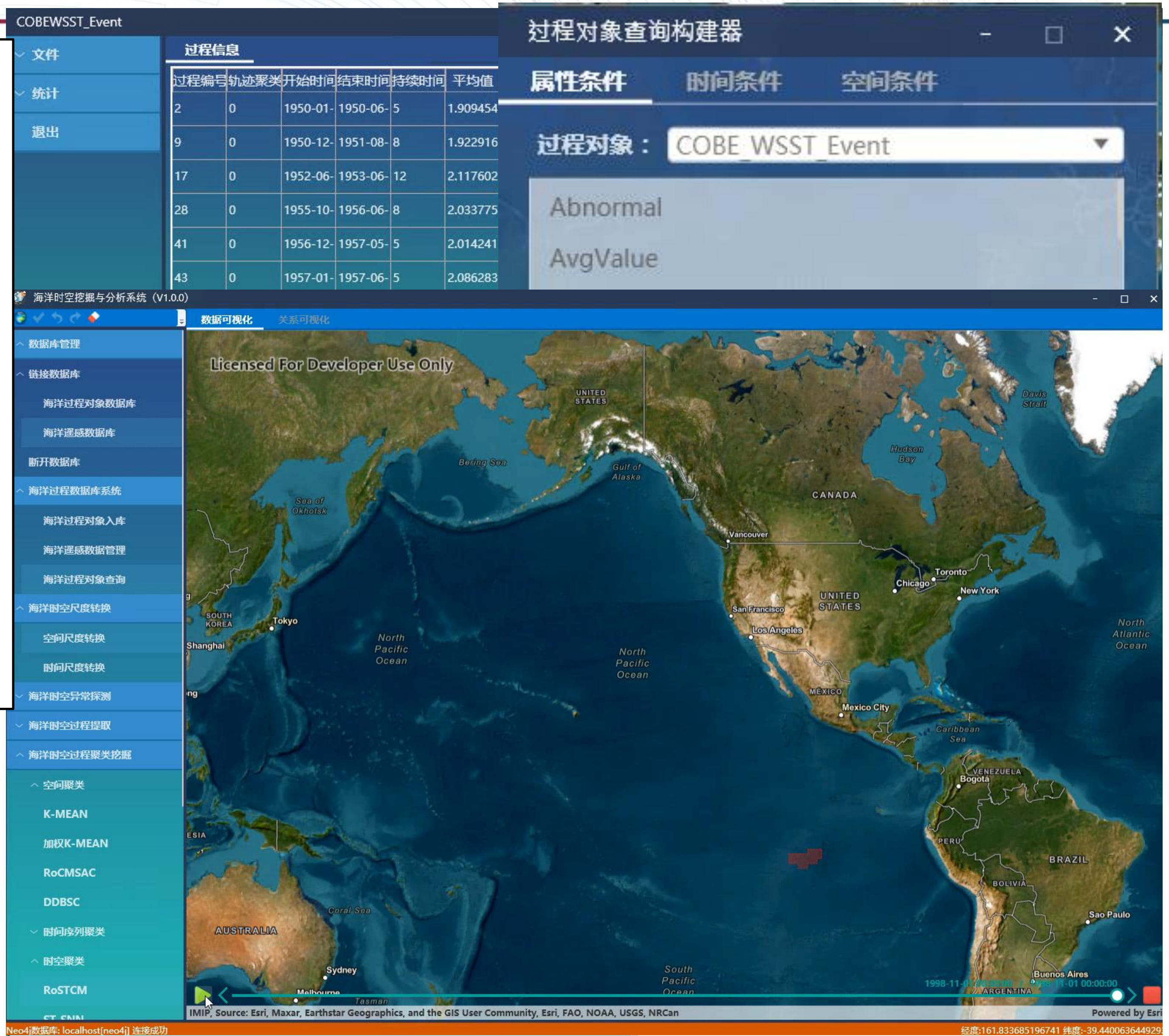
- Challenges and demands from Marine SDGs
- Marine data-added mining technologies
- Digital twin of the ocean promotes 'what-if' scenarios simulations
- Marine information system for marine SDGs
- Case studies of monitoring and evaluating marine SDGs

Marine information system

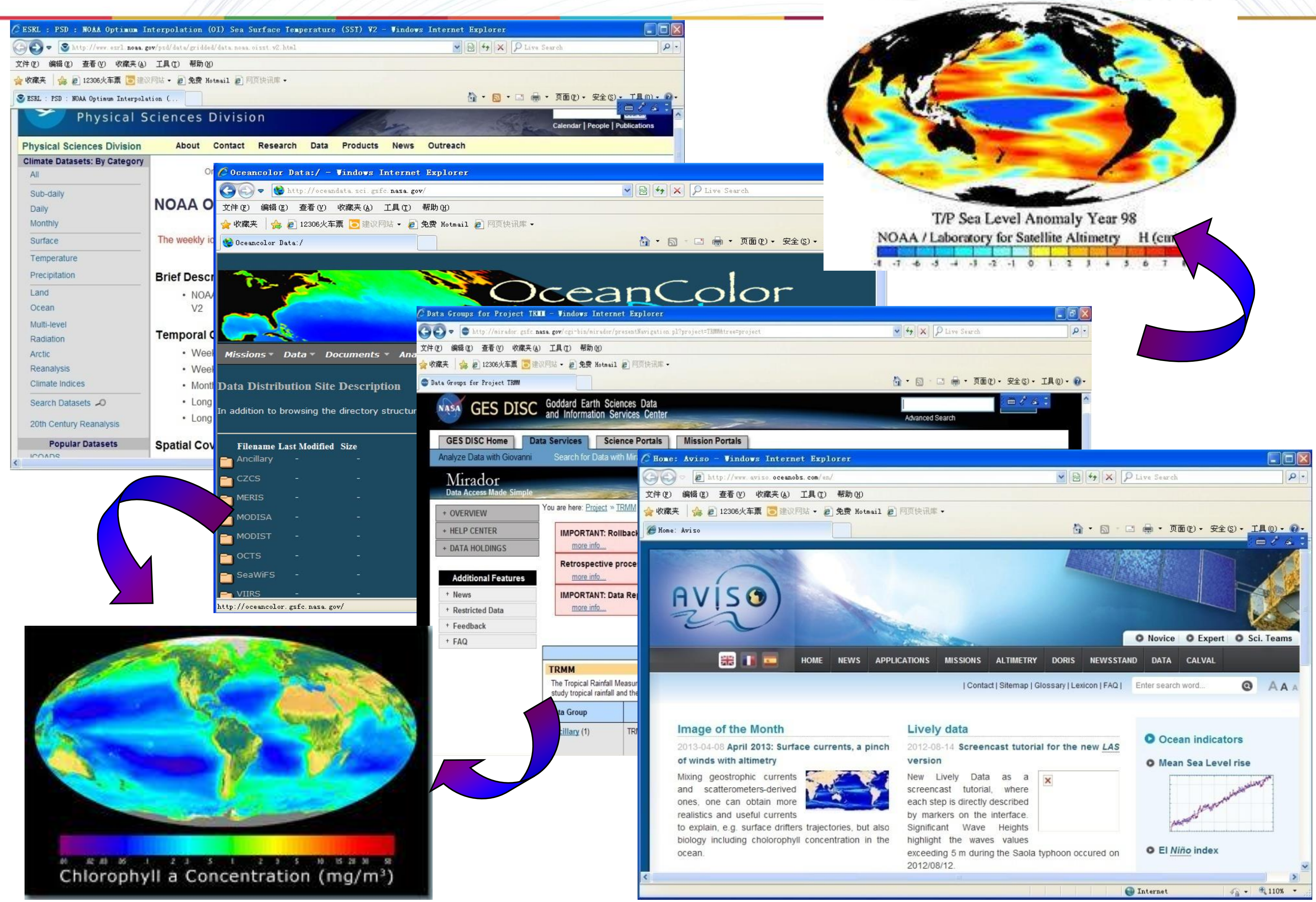
- **Marine data processing**
  a. A custom unified dataset
  b. Transformation of four temporal and five spatial scales
  c. Marine process object extraction from six kinds of marine parameters
  d. Capability of dealing with more than 70 years datasets
- **Marine spatiotemporal pattern mining**
  a. Capability of clustering six kinds of marine parameters
  b. Capability of abnormal detecting six kinds of marine parameters
  c. Capability of rule mining seven kinds of marine parameters
  d. Capability of dealing with more than 20 marine objects

Marine information system

- **Marine data management & dynamic visualization**
  - a. Process-oriented data storage
  - b. Process-oriented object inquiry
  - c. Time-Space-Thematic inquiry
  - d. Object visualization
  - e. Dynamic visualization
  - f. Evolutionary relationship visualization
  - g. Associated visualization of objects and relationships



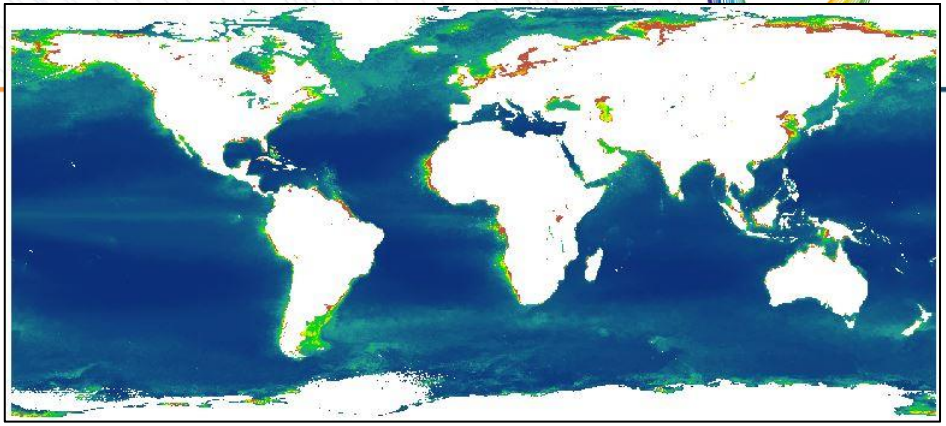
Marine information system- Unified dataset

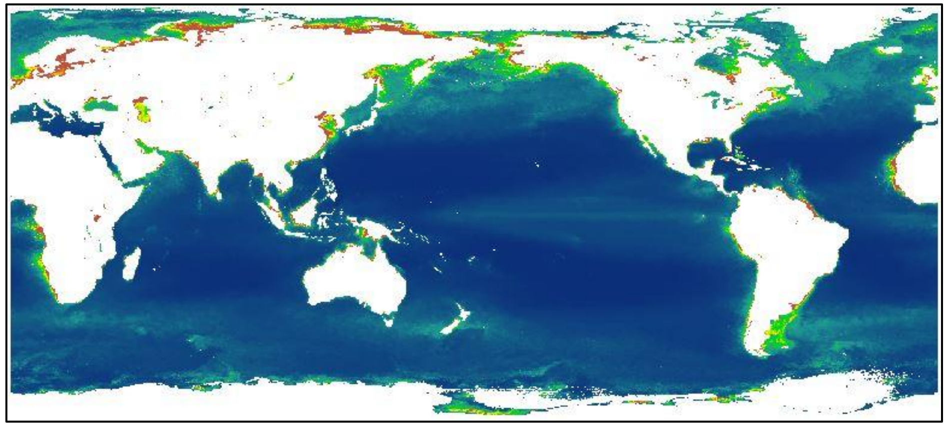


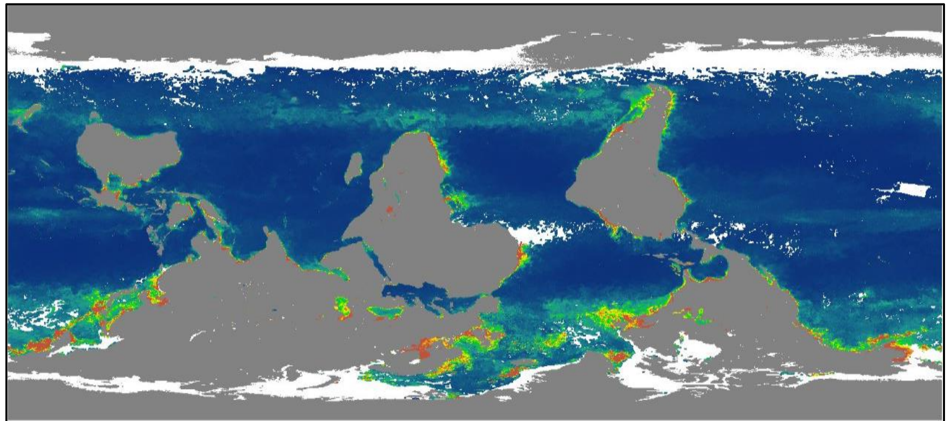
- Marine bio-optical Parameters
- Marine dynamic parameters
- Global coverage
- From surface to middle-depth layer
- From kilometers to meters
- Capabilities of continuous update

Marine information system- Unified dataset

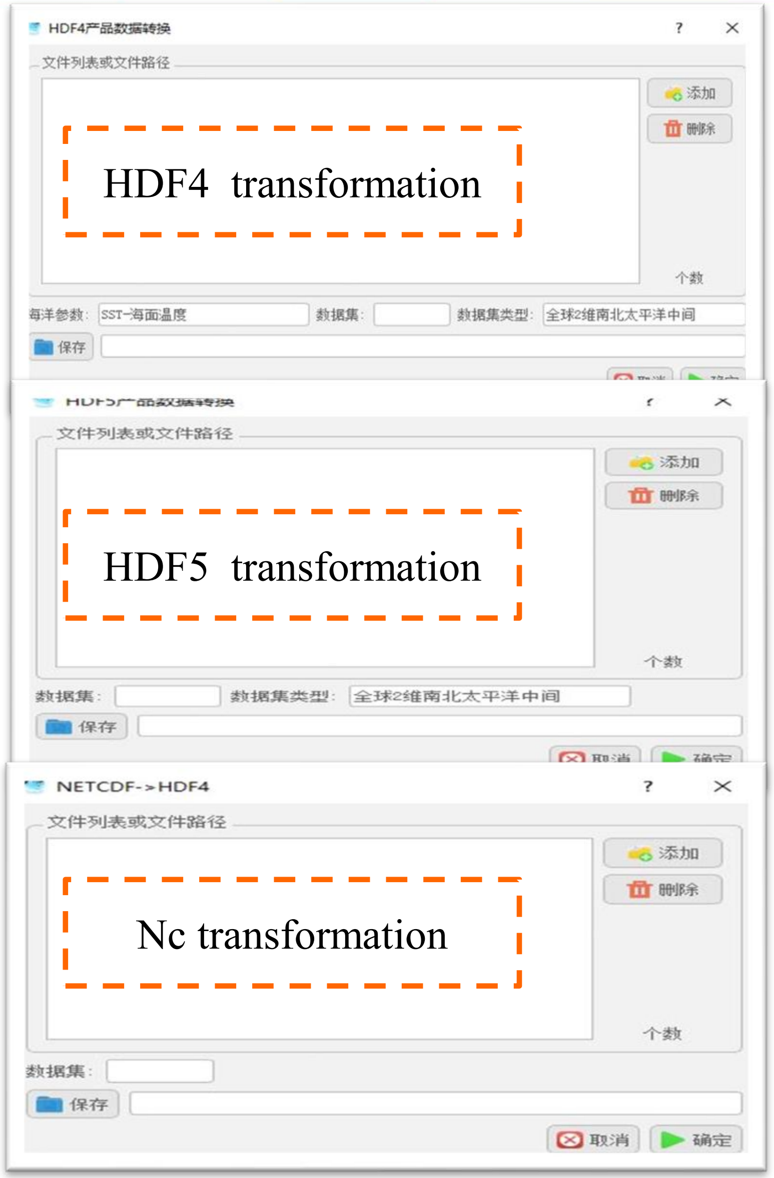
Parameters	Sources	Spatial Resolution	Format
SST	MODIS/ERSST	4km、9km、0.5° 、1.0°	.nc /.hdf4
SSP	TRMM/GPM	0.25°	.hdf4 /.hdf5
SLA	AVISO	0.25°	.hdf4
Chla	MODIS/SeaWifs	4km、9km	.hdf4
SSS	SMOS/SMAP	0.25°	.nc /.hdf4
SSW	CCMP/QuikSCAT	0.25°	.hdf4

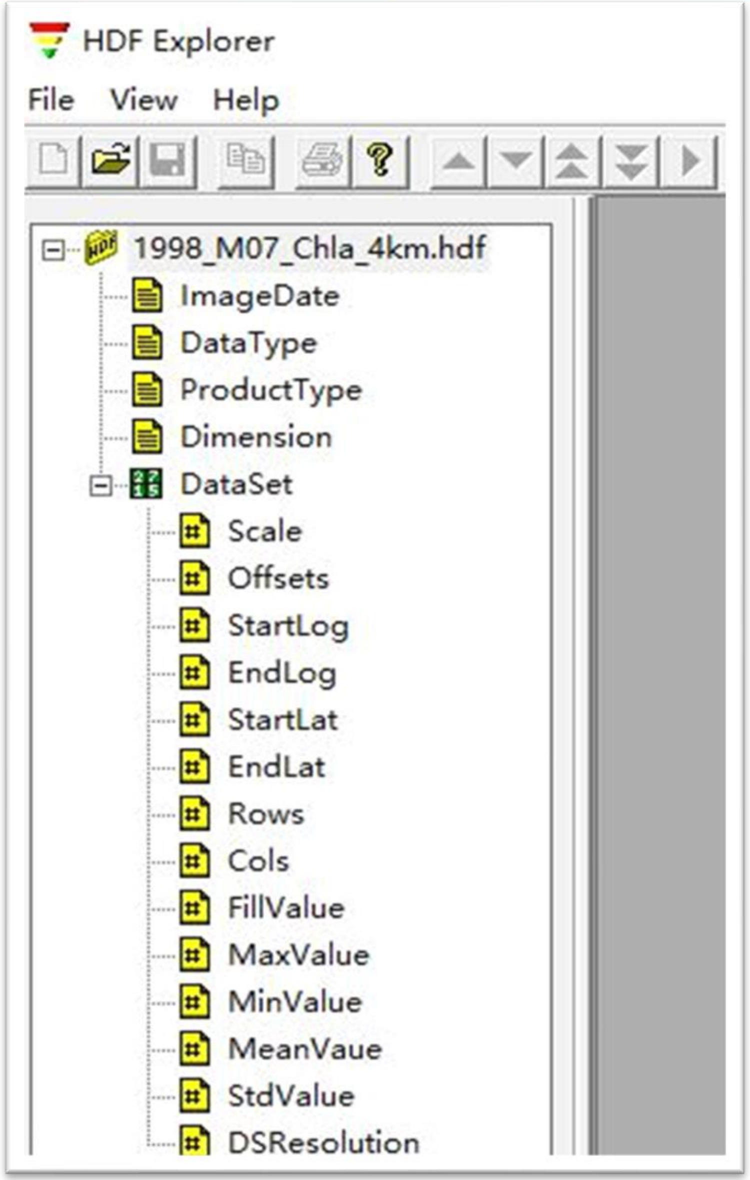


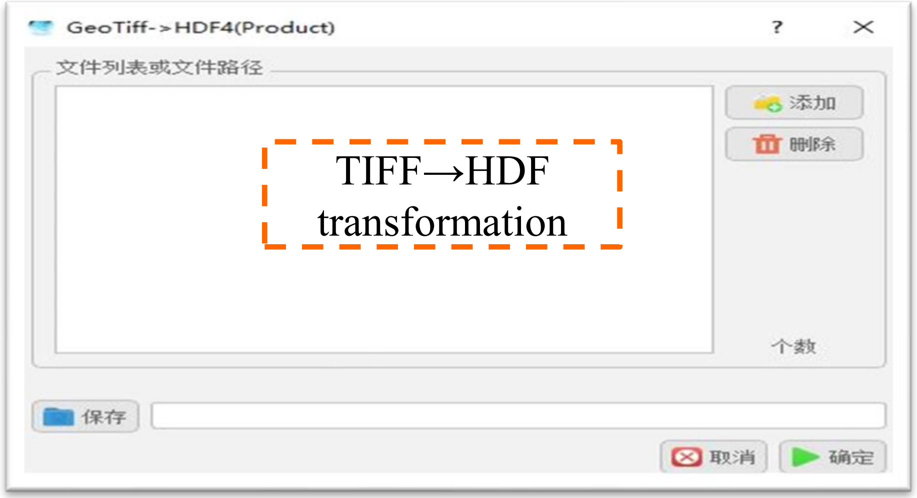


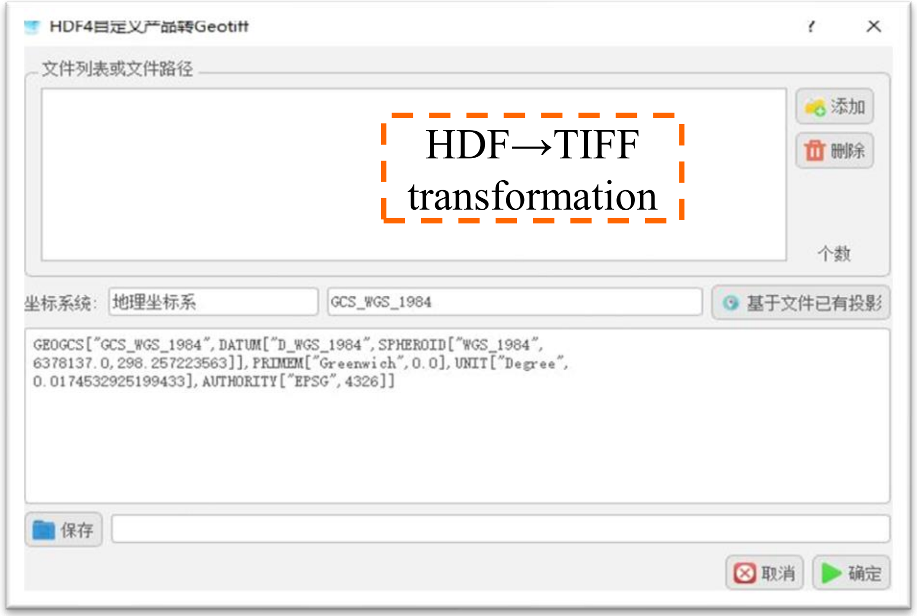


Marine information system- Unified dataset





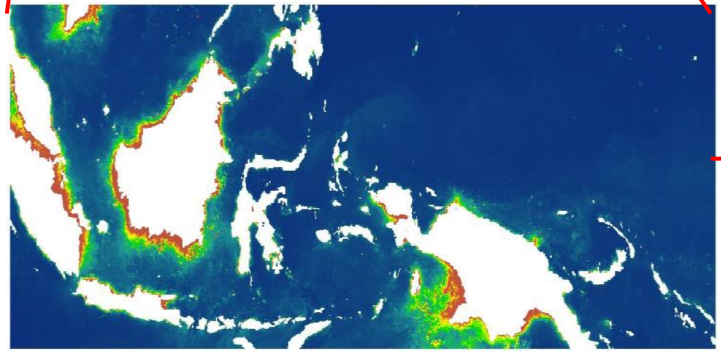




Marine information system- Unified dataset

序号	产品类型	产品名称	空间分辨率	时间分辨率	更新周期	投影方式	数据类型	数据格式	地域范围	数据量	提交时间	共享方式	时间范围
1	遥感融合产品	海表叶绿素a浓度	4km	月/季/年	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	319GB	2019	开放共享	1998-2021年
2	遥感融合产品	海洋表面降雨	0.25度	月/季/年	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	5.71GB	2019	开放共享	1998-2021年
3	遥感制作产品	重点区域悬浮颗粒浓度	1km	月/季/年	1年	UTM投影	栅格	GeoTIFF	黄河口、密西西比河、尼罗河	3.75GB	2019	开放共享	2002-2020年
4	遥感制作产品	重点区域黄色物质	1km	月/季/年	1年	UTM投影	栅格	GeoTIFF	黄河口、密西西比河、尼罗河	3.75GB	2019	开放共享	2002-2020年
5	遥感再分析产品	海洋表面温度距平	4km	月/季/年	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	283GB	2020	开放共享	1982-2021年
6	遥感再分析产品	海面降雨距平	0.25度	月/季/年	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	4.73GB	2020	开放共享	1998-2021年
7	遥感再分析产品	海面高度异常距平	0.25度	月/季/年	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	5.71GB	2020	开放共享	1993-2021年
8	遥感再分析产品	海表叶绿素a浓度距平	4km	月/季/年	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	4.73GB	2020	开放共享	1998-2021年
9	遥感再分析产品	海洋初级生产力距平	4km	月/季/年	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	42.5GB	2020	开放共享	1998-2021年
10	遥感再分析产品	海表盐度距平	25km	月/季/年	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	2.41GB	2020	开放共享	2010-2021年
11	遥感再分析产品	海洋表面温度异常变化对象集	4km	月	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	380GB	2021	开放共享	1998-2019年
12	遥感再分析产品	海面降雨异常变化对象集	0.25°	月	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	6.12GB	2021	开放共享	1998-2019年
13	遥感再分析产品	海面高度异常变化对象集	0.25°	月	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	3.06GB	2021	开放共享	1993-2019年
14	遥感再分析产品	海表叶绿素a浓度异常变化对象集	4km	月	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	220GB	2021	开放共享	1998-2019年
15	遥感再分析产品	海表盐度异常变化对象集	0.25°	月	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	3.06GB	2021	开放共享	2010-2019年
16	遥感再分析产品	海洋初级生产力异常变化过程对象数据集	9km	月	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	55GB	2021	开放共享	1998-2019年
17	再分析产品	全球气候态月尺度溶解氧数据集	1度	气候态月尺度	1年	经纬格网	栅格	GeoTIFF、HDF4	全球	135MB	2022	开放共享	2010-2021年

Marine information system- Spatiotemporal statistic

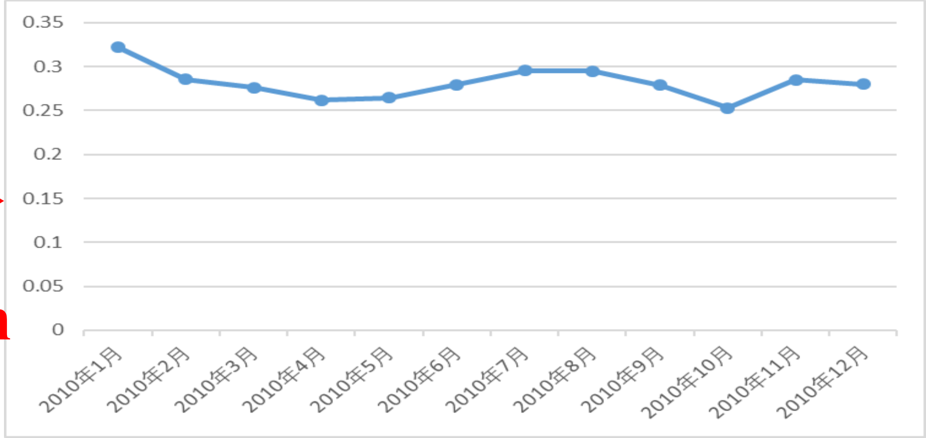


Max

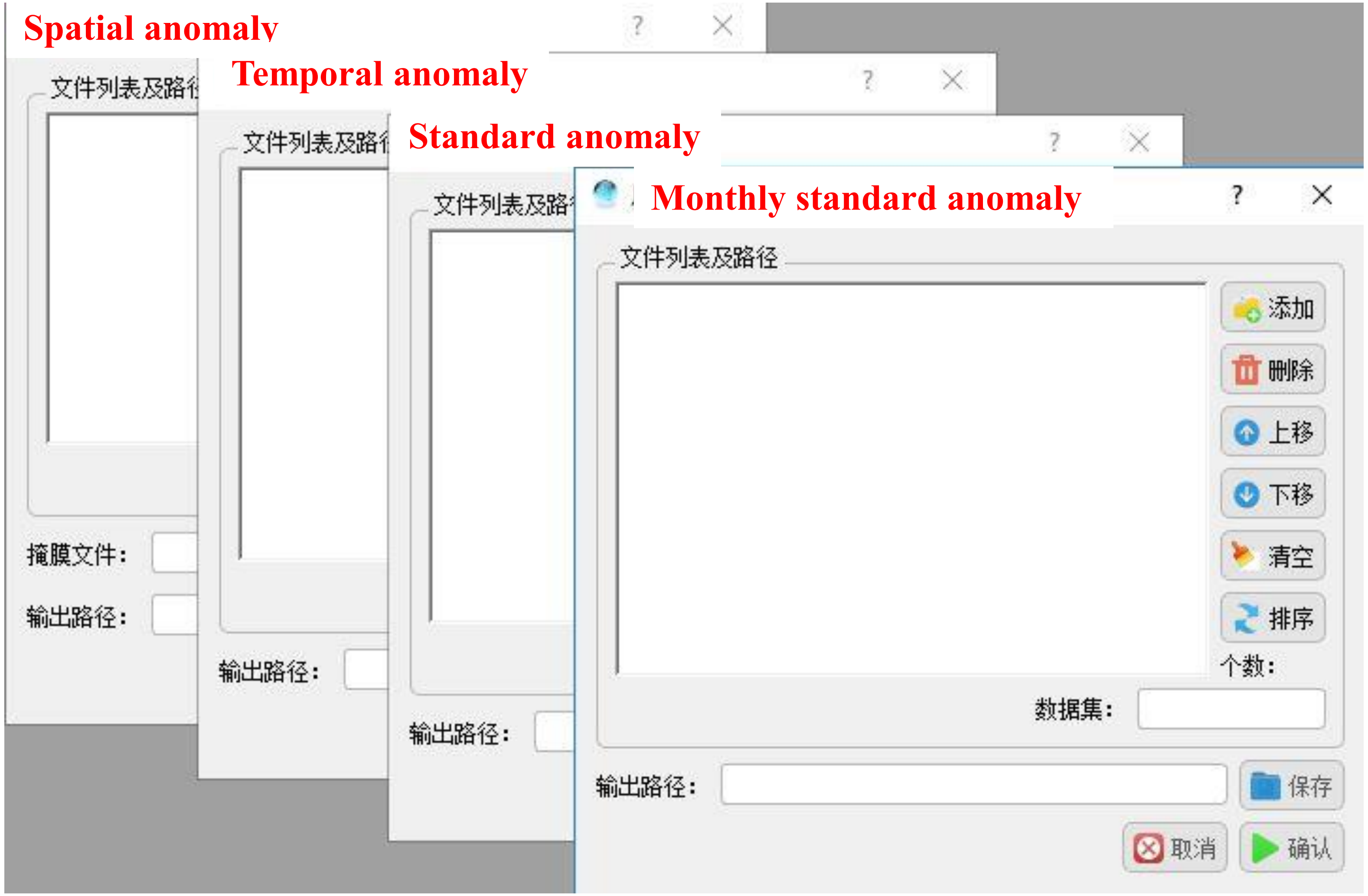
Min

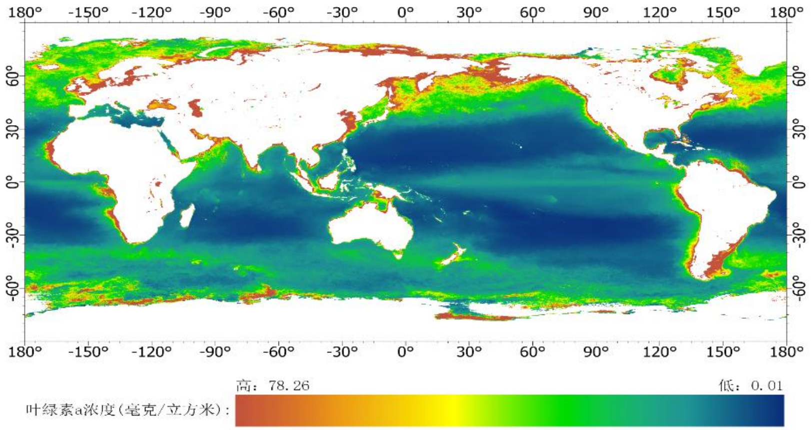
Mean

Standard devriation

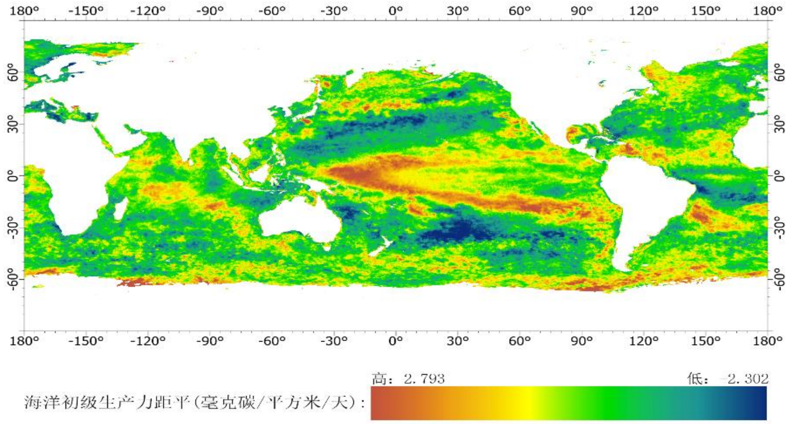


Marine information system- Anomaly analysis





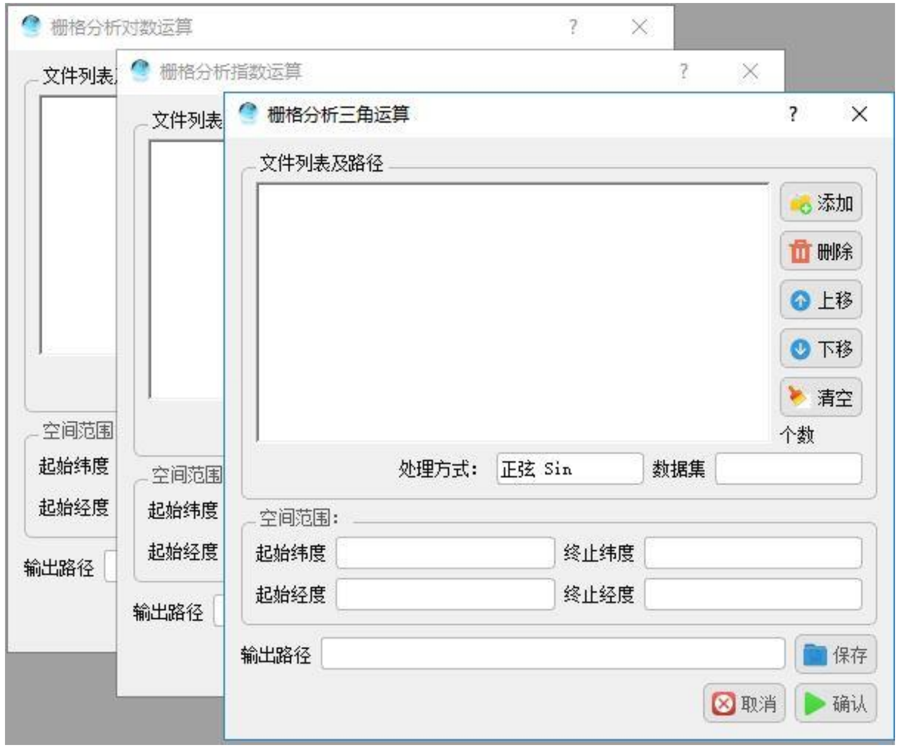
Monthly Chl_a concentration



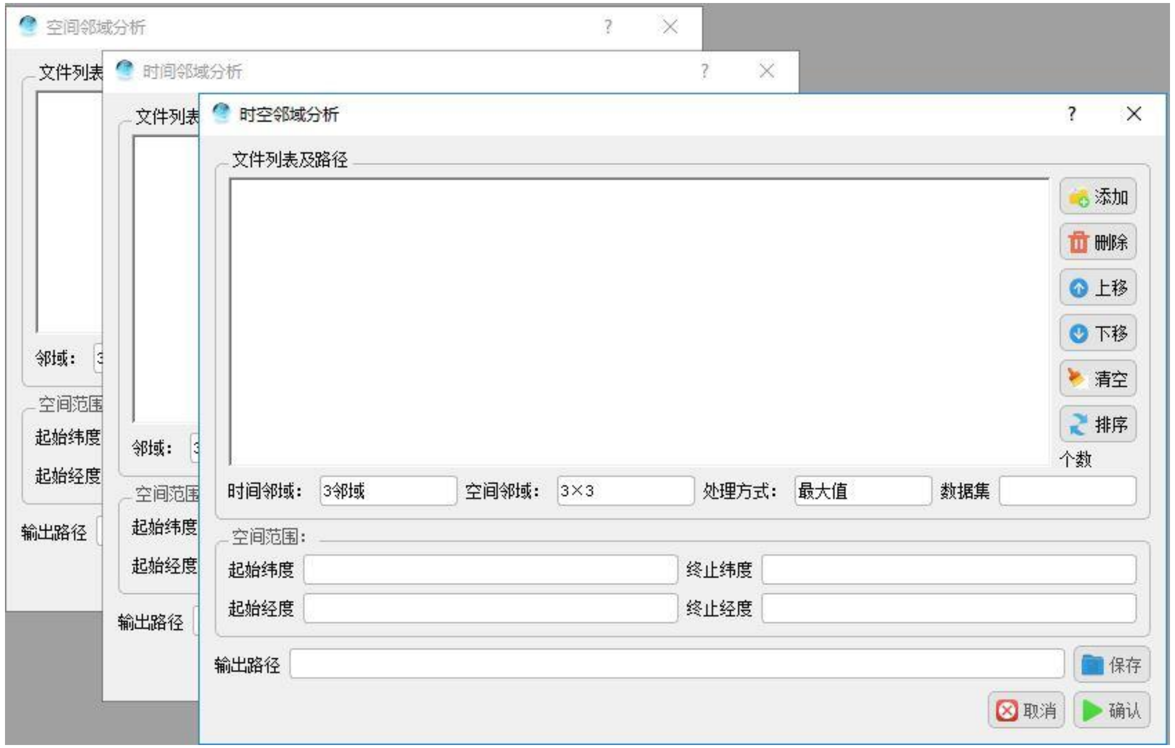
Monthly anomaly of Chl_a concentration

Marine information system- Spatiotemporal calculations

Math calculation



Neighborhood analysis



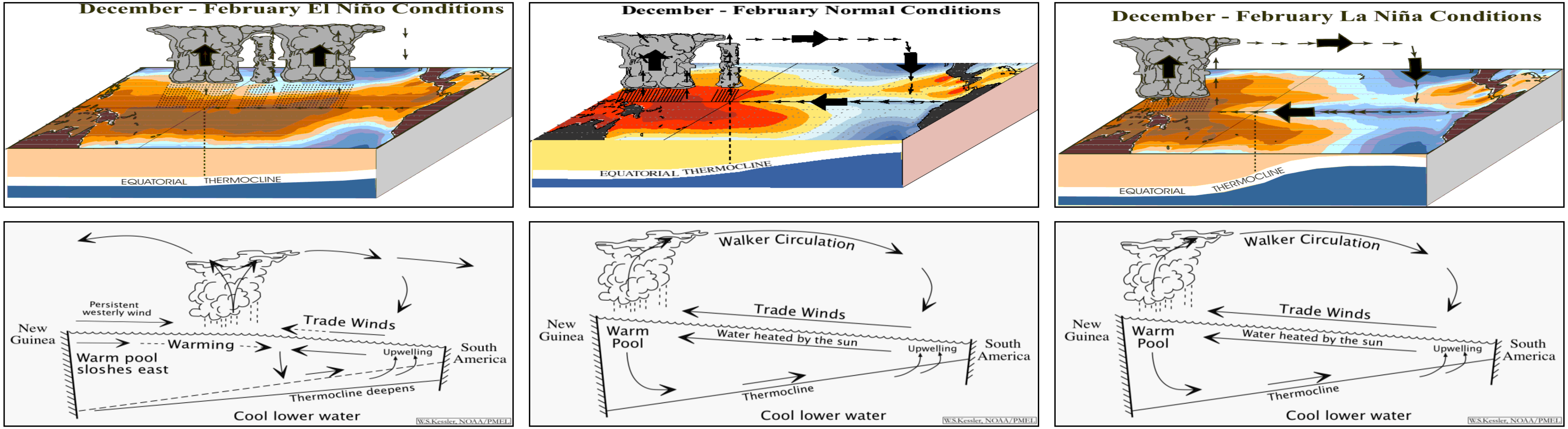
Outlines

- Challenges and demands from Marine SDGs
- Marine data-added mining technologies
- Digital twin of the ocean promotes 'what-if' scenarios simulations
- Marine information system for marine SDGs
- Case studies of monitoring and evaluating marine SDGs

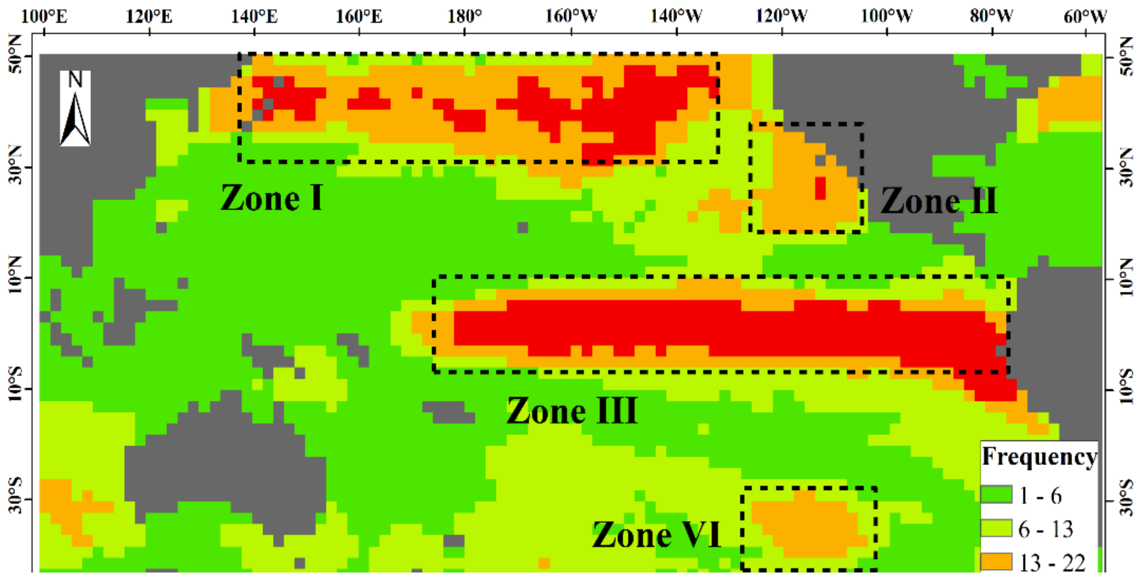
Case studies-SSTA dynamics in Pacific Ocean and its relationships with ENSO

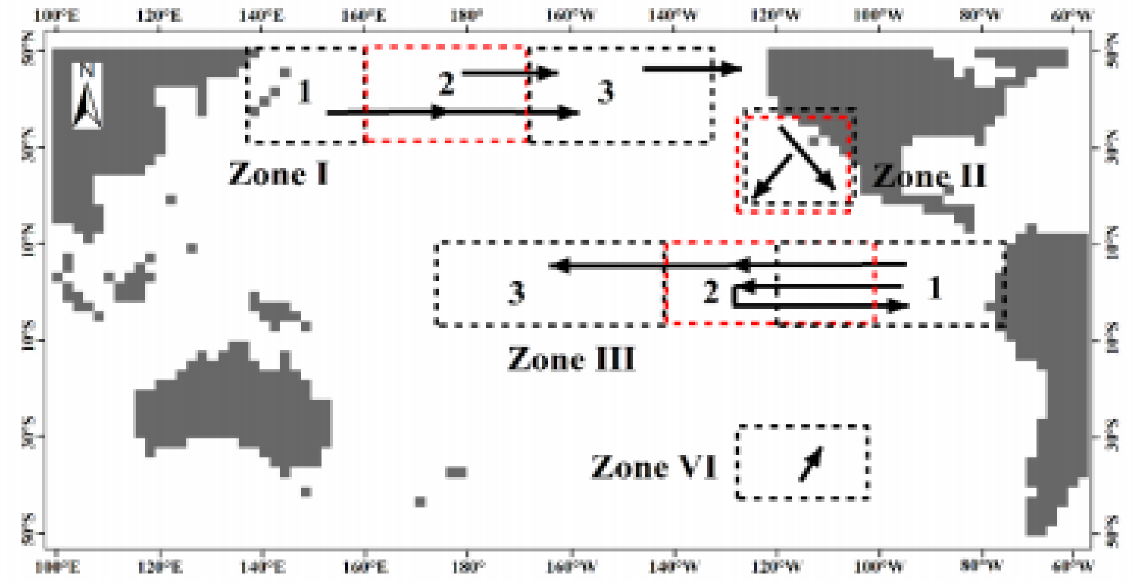
ENSO(El Niño-Southern Oscillation) is **one of the most important modes of variability** impacting the global climate, which has a close **mutual relationship with marine environment anomalies**. And **its origination, development and shrinkage** play an important role on **predicting extreme climate events, managing marine resources, and protecting marine ecosystems.**

The changing behaviors of marine environment anomalies will **strengthen or weaken the intensity of ENSO events**, which may serve as an important reference for research on **the mutual response and driving mechanisms** behind global climate change and abnormal marine variations.



Case studies-SSTA dynamics in Pacific Ocean and its relationships with ENSO





SST abnormal variations mainly covers four zones

Zone I： mid latitude North Pacific

Zone II： southwestern and coastal California

Zone III： central and eastern equatorial Pacific

Zone IV： mid latitude Southeast Pacific

Different zones show different movement patterns

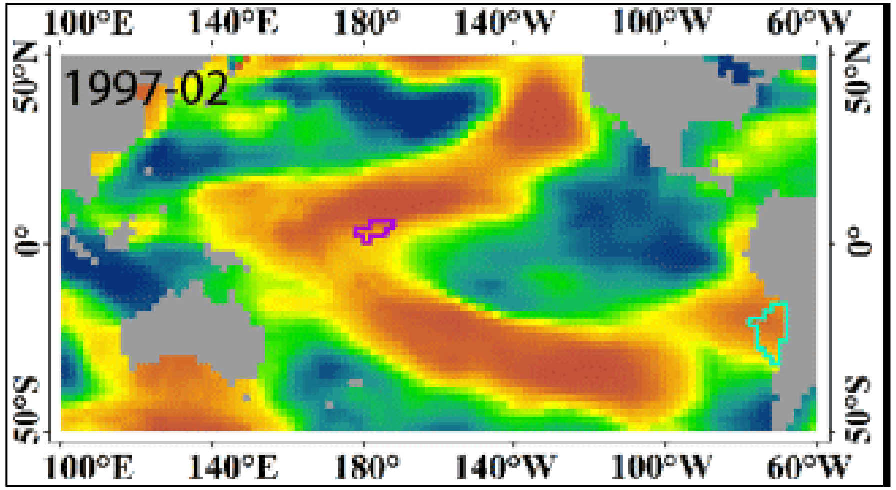
Zone I： From the west to the east

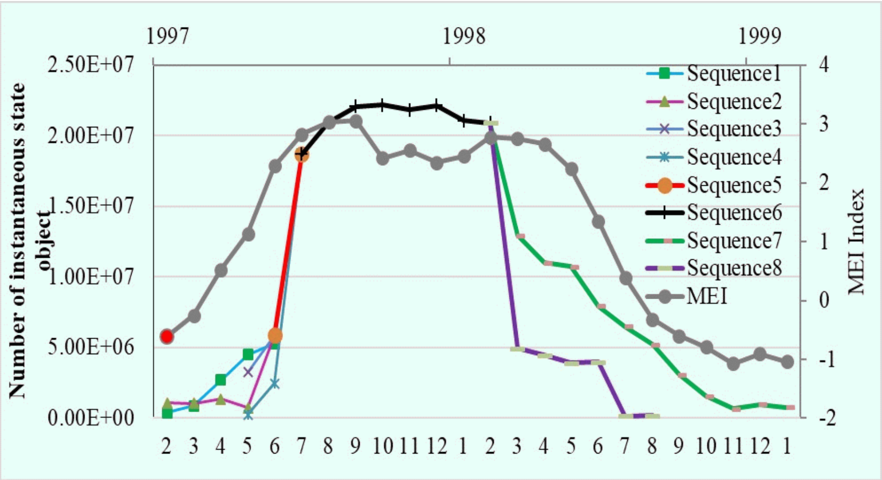
Zone II： Along the coast

Zone III： From the east to the west; from the east to west then return to the east.

Zone IV： from the south to the north

Case studies-SSTA dynamic evolution in Pacific Ocean





In February 1997, Sequence 1 is **generated** in the Eastern Pacific Ocean and **expands** in space from March to May, and Sequence 2 is **generated** in the Central Pacific Ocean and **expands** from March to June.

In May, Sequence 3 and Sequence 4 are **generated** in the Eastern Pacific Ocean.

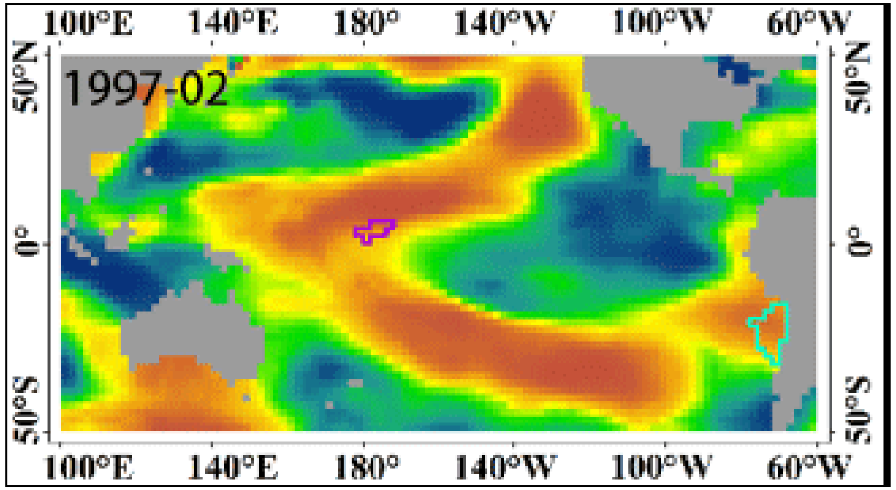
In June, Sequence 1 and Sequence 4 **merge into** Sequence 5.

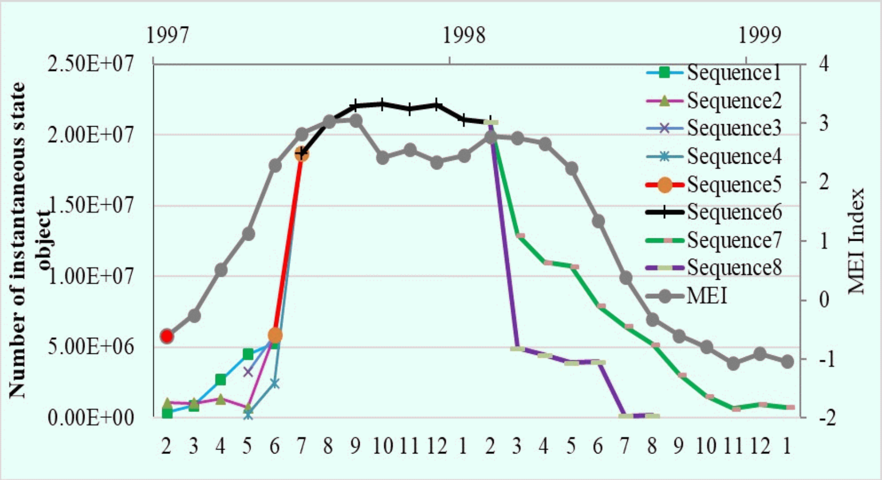
In July, Sequence 1, Sequence 3 and Sequence 5 **merge into** Sequence 6 in the Central Eastern Pacific Ocean, where Sequence 6 continues to **develop** from July 1997 to February 1998.

In March 1998, Sequence 6 **splits** Sequence 7, in the Central Eastern Pacific Ocean, and Sequence 8, in the Eastern Pacific Ocean.

From April, Sequence 7 from the Central Pacific Ocean **begins to shrink** in space and **disappears** in January 1999 in the Eastern Pacific Ocean, and Sequence 8 decreases and **disappears** in August 1998 in the Eastern Pacific Ocean.

Case studies-SSTA dynamic evolution in Pacific Ocean





The **generation, development, merging, splitting and disappearance** of the abnormal marine process are generally **consistent** with the evolution of an El Niño event.

When Sequence 1 and Sequence 2 are being **generated**, the El Niño event is also about to **begin**.

When the time of Sequence 1, Sequence 3 and Sequence 5 **merging** in a space, the strength of the El Niño event **reaches its maximum**.

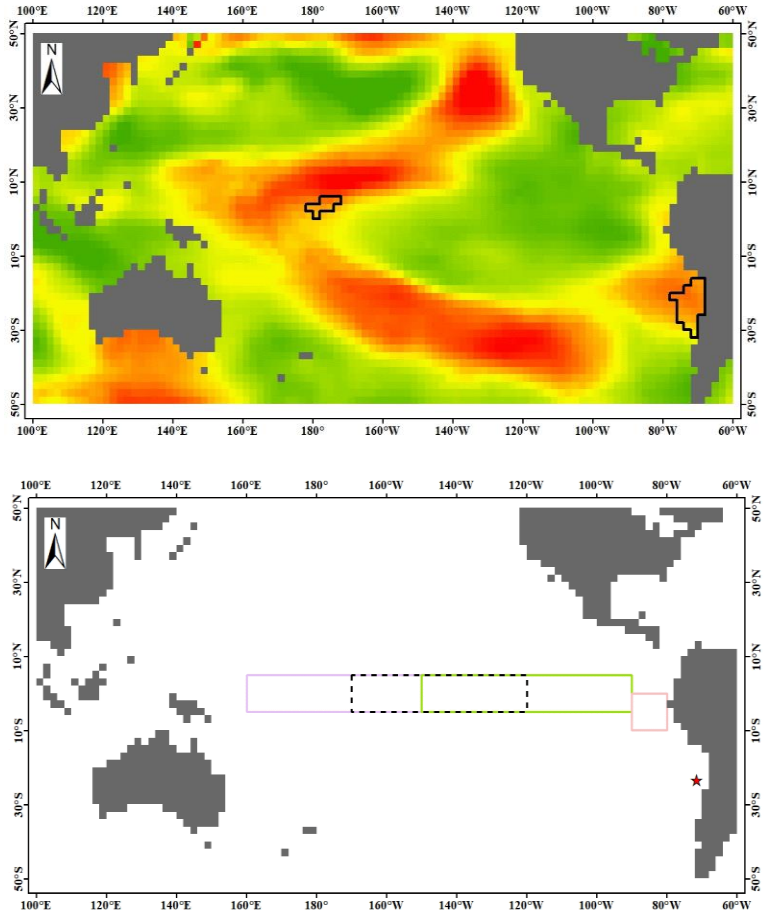
During the period of the **development** of Sequence 6, the El Niño event **remains stable**.

when the process **splits** and comes to **disappear**, after three months, the El Niño event **begins to decrease** and then **disappears** over time.

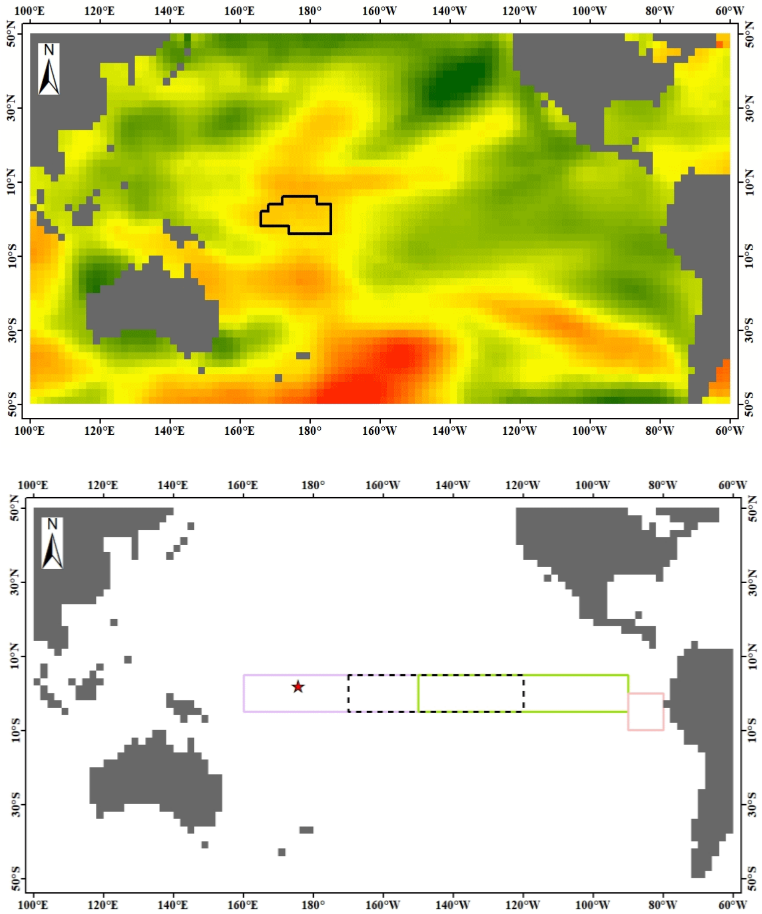
EP-ENSO: SSTA generates, develops and disappears in Eastern Pacific ocean

CP-ENSO: SSTA generates, develops and disappears in Central Pacific ocean

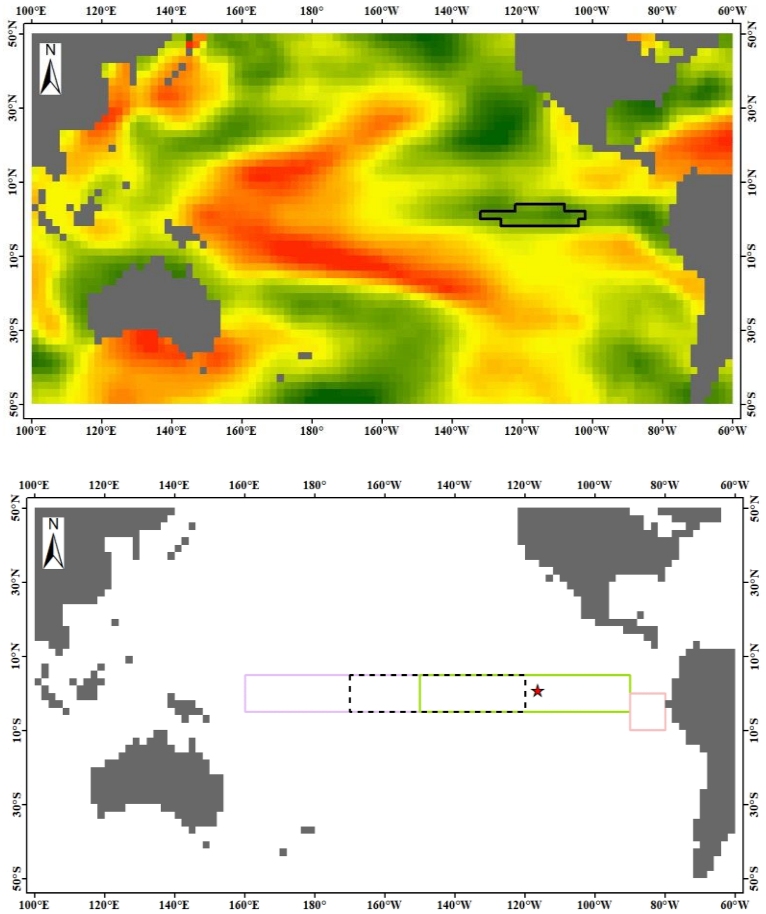
Mix-ENSO: SSTA generates in Eastern Pacific ocean, and develops from east to west, and disappears in Central Pacific ocean



EP-ENSO *Vs* Evolution of SSTA



CP-ENSO*Vs* Evolution of SSTA



MIX-ENSO *Vs* Evolution of SSTA

EP-ENSO: SSTA generates, develops and disappears in Eastern Pacific ocean

CP-ENSO: SSTA generates, develops and disappears in Central Pacific ocean

Mix-ENSO: SSTA generates in Eastern Pacific ocean, and develops from east to west, and disappears in Central Pacific ocean

A new rule to identify ENSO event is proposed according to the origination, development and disappears of SSTA

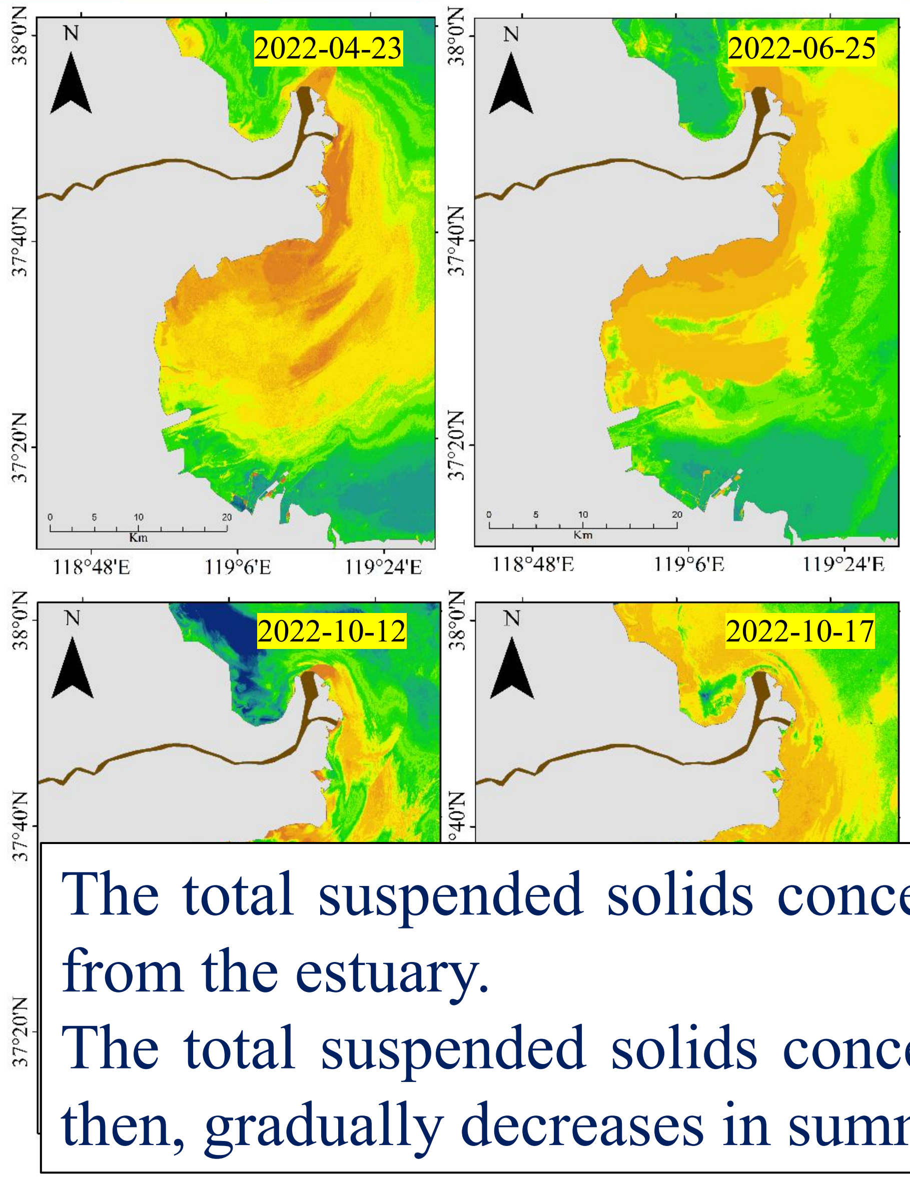
- **If SSTA generates, develops and disappears in region of Niño 1+2 or 3, then EP-ENSO**
- **If SSTA generates, develops and disappears in region of Niño 3.4 or 4, then CP-ENSO**
- **If SSTA generates in region of Niño 1+2 or 3 , and develops from east to west, and disappears in region of Niño 3.4 or 4, then Mix-ENSO**
- **If SSTA generates in region of Niño 3.4 or 4, and develops from west to east, and disappears in region of Niño 1+2 or 3, then Mix-ENSO**

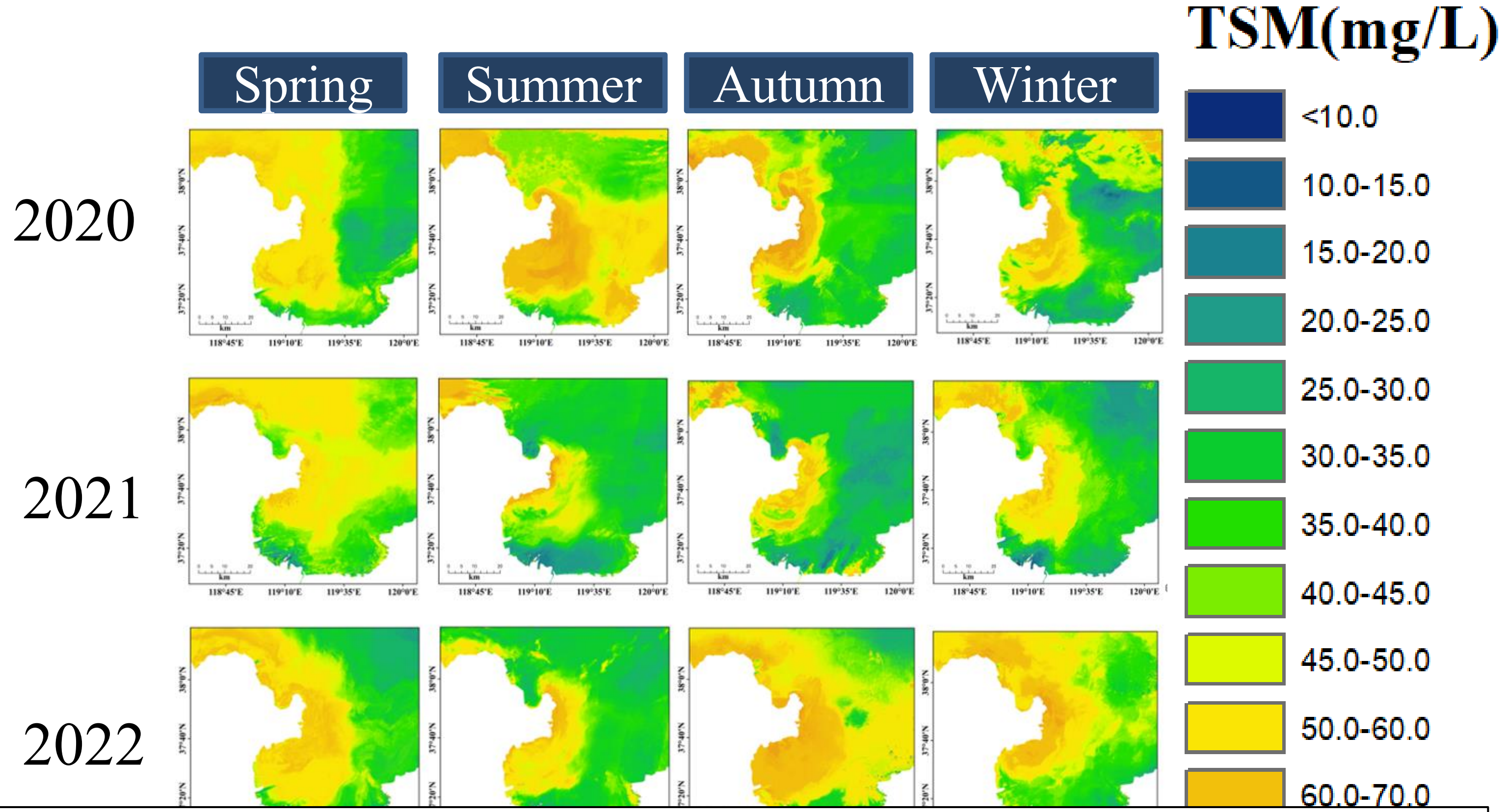
Well known EP/CP ENSO, all the algorithms have the same results

Disagreed EP/CP ENSO, the algorithms have the different results

	SSTAV Evolution-based	*EMI*	*PTN-based (Yu et al., 2013)*	I_{cp} / I_{ep}	*EOF-based (Kao & Yu, 2009)*	N_{CT}/N_{WP}	*CT/WP (Li. et al,2017)*
1951	EN-EP	EP	MIX		EP	EP	EP
1976	EN-EP	EP	MIX	EP	EP	EP	EP
1954-1957	LN-MIX			MIX	MIX	EP	MIX
1957-1958	EN-MIX	CP	CP	MIX	EP	EP	EP
1961-1962	LN-EP					EP	
1963-1964	EN-CP	CP	CP		CP	EP	EP
1964-1965	LN-EP			EP	EP	EP	
1965-1966	EN-MIX	CP	CP	MIX	MIX	EP	EP
1970-1972	LN-EP				MIX	MIX	EP
1972-1973	EN-EP	EP	MIX	EP	EP	EP	EP
1973-1976	LN-MIX			MIX	CP	CP	EP
1982-1983	EN-EP	EP	EP	EP	EP	EP	EP
1984-1985	LN-EP			EP	CP	EP	
1986-1988	EN-MIX	EP	MIX	MIX	CP	EP	MIX
1988-1989	LN-MIX			MIX	MIX	CP	MIX
1991-1993	EN-CP	CP	CP	CP	CP	EP	EP
1994-1995	EN-CP	CP	CP		CP	CP	CP
1997-1998	EN-EP	EP	EP	EP	EP	EP	EP
1998-2000	LN-CP			CP	CP	MIX	
2002-2003	EN-CP	CP	CP	CP	CP	CP	CP
2004-2005	EN-CP	CP	CP	CP	CP	CP	CP
2006-2007	EN-MIX	EP	MIX		CP	CP	EP
2007-2008	LN-MIX			EP		MIX	
2009-2010	EN-CP	CP	CP	EP		CP	CP
2010-2011	LN-EP			CP		CP	EP
2011-2012	LN-EP			EP		CP	
2014-2016	EN-MIX			MIX			

Case studies-RS Imagery-based water quality monitoring model in coastal regions





The total suspended solids concentration near the Yellow River decreases with the distance from the estuary.

The total suspended solids concentration reaches a high value in spring season every year, then, gradually decreases in summer and autumn, and increases from autumn to winter.

Case studies-RS Imagery-based water quality monitoring model in coastal regions

DOC based on Sentinel-2

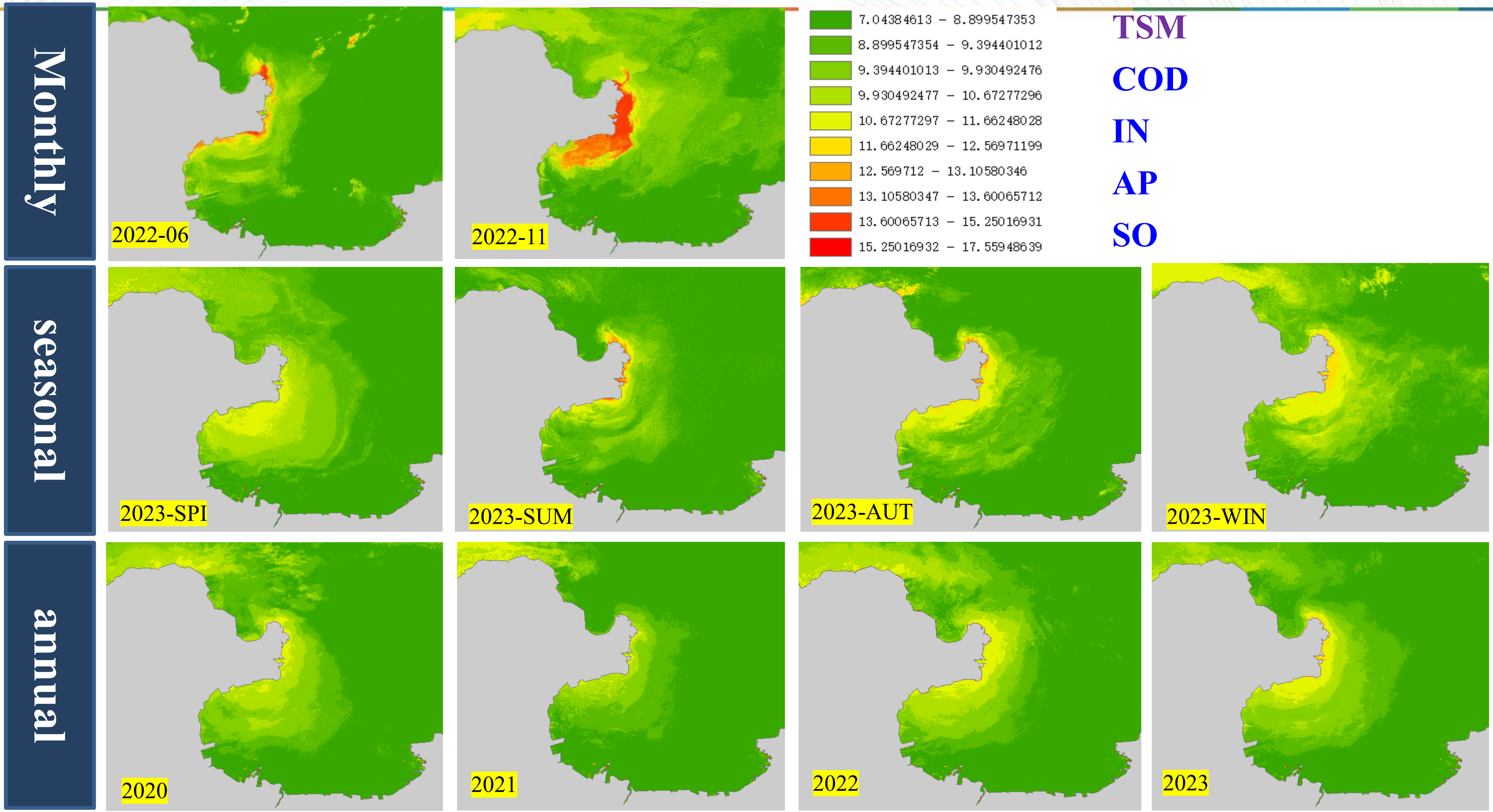
TSM

COD

IN

AP

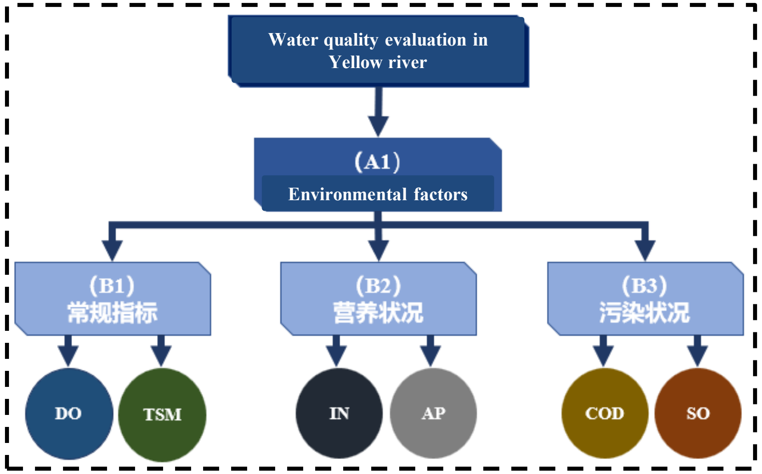
SO

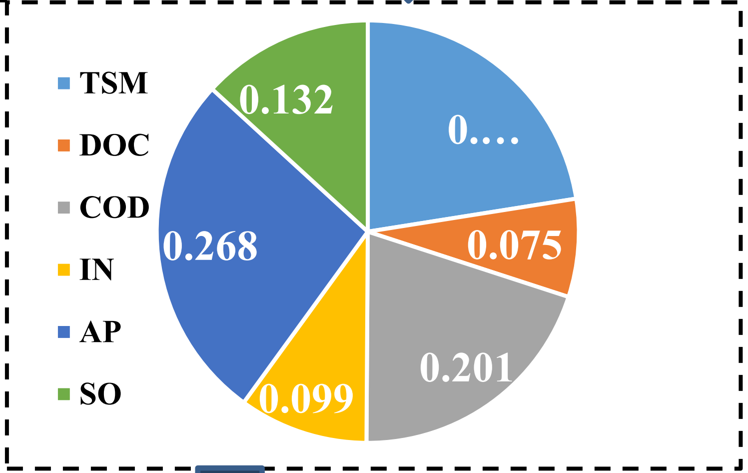


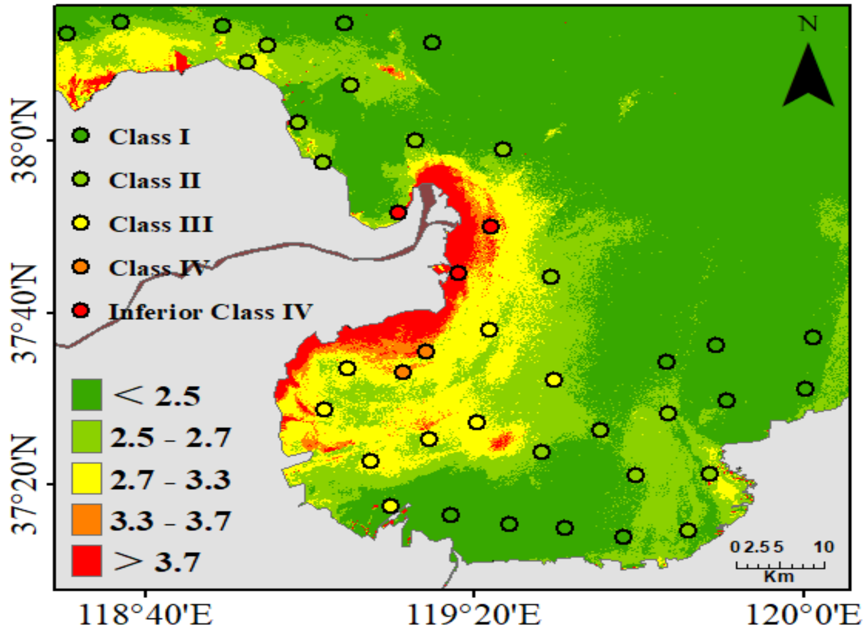
Case studies-RS Imagery-based water quality monitoring model in coastal regions

AHP Weights

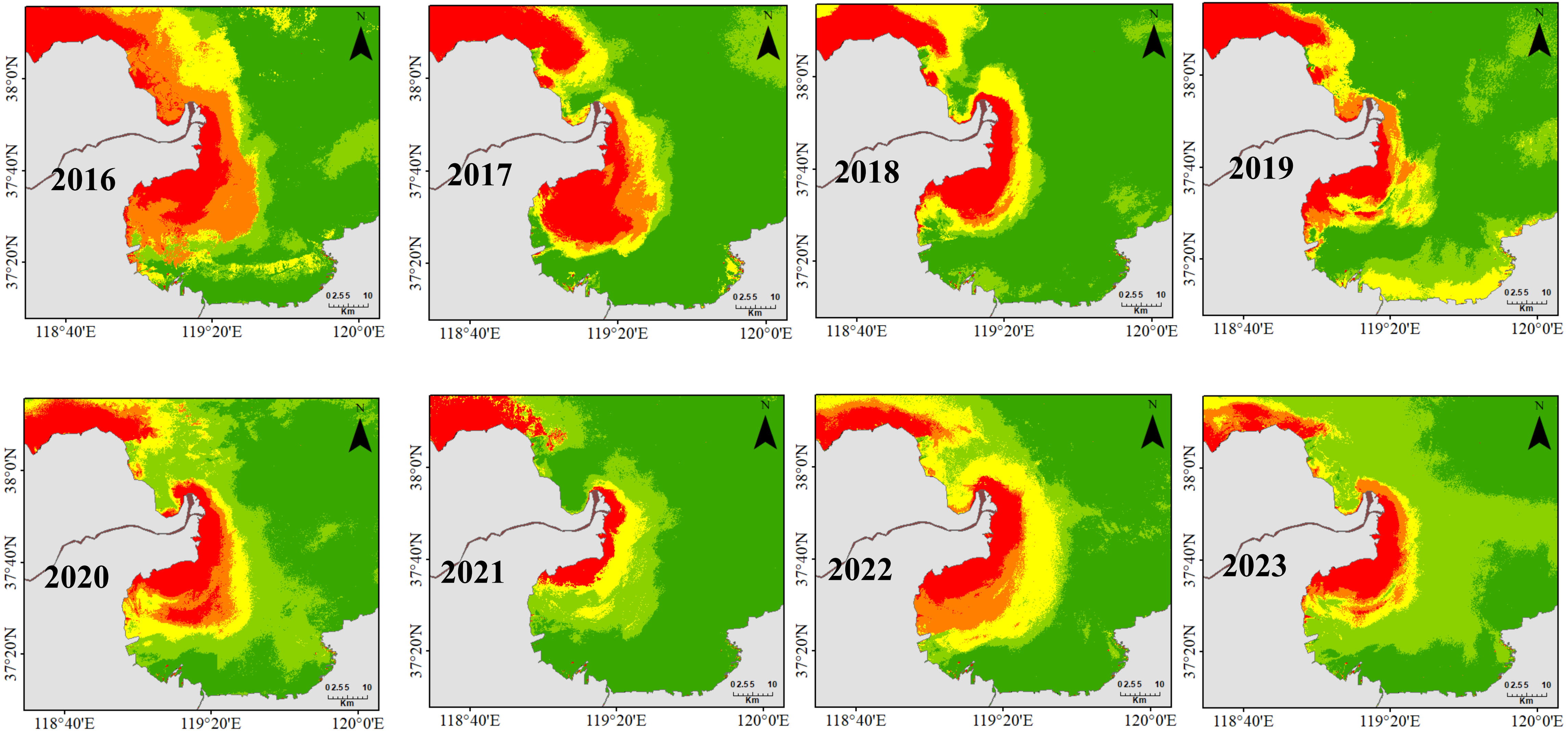
2nd factors	3rd factors	Matrix Weights	
Common(B1) 0.30	DOC (C1)	1	3
	TSM (C2)	1/3	1
Nutritional status (B2)0.30	IN (C3)	1	2
	AP (C4)	1/2	1
Pollution status(B3) 0.40	COD (C5)	1	2
	SO (C6)	1/2	1





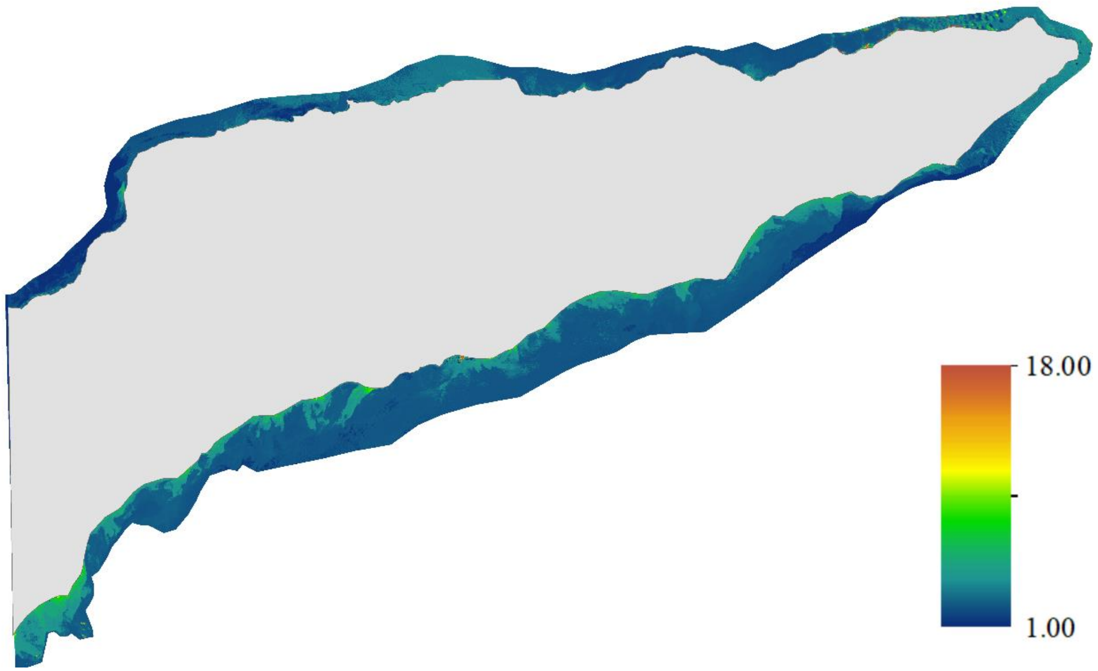


Case studies-RS Imagery-based water quality monitoring model in coastal regions

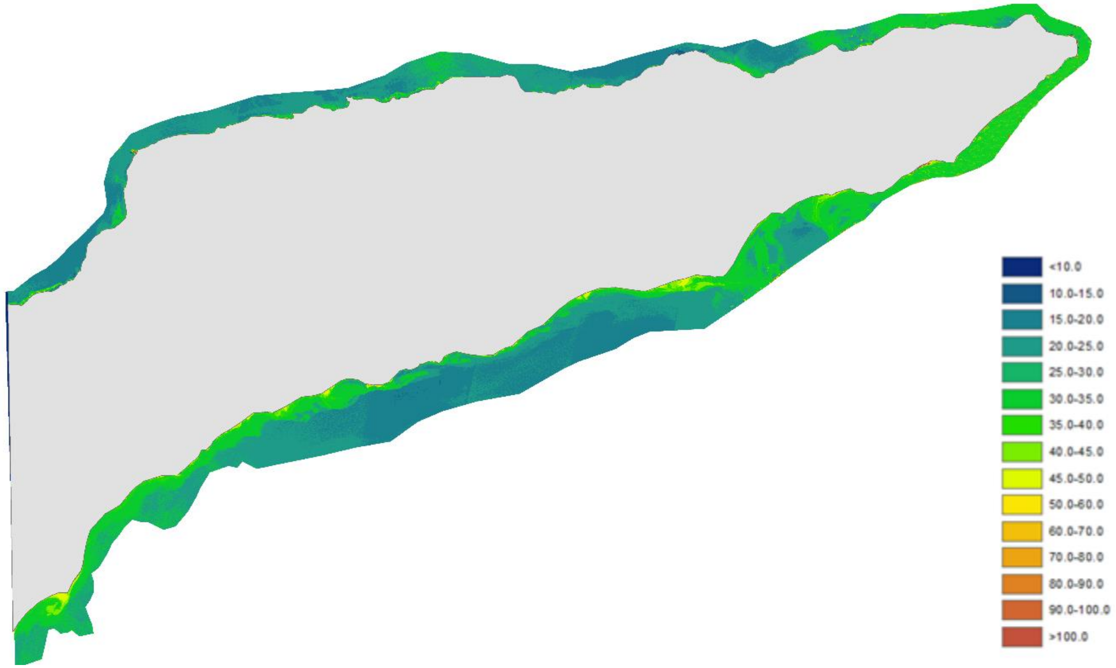


四、Case studies- water quality in Timor-Leste

- QAA-RF model based on real data in offshore China
- Sentinel-2 image



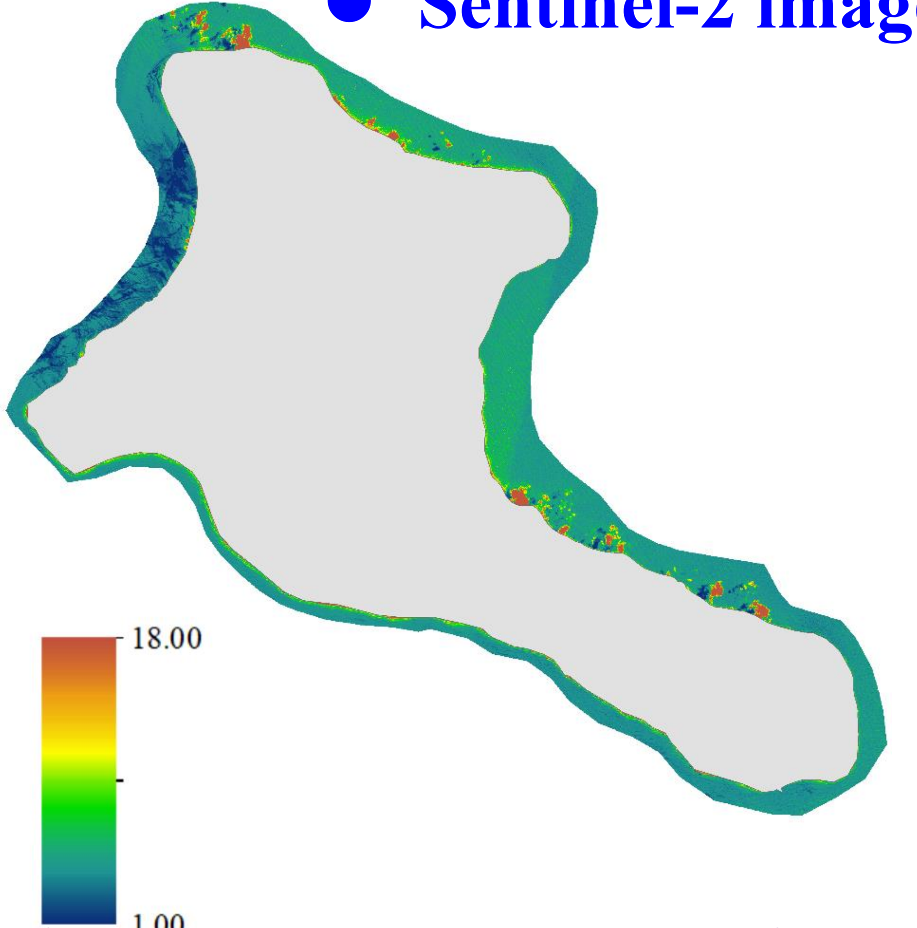
Dissolved oxygen concentration



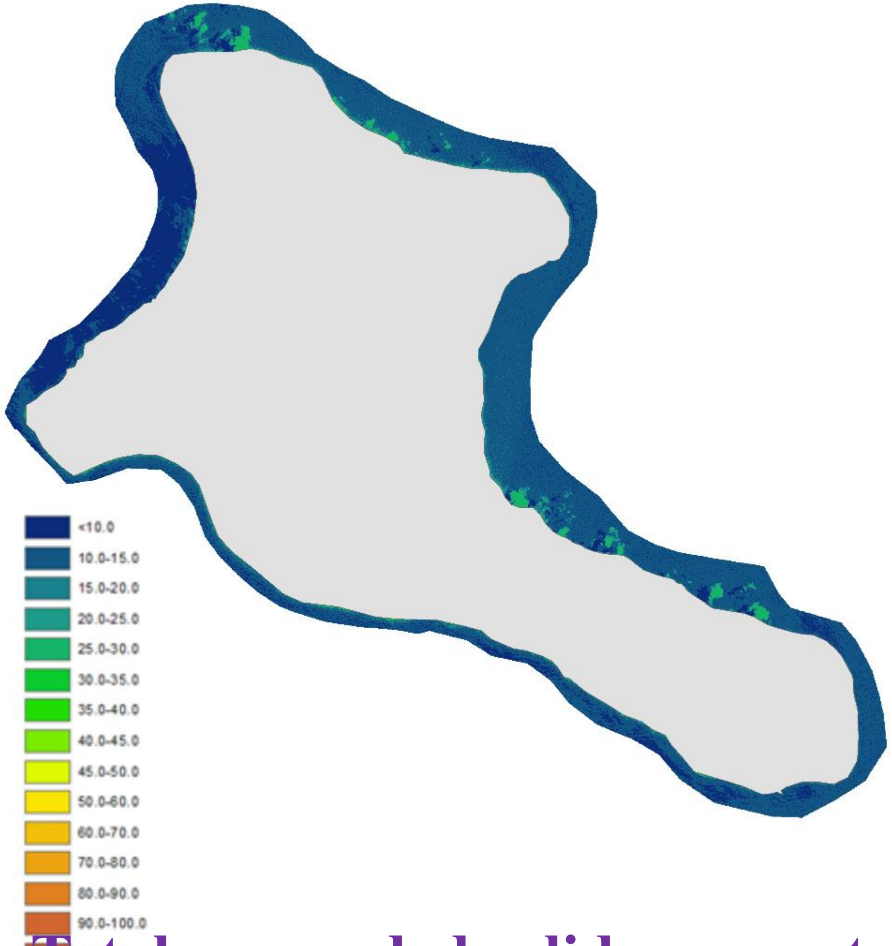
Total suspended solids concentration

四、Case studies- water quality in Republic of Kiribati

- **QAA-RF model based on real data in offshore China**
- **Sentinel-2 image**



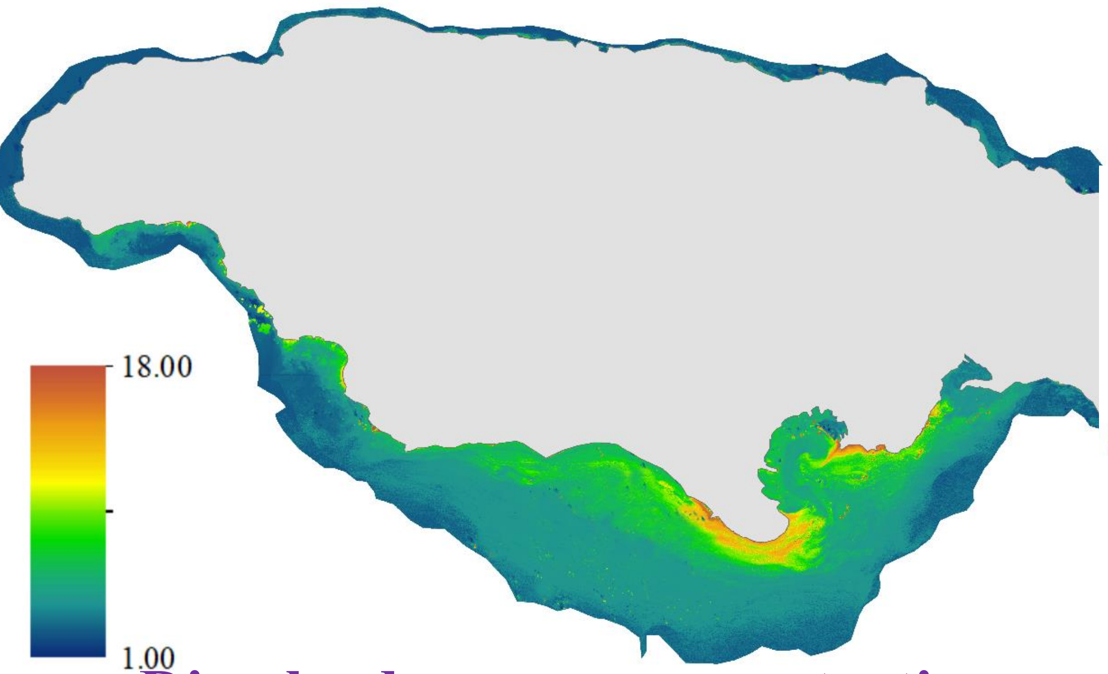
Dissolved oxygen concentration



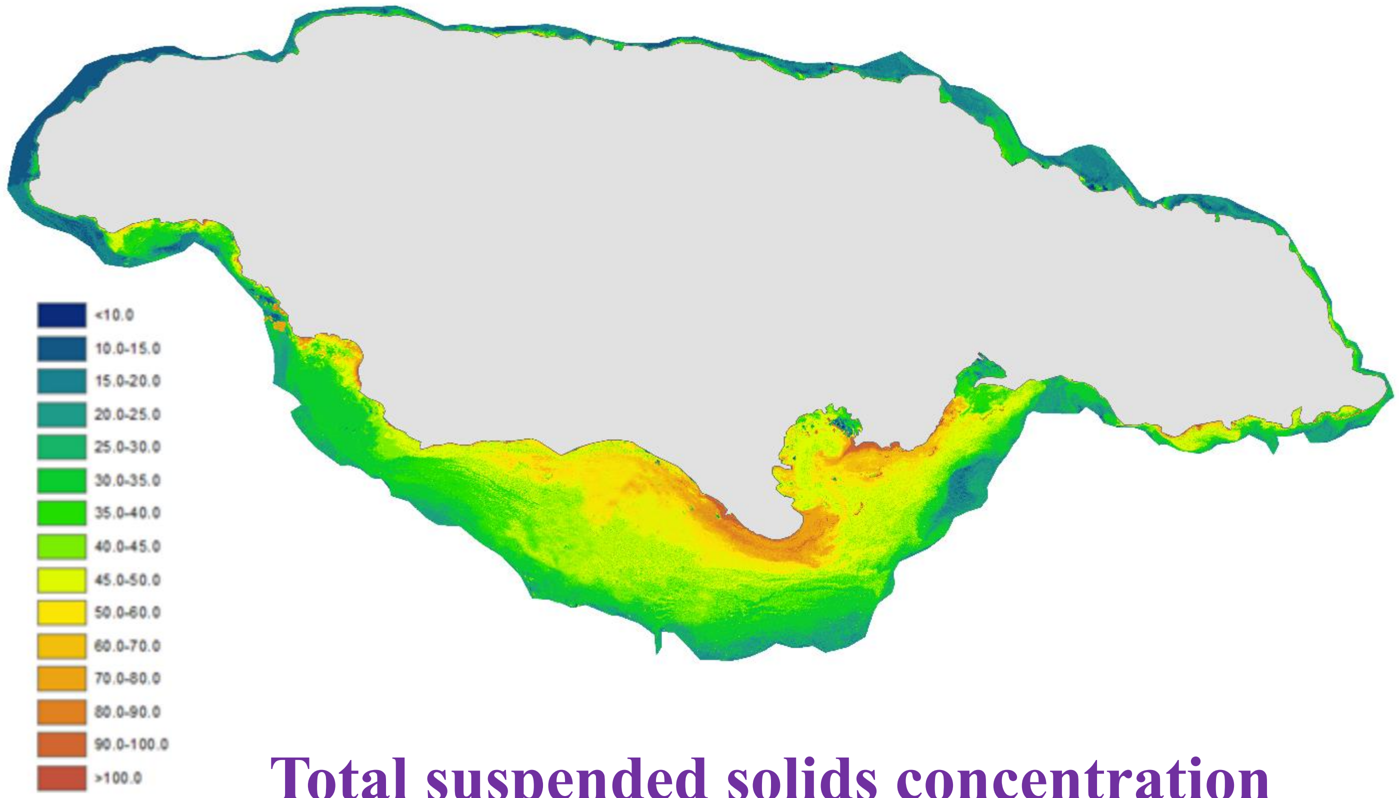
Total suspended solids concentration

四、Case studies- water quality in Jamaica

- QAA-RF model based on real data in offshore China
- Sentinel-2 image



Dissolved oxygen concentration



Total suspended solids concentration

Case studies-Digital twin of the Island

Layer Management

Timor-Leste

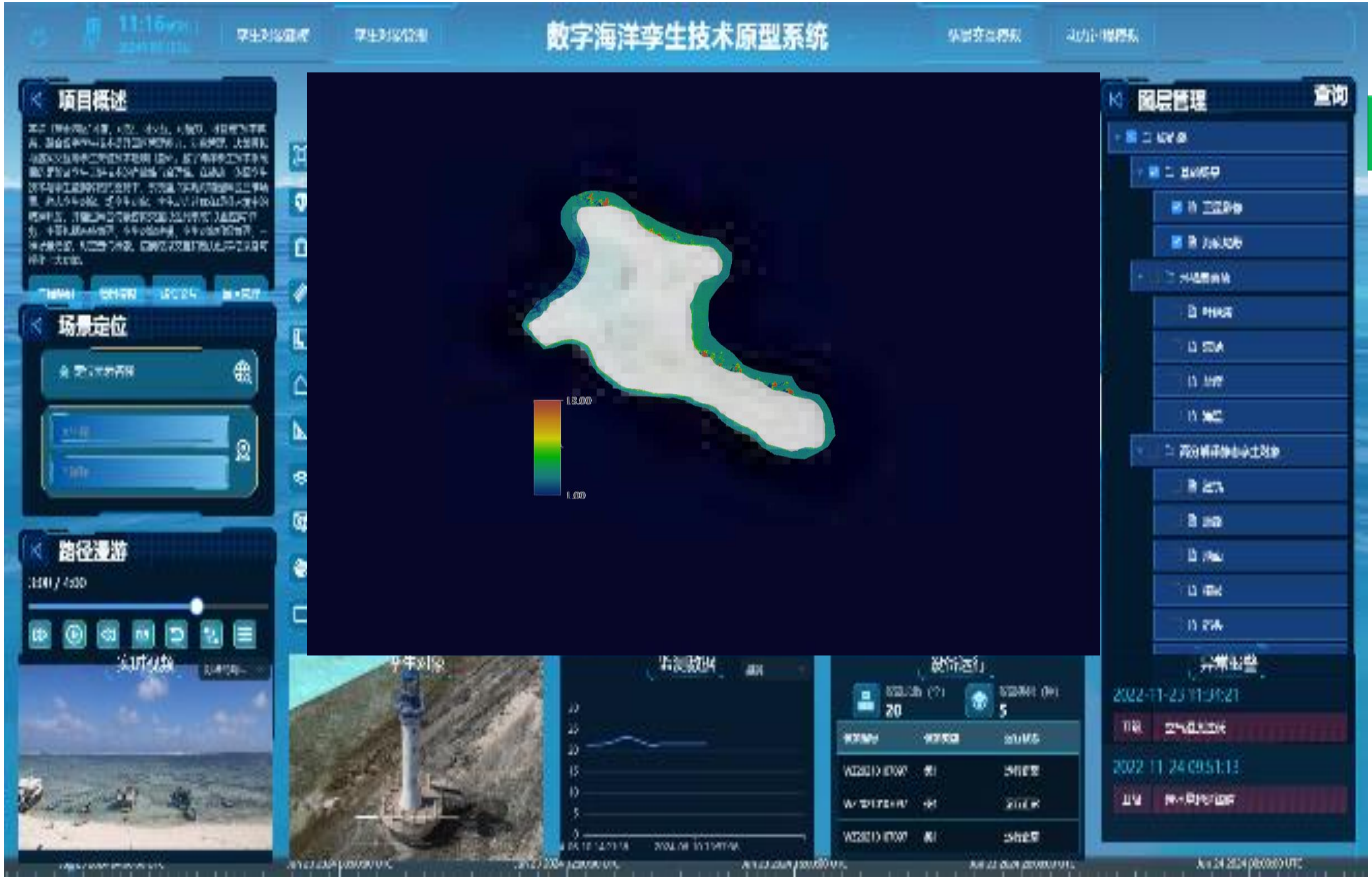
DOC

TSC

Kiribati

Jamaica

Case studies-Digital twin of the Island



Layer Management

- Timor-Leste
- Kiribati
- DOC
- TSC
- Jamaica

Case studies-Digital twin of the Island

Layer Management

- Timor-Leste
- Kiribati
- DOC
- TSC
- Jamaica

Summaries

- **From The ocean we have to The ocean we want, there is a huge technological gap, from the data processing through information mining to information system.**
- **Marine big data mining and digital ocean of the ocean play a technological foundation for dealing with the challenges of oceanic sustainable development.**
- **SDGSAT-1 and its subsequent satellite series provide an important source of monitoring and evaluating the SDG14 and others.**
- **International Research Center of Big Data for Sustainable Development Goals (CBAS) has been in actions, and the established models and information systems are well portable and can effectively serve the marine sustainable development of small island countries.**

INTERNATIONAL RESEARCH CENTER OF BIG DATA FOR SUSTAINABLE DEVELOPMENT GOALS
可持续发展大数据国际研究中心

Thank you