

CONFÉRENCE

TIC et mer: nouveaux défis et solutions

Les technologies de l'information au service de la recherche marine

Gérer des bases de données de plus en plus grandes et complexes

Partage et interactions de bases de données

Méthodes de fouille et d'analyse

9H45: ACCUEIL

10H00-16H00: PRÉSENTATIONS

MATHIAS HERBERTS, JEAN-FRANÇOIS PIOLLÉ,
STÉPHANIE MAHÉVAS, GUILLAUME MAZE,
THOMAS LOUBRIEU, GILBERT MAUDIRE,
PHILIPPE LENCA, RONAN FABLET

16H00: TABLE RONDE AVEC RENÉ GARELLO

Pierre Tandeo

Initialement par John Puentes

(Telecom Bretagne)

“Processing legacy data”

26 Novembre 2013, Ifremer, Brest



<http://wwz.ifremer.fr/bigdata>

Pertinence of processing legacy data Application to the SST

IFREMER Brest
Conference on Big Data

P. Tandeo (1), J. Puentes (1), B. Chapron (2),
E. Autret (2), R. Fablet (1) and R. garello (1)
`pierre.tandeo@telecom-bretagne.eu`

(1) Lab-STICC, Telecom Bretagne
(2) LOS, IFREMER

26th of November 2013

Spatial oceanography

- 20 years of satellite data:
 - different sensors
 - various physical variables
 - large missing data
- Currently:
 - only $<10\%$ used
 - resolution increases

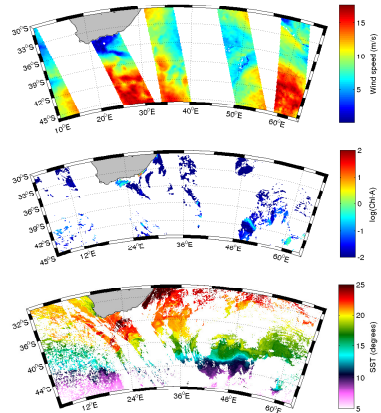


Figure: Sea surface wind, colour and temperature

Sea Surface Temperature (SST)

Name	Description	Spatial resolution	Temporal resolution
AMSRE	Microwave	~ 25 km	~ 24 h
SEVIRI	Infrared	~ 10 km	~ 1 h
METOP	Infrared	~ 5 km	~ 12 h

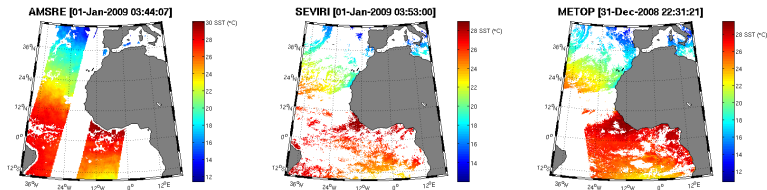


Figure: Different satellite sensors to measure SST (~ 250 Go/day)

Goal & Issues

- Goal:
 - reduce information
 - synthetic/realistic
- Uses:
 - GCM input
 - ocean circulation
 - others (e.g., ecology, climate)
- Issues:
 - blend different data?
 - interpolate missing data?
 - use historical data?

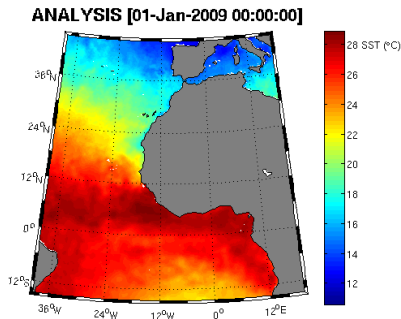


Figure: Synthetic SST field

Compute statistics

- First order:

- global mean
- climatology

- Second order:

- temporal covariance
- spatial covariance

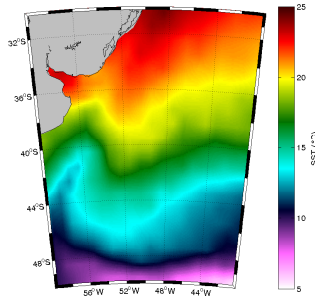


Figure: SST climatology

Compute statistics

- First order:
 - global mean
 - climatology
- Second order:
 - temporal covariance
 - spatial covariance

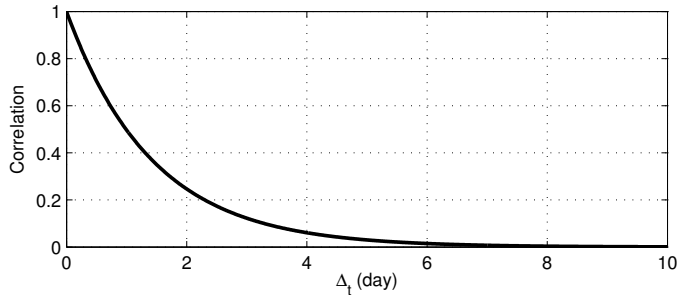


Figure: Temporal correlations

Compute statistics

- First order:

- global mean
- climatology

- Second order:

- temporal covariance
- spatial covariance

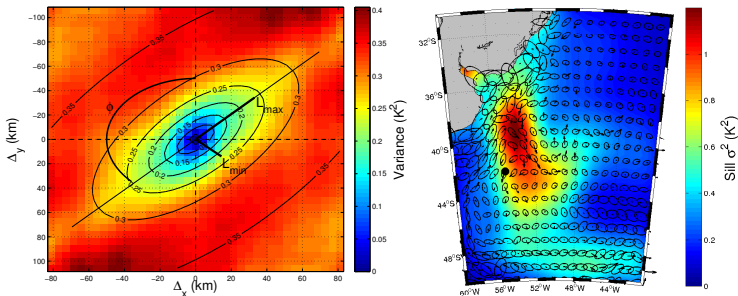


Figure: Anisotropic spatial variograms

Relevant to store data with large missing values?

- Kalman filtering:
 - spatio-temporal covariances
 - different resolutions (5 km, 10 km, 25 km)
- Case of study:
 - highly dynamical region
 - year 2008
- Different experiments:
 - missing data rates
 - from 60 % to 80 %

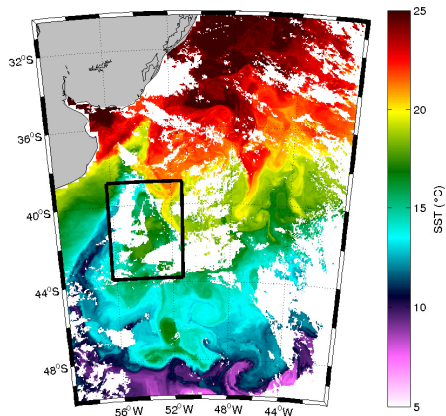
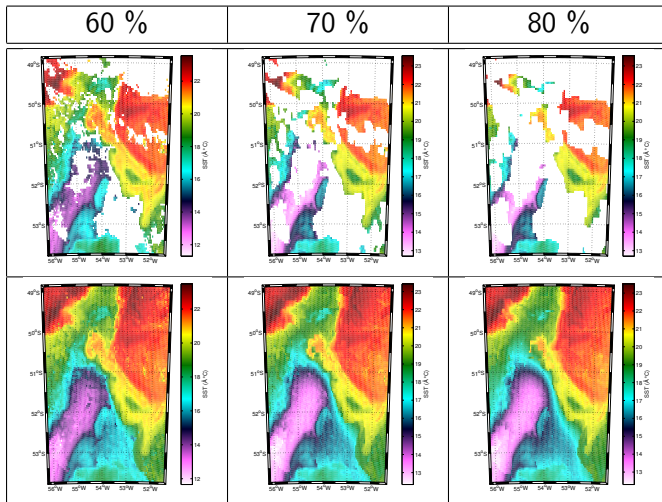


Figure: Zone of study off the Argentinian coasts

Missing data rate changes



Missing data rate changes

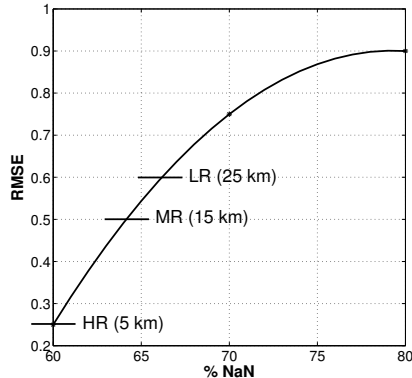


Figure: RMSE as a function of missing values

Issues

- Historical data:
 - large database
 - various resolutions/variables
 - only <10% used
- Question:
 - how to use all the information?
- Future data:
 - high spatial resolution
 - high storage cost
- Questions:
 - store everything?
 - criteria to select relevant data?

Use historical data

- Find similar features:
 - patch-based approach
 - distance between images
 - project on reduced bases
- Problems:
 - physical meaning?
 - realistic interpolation?
 - temporal consistency?

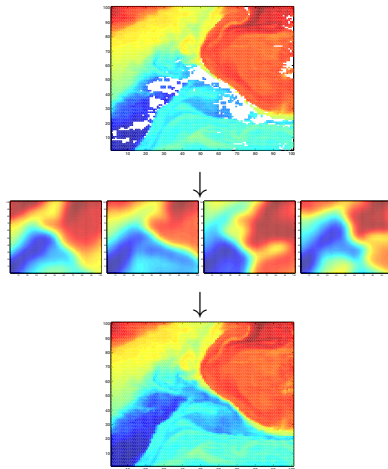


Figure: Inpainting method

Use historical data

- Learn relationships between:
 - low resolution data
 - high resolution data
- Methods:
 - downscaling
 - learn contour statistics

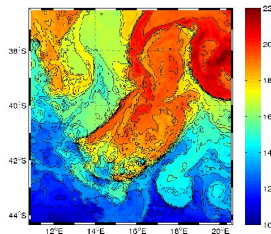
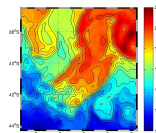


Figure: Super resolution method

Use historical data

- Learn relationships between:
 - different variables
 - using latent variables
- Recent projects:
 - ANR EMOCEAN:
 - IFREMER
 - Telecom Bretagne
 - Ocean Data Lab
 - OUC (China)
 - Nephelae platform:
 - IFREMER / CERSAT
 - data access
 - computing power

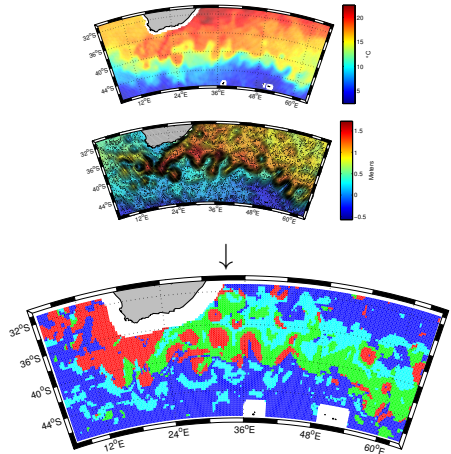


Figure: Synergy between SST and SSH