

# Datasets

- As part of a (COST?) action, gather existing datasets with complete input and data for comparison
- EMODNET: make available datasets
- We need to be more selective: specific datasets for specific functionalities
- Portal for validation; link with existing platforms (EMODNET, ..., CSDMS..)
- Only well-documented and verified data

# Datasets(2)

- Topology of data needed?
  - Coastal GOOS report summarizes
- Documenting needs to test model application at different locations.
  - Depending on region
  - Depending on scales
- Developing strategies for data collection for validation
- Lots of data on temperature and salinity, much less on currents

# Optimizing data collection

- Different ways: difficult, through data assimilation with synthetic datasets, or simple, based on covariance matrices.

# Methodology for validation

- How to move forward
- Not just numbers, find out how to verify the processes
- Analyse the propagation of errors through the system
- Collaborate with CSDMS consortium

# Methods

- Pattern recognition techniques to develop metrics
- Summarise the techniques that people use for specific processes
- Models must be rigorously tested for the right parameters (e.g. mean fluxes can be wrong even if water levels are perfect)
- Validation must be related to specific scales (e.g. seasonal vs interannual)
- Easiest: synoptic; then interannual; then seasonal?
- Need to stop thinking in deterministic way; use probabilistic representation
- Ensemble Kalman filtering can add much to understanding model error
- Identify the modelling and schematization choices in different types of models

# Annotating Graphs

- Title
- Code
- Objective of the graph
- Targeted end user
- Criterion of a usable/successful model output
- Strength/weakness analysis
- Future directions/developments (of the graph)
- Applicability.transportability
- Extra data needed?
- How can the fit be improved
- Statistical foundation and assumptions