SANGOMA

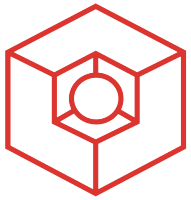
WP2: Sharing and Collaborative Development

Lars Nerger

Alfred Wegener Institute for Polar and Marine Research
Bremerhaven, Germany

and
Bremen Supercomputing Competence Center BremHLR
Bremen, Germany

lars.nerger@awi.de



BremHLR
Kompetenzzentrum für Höchstleistungsrechnen Bremen

Synopsis

Provide tools of interest to the Data Assimilation community and avoid redundant developments

by

adapting existing and developing new tools according to data model and interface standard from WP1.

Possible Tools

- Identification by WP1
- Preliminary inventory of existing tools generated when proposal was formulated
- New tools identified by WP3 and WP4 during project

5 categories

- Diagnostic tools
- Perturbation tools
- Transformation tools
- Utilities
- Analysis steps ➜ WP 3

Existing tool boxes

Existing tools spread over range of tool boxes:

- Beluga/Sequoia (Toulouse)
- OpenDA (Delft)
- PDAF (AWI Bremerhaven)
- SESAM (Grenoble)
- NERSC EnKF repository (Bergen)
- OAK (Liege)
- [DART (NCAR, Boulder, CO, USA)]

- Tool boxes developed for their particular requirements

  ➜ Keep the tool boxes, but harmonize tools in them

Examples

- Diagnostic tools
  - statistical consistency checks (innovation, etc.)
  - checking for unbiased innovations (not yet available)
- Perturbation and stochastic modeling tools
  - generate perturbations for initial ensembles
  - stochastic sources of uncertainty in models
- Transformation tools
  - Gaussian Anamorphosis
  - EOF calculations
- Utilities
  - sophisticated observation operators
  - data manipulation tools for DA
- Analysis steps (for algorithms developed in WP3)

Adapting and Developing DA tools

- WP1 identifies existing and required new tools
- WP2
  - adapts existing tools
  - develops new tools
- Follow data model and interface specified in WP1

- Provide tools together with
  - documentation
  - simple test routines
  - use ‘make’ for complex test cases

Programming Languages

Matlab/Octave .m

- reduced development time
- if CPU performance is not essential
- Matlab or Octave frequently used for
  - testing
  - data manipulation
  - post-processing

Fortran

- for tools tightly coupled to numerical models
- if CPU performance is essential
- Fortran frequently used for large-scale numerical models (NEMO, TOPAZ, HYCOM, etc.)

Adaptation of existing tools

- Various tools already exist in DA software of consortium partners
- Implementations vary
  - limited re-use
  - harmonization required

- Adapt tools to the specifications of WP1
- Performed by originator of tools (spread relatively uniform)

Development and implementation of new tools

- WP1 identifies necessary additional tools
  - required by WP4

- Discuss new tools in developer's forum to meet requirements
- Implement new tools according to standards from WP1
- Dispatch work between all partners (WP leader in charge of balanced workload)

Work distribution

- Main contributors: AWI and TU Delft
  (both also strongly involved in WP1)
- Collection of tools from all partners
- All partners involved in adaption and development

Partner	Ulg	UREAD	AWI	TUD	CNRS	NERSC
man-months	2	2	6	4	4	2

Timing of Tasks

- Creation of SVN Server (M1)
- Initial filling of SVN repository (until M6)
  - with (some) existing implementations
- Adaption and development of tools (M7 to M48)
- Codes in SVN repository updated continuously

Milestones & Deliverables:

- SVN server description (M1)
- Preparing public bundled versions (M12, M30, M48)
  - 3 versions: V0, V1, V2
- Software reports for V0, V1, V2 (Deliverable)

Green: completed

SVN Server (Task 2.1)

- SVN: version control system
- Central server for shared development
- Used internally
- Storage for
  - Documents (www, templates, reports, etc.)
  - Software codes
- Description for SVN server and structure (D2.1)
  - Standard organization for code
    (trunk/, tags/, branches/)
  - Directories for documents, templates, etc.

Software release V0 – included tools

Diagnostic tools

- sangoma_ComputeHistogram
- sangoma_ComputeEnsStats
- mutual_information
- relative_entropy
- sensitivity

Perturbation tools

- sangoma_MVNormalize
- sangoma_EOFCovar
- weakly constrained ensemble perturbations

Transformation tools

- anam_setup
- anam_transform
- anam_invtransform

Utilities

- hfradar_extractf
- PodCalibrate
- EnKF

Language: Fortran, Matlab/Octave, Java

Software release V0

- Intended to test the collaborative development
- Codes not yet adapted to data model

- Required work:
  - Adaptation to data model
  - Ensure independence from assimilation system
  - Uniform naming scheme
  - Categorization

    (add ‘AnalysisStep’ – for inputs from WP 3)

Next steps – toward release V1 (month 30)

- Adaption and addition of tools
- How to extend the selection of tools?
  - Discussion:
    - Which tools are important?
    - If it exists, which partner can provide/adapt it?
    - If new, who can implement it?
    - Fortran or Matlab? Java?

Conclusion

- WP2 results in
  - Collection of harmonized existing DA tools
  - Addition of new tools with standard data model and interface
  - Publicly available bundle of “Sangoma-Tools”
- Expected achievements
  - Improved re-use of DA tools
  - larger selection of available tools
  - simplified use of tools (documentation, test cases)
- Current status
  - initial code release (V0)
  - Adaption and extension of tools for next release (V1)