Deltares
Enabling Delta Life

Data assimilation in coastal ocean modelling and forecasting

"Application aspects - beyond the user manual"

<u>Herman Gerritsen</u>, Martin Verlaan, Julius Sumihar, Firmijn Zijl

thanks to various Deltares and SDWA colleagues

OpenDA
www.openda.org


Outline of the talk

Overview of OpenDA

Status of available model wrappers and content

1. Information content of the data
2. Calibration = finding a mathematical optimum
3. Model parameters and their interdependence
4. Ranking parameters to their uncertainty impact
5. Correlation scales: local and spatial effects
6. Kalman twin experiment - impact of observations
7. Correlation of uncertainties in time and space
8. Adequate identification of uncertainties avoids errors due to overadjustment

Summary

OpenDA in short

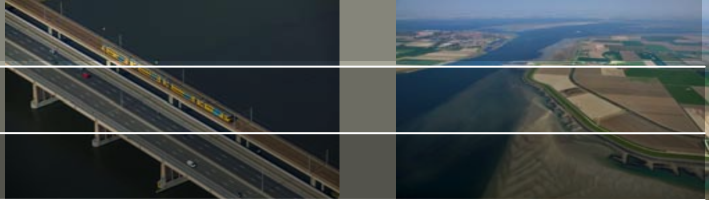
What is OpenDA?

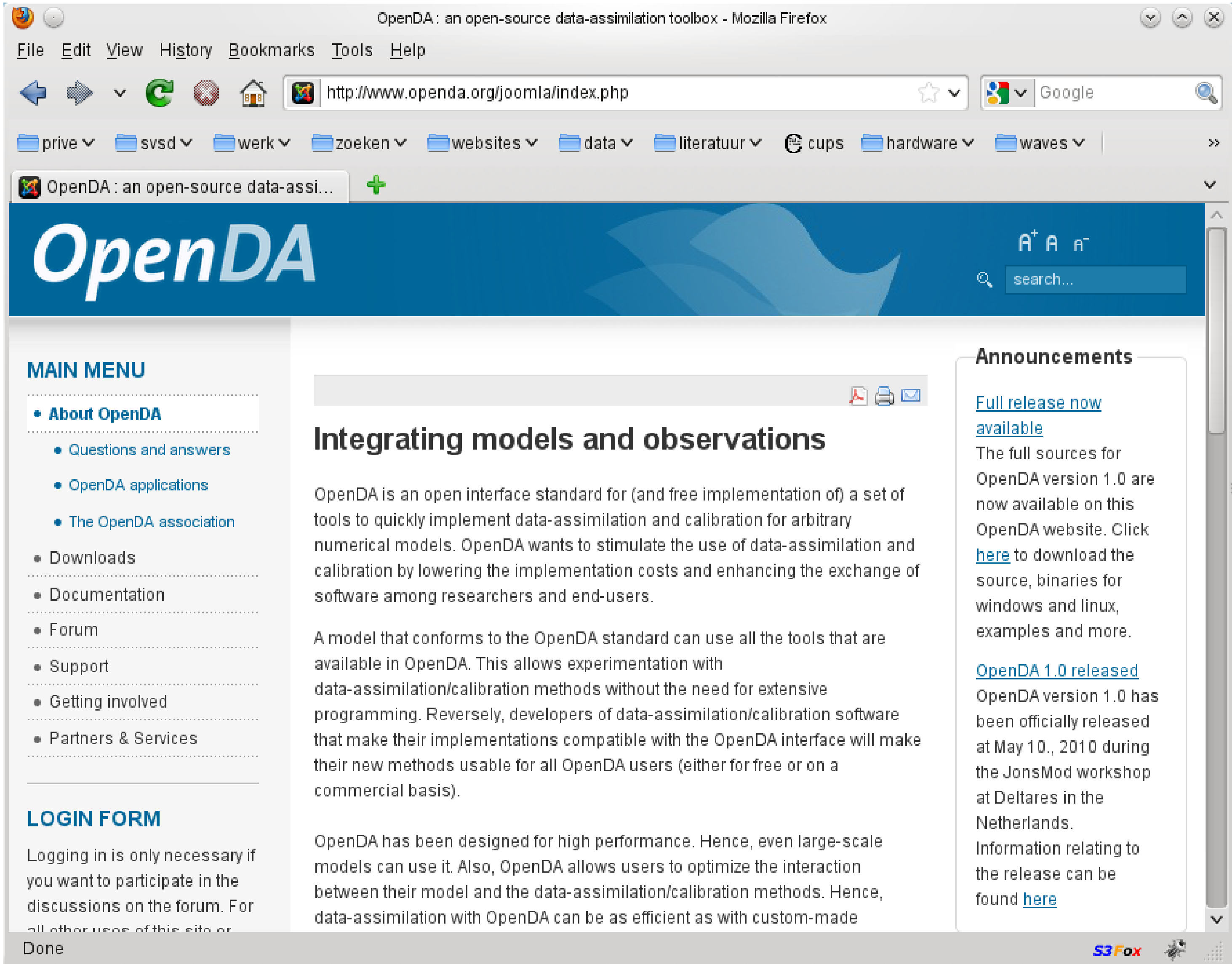
- it is a portable DA environment, suitable for any process model
- it needs one time interface development with the process model
- provides tools for uncertainty analysis, model calibration, forecast optimisation
- it is fully user configurable
- it was initially developed by TUDelft, VORtech and Deltares
- it is the merging of the earlier COSTA and DATools routines
- it contains many well-known optimisation and filter algorithms
- it is open source (www.openda.org)
- it is coupled to and embedded in FEWS
- it is used daily in several operational forecast environments
- it was launched during Jonsmod2010 (version 1.0)
- being extended continuously
- helpdesk support available after prior arrangement

OpenDA

Deltares

www.openda.org





- Download
- Documentation
- Association

Additional interace or wrapper module needed for each model

OpenDA
Deltares

Present status of available algorithms and wrappers

Calibration algorithms:

DUD, Sparse DUD, Simplex, Powell, GLUE, SCE, BFGS, L-BFGS, CG

Filter algorithms:

EnkF, SSEnKF, PF, EnSR, 3Dvar

Wrappers for calibration:

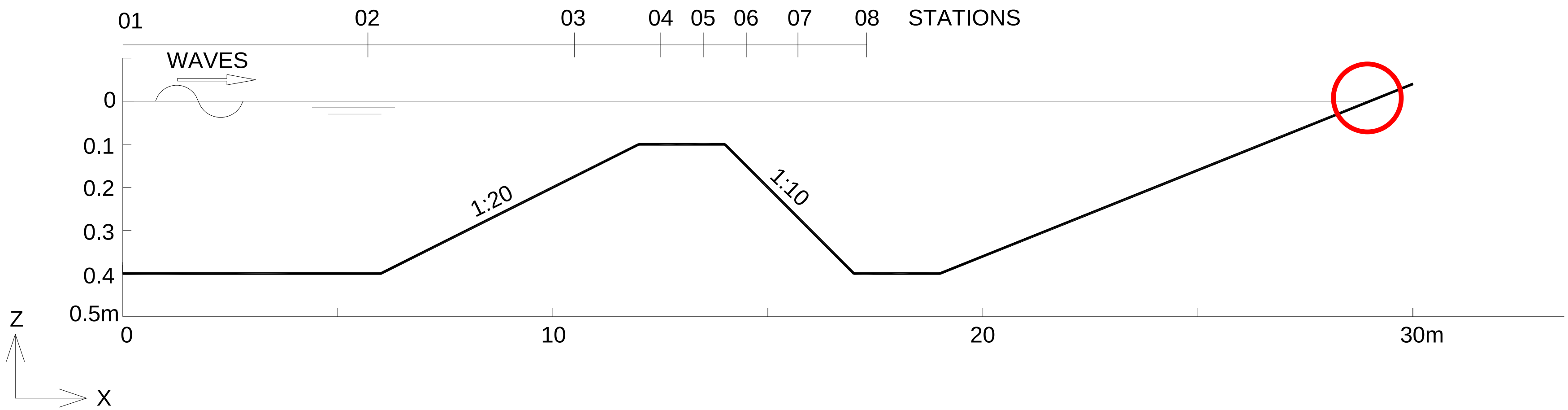
SOBEK-RE, HBV, WAQUA, Delft3D-flow, (~Delft3D-waq), SWAN(-CI), WANDA,

Wrappers for filtering:

SOBEK-RE (DAtools based), Delft3D-flow, Delft3D-waq, WAQUA, SWAN,
Chimère (VORtech+Argoss), Lotos-euros (VORtech+TNO)

- Specification of uncertainty = control variables

Information content of the data; no constraints

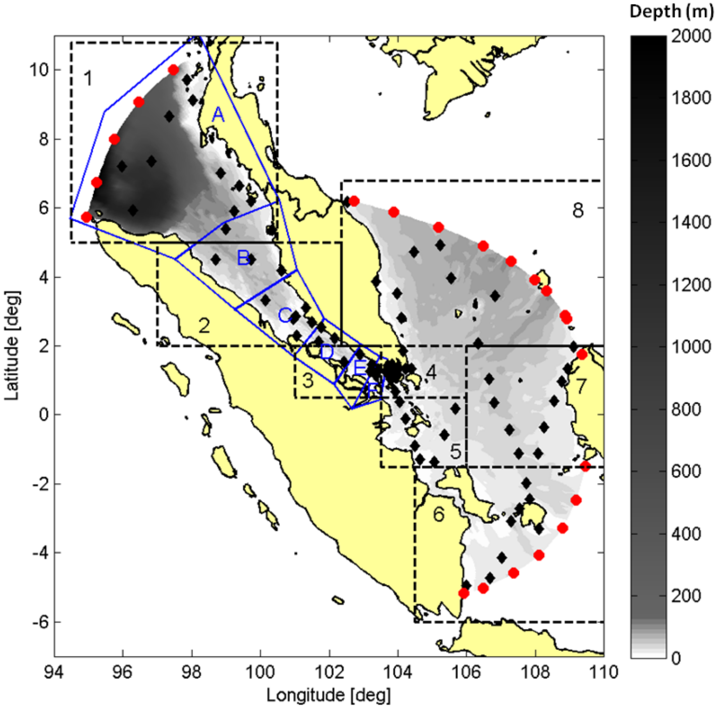


- Triad wave wave interactions in the classical one-dimensional Beji Battjes bar (flume experiment, 1993)
- Simultaneous estimation of cfjon (FRIC), gamma (BREA), trfac (TRIAD), cutfr (TRIAD)
- Breaking occurs in the surf zone (red circle)
- Data on density spectra, H_{m0} , T_p, $T_{m-1,0}$ in stations 01 – 08
- Gamma value ~20,000 (from 0.7000); other values within expectation
- → Data provided contain no information on wave breaking
- → Data are suitable for calibration of triad parameters only

OpenDA

Deltares

Calibration = Optimising a mathematical function (GoF)



- Objective: Optimise the M2 and S2 tidal amplitudes and phases (H,G) at the 5 boundary locations along the NW open boundary
- Configure the GoF criterion, taking into accounting data properties

$$GoF = \frac{1}{2}\sum_{r=1}^{r=R\max}\sum_{s=1}^{s=S\max}\sum_{n=1}^{n=N\max} w_{r,s}\left(H_{r,s,n}^{sim}(t) - H_{r,s,n}^{obs}(t)\right)^2 \Big/ \left(\sigma_{Hobs}\right)^2$$

H is the water level measured at time ***t***, ***sim*** refers to results obtained from model simulations, ***obs*** are observed values, ***Nmax*** is the number of timesteps in the time series, ***Smax*** is the number of stations in region ***r***, ***Rmax*** indicates the regions for which observations are included while σ_{Hobs} denotes the uncertainties assigned to the observations (here: tidal prediction values)

OpenDA Deltares

Calibration = Optimising a mathematical function (GoF)

Test 3: We optimise the amplitudes and phases at all 5 individual boundary support points (BSP) simultaneously,
→ a total of 5*2*2 = 20 variables

BSP	Changes made to Boundary Conditions at Different Support Points							
	M2				S2			
	Test 3		Test 3_C		Test 3		Test 3_C	
	Amp. Factor	Δ Phase	Amp. Factor	Δ Phase	Amp. Factor	Δ Phase	Amp. Factor	Δ Phase
1	1.018	-29.093	0.965	-5.542	0.887	-50.302	0.913	-12.464
2	1.100	11.720	0.965	-5.542	1.095	1.571	0.913	-12.464
3	1.117	-6.956	0.965	-5.542	1.020	-28.076	0.913	-12.464
4	1.048	0.945	0.965	-5.542	1.063	-7.598	0.913	-12.464
5	1.154	-2.128	0.965	-5.542	1.079	31.719	0.913	-12.464

Test 3_C: We optimise the amplitudes and phases at the 5 coupled boundary support points (BSP) simultaneously,
→ a total of 1*2*2 = 4 variables

OpenDA

Deltares

15

Need for parallel assessment of Process Behaviour

→ Evaluation in terms of the process physics is needed:

A very practical process indicator or error measure for tidal constituent *k* is the summed vector difference (SVD) over selected regions or for the entire model

$$VD_{k,r,s} = \sqrt{\left[\left(H_{c,k}\cos G_{c,k} - H_{o,k}\cos G_{o,k}\right)^2 + \left(H_{c,k}\sin G_{c,k} - H_{o,k}\sin G_{o,k}\right)^2\right]}$$

$$SVD_k = \sum_{r=1}^{r=R\max}\sum_{s=1}^{s=S\max} VD_{k,r,s}$$

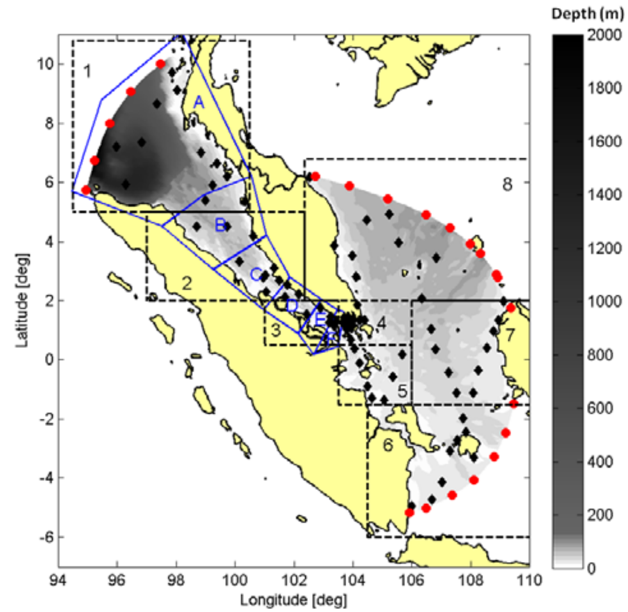
where $H_{c,k}$, $H_{o,k}$, $G_{c,k}$, $G_{o,k}$ are the computed and observed tidal amplitudes and phases of a given tidal constituent *k.*

16

Calibration = Apply GoF plus Process Indicators

Test	Parameter (*p*) varied	Observer Regions used	*P*	Iter.	GoF Initial	GoF %IMP	Remarks
1	Phase of M2 & S2	1,2	2	5	9.63E+05	12.09	
2	Amp. Of M2 & S2		2	6		29.10	
3	Phase & Amp. Of M2, S2		4	5		36.30	
4	Phase & Amp. Of M2, S2	1	4	5	2.36E+05	33.71	
5	Phase & Amp. Of M2, S2	1,2,3	4	11	1.85E+06	41.26	
6	Phase & Amp. Of M2, S2	1,2,4	4	11	2.34E+06	37.72	
7	Depth in Region 3	3	1	4	5.48E+05	57.50	Starting point was the optimum result in Test 3
8	Friction in Region 3	3	2	18		3.19	
9	Depth, Friction (Region 3)	3	3	17		58.65	

Test	M2 Overall Model Initial SVD (m)	M2 Overall Model %IMP	M2 Observer only Initial SVD (m)	M2 Observer only %IMP	S2 Overall Initial SVD (m)	S2 Overall %IMP	S2 Observer only Initial SVD (m)	S2 Observer only %IMP
1	11.56	11.00	3.39	8.24	5.61	2.07	1.87	0.75
2		14.14		30.51		15.27		26.00
3		22.10		32.52		15.02		27.88
4		9.17	1.17	29.86		0.46	0.81	42.31
5		25.18	5.36	31.42		17.63	2.64	24.71
6		24.74	6.75	28.32		17.85	3.34	22.54
7	9.00	22.35	1.97	69.40	4.77	25.17	0.77	66.35
8		2.13		4.42		2.91		5.66
9		24.13		66.67		26.95		67.61

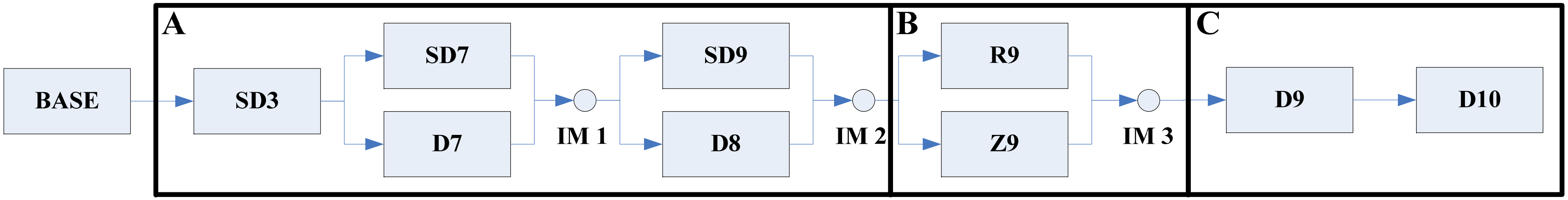


Observer regions.

- Overall assessment
- GoF but also tide (M2 S2)
- %IMP and SVD

OpenDA
Deltares

Calibration = Rank parameters and iterate



The SRM tidal model:

- 3 open boundaries: AS, SCS, JS (17 BSP's)
- 8 evaluation regions
- 6 Malacca Strait "bed blocks"

Number of uncertain parameters:

- (H, G) of 4 semidiurnals: M2, S2, N2, K2 in 17 BSP's
- (H, G) of 4 diurnal tides: O1, K1, Q1, P1 in 17 BSP's

- Bed level factor (6 blocks)
- Bed friction factor (6 blocks)

Total uncertain model parameters:

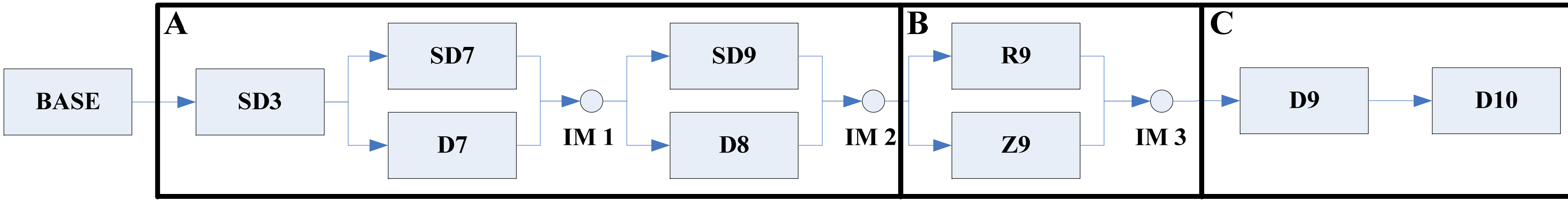
2*4*17 + 2*4*17 + 6 + 6 = 284

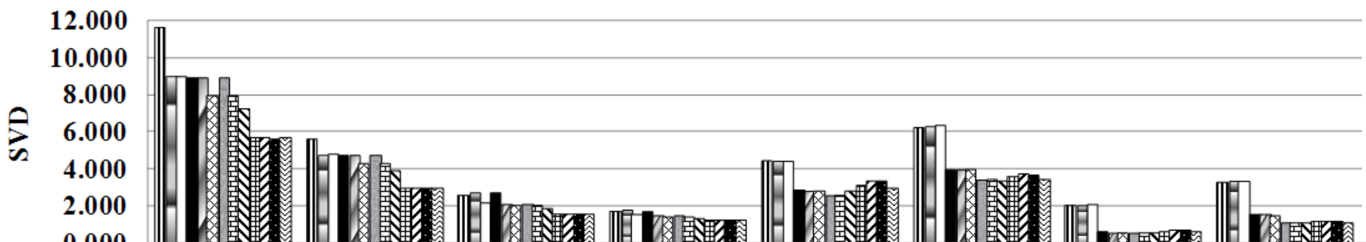
→ Rank the parameters to the expected impact of their uncertainty

→ Start with a first set, then the next set; then iterate

→ This requires process insight!

Calibration = Rank parameters and iterate



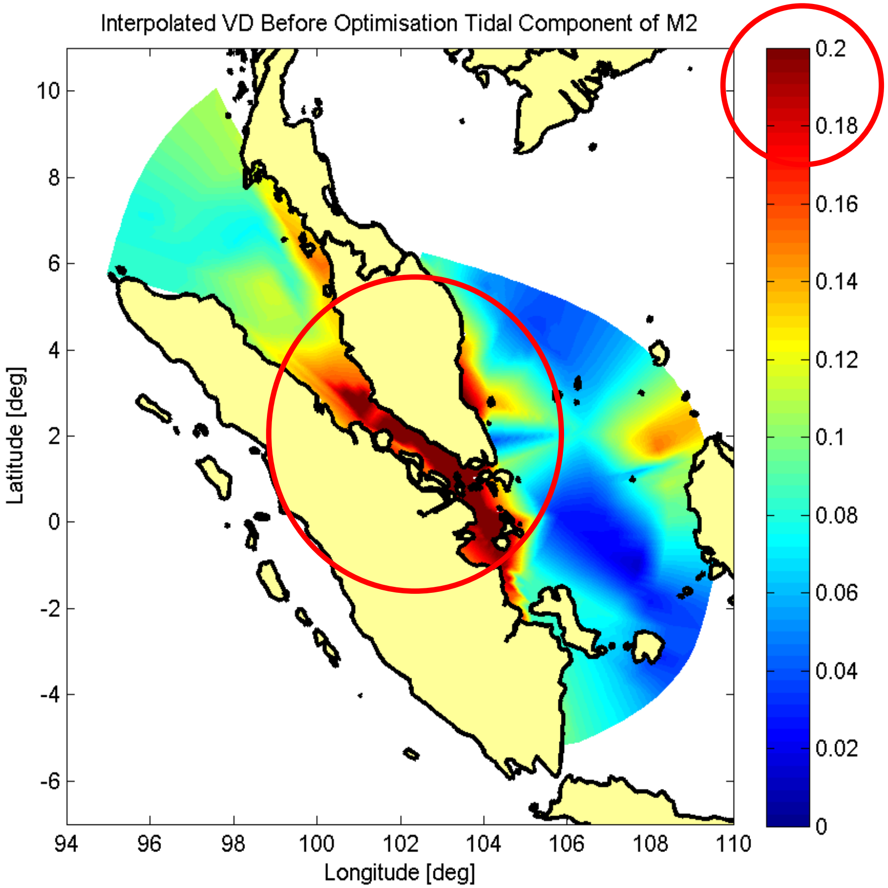


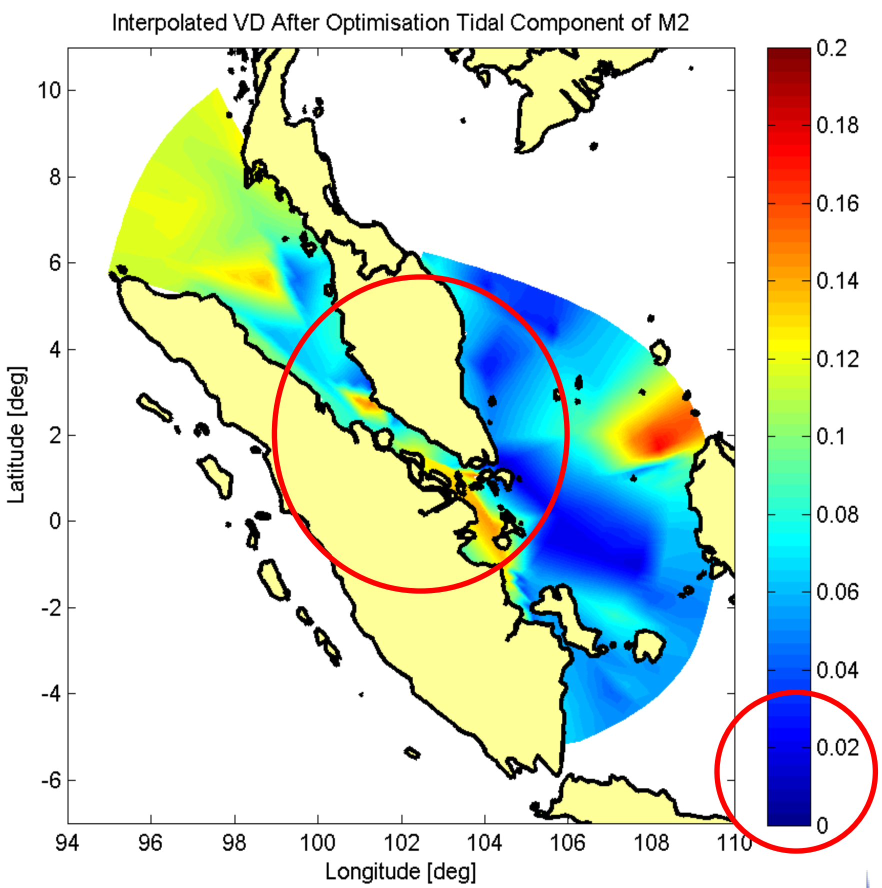
	M2	S2	N2	K2	O1	K1	Q1	P1	Total SVD
Base	11.558	5.614	2.634	1.743	4.478	6.267	2.023	3.292	37.609
SD3	9.004	4.771	2.786	1.791	4.435	6.319	2.050	3.335	34.491
SD7	8.996	4.769	2.204	1.560	4.439	6.336	2.073	3.342	33.719
D7	8.926	4.754	2.737	1.737	2.857	3.969	0.614	1.558	27.152
IM 1	8.920	4.751	2.158	1.519	2.854	3.990	0.618	1.560	26.370
SD9	7.969	4.278	2.065	1.443	2.831	3.974	0.614	1.543	24.717
D8	8.920	4.752	2.157	1.516	2.614	3.479	0.597	1.156	25.191
IM 2	7.966	4.284	2.065	1.449	2.587	3.457	0.591	1.131	23.530
R9	7.268	3.879	1.887	1.326	2.801	3.409	0.607	1.117	22.294
Z9	5.692	2.984	1.623	1.268	3.181	3.624	0.684	1.184	20.240
IM 3	5.659	2.950	1.591	1.276	3.375	3.765	0.712	1.237	20.565
D9	5.654	2.952	1.592	1.277	3.361	3.721	0.708	1.230	20.495
D10	5.681	2.957	1.593	1.259	2.966	3.479	0.639	1.147	19.721

OpenDA

Deltares

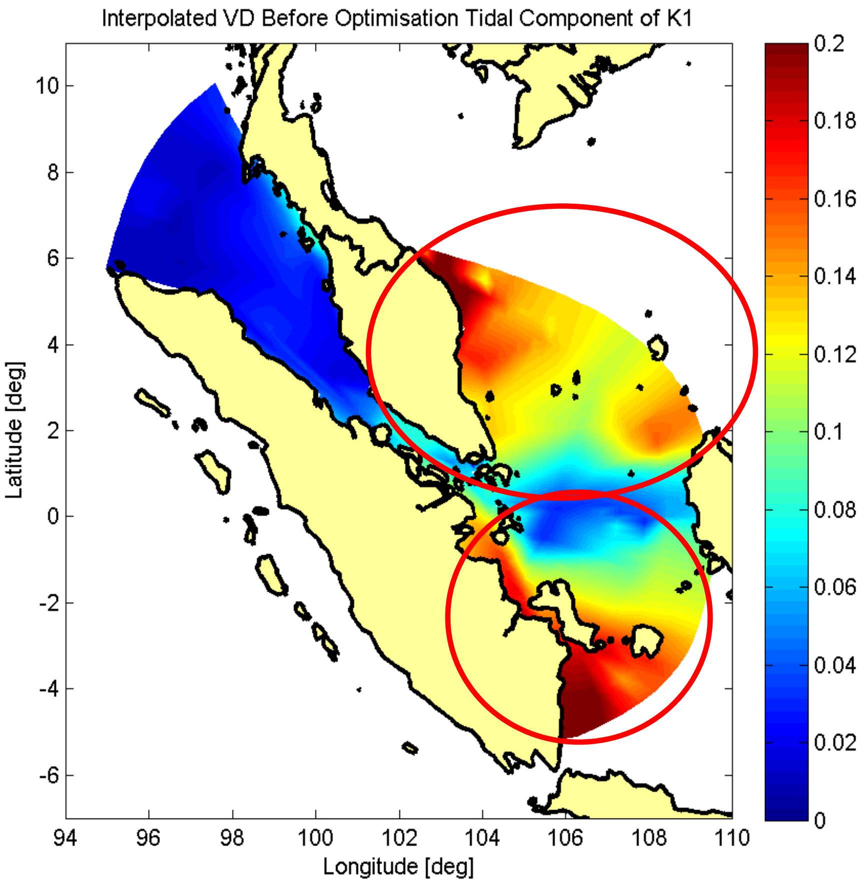
Calibration - Improvement of M_2 (in terms of VD)

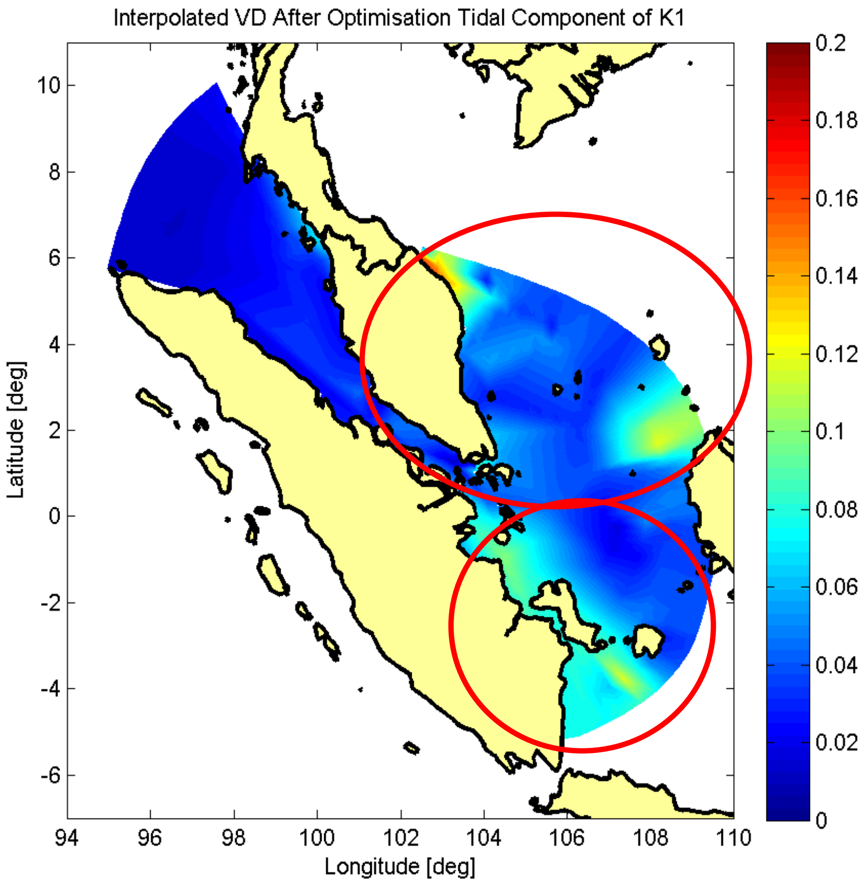




OpenDA
Deltares

Calibration - Improvement of K_1 (in terms of VD)





OpenDA

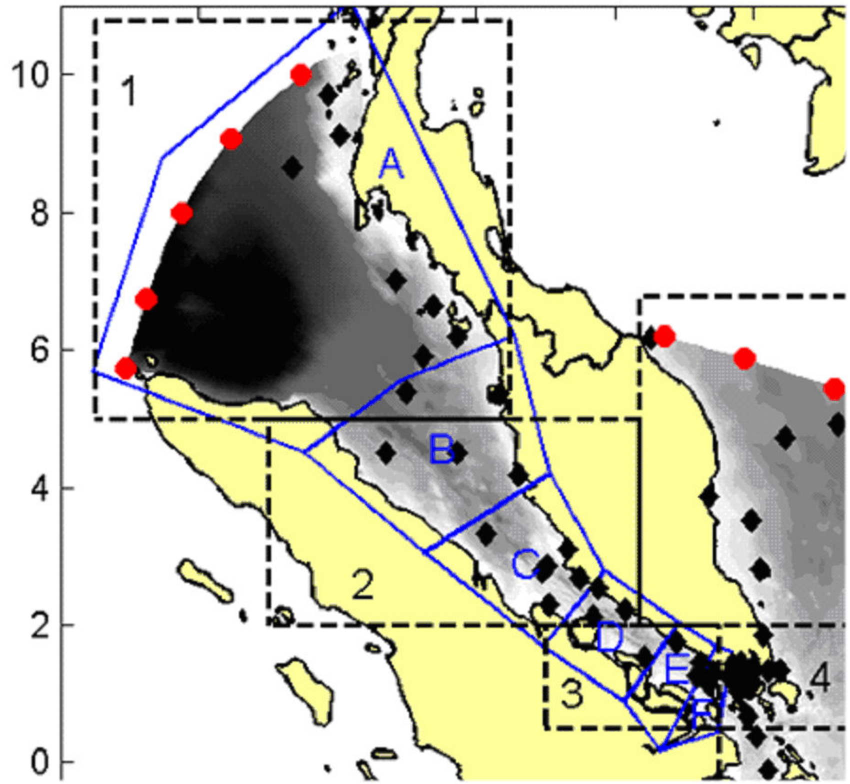
Deltares

Calibration – local bed effects and spatial impact

Sensitivity analysis bed level uncertainty (factor α)

- Blocks A – F; α_A, α_{AB}, α_{DEF}, etc.
- In terms of local GoF improvement

Variable	#Obs	#Iter	%GoF
α_A	7	20	-14.9 %
α_B	5	4	- 8.7 %
α_C	6	4	- 6.0 %
α_D	4	3	- 1.1 %
α_E	3	6	-62.3 %
α_F	3	23	- 5.1 %
α_{AB}	12	4	- 6.8 %
α_{CD}	10	3	- 8.7 %
α_{EF}	6	5	-72.8 %
α_{ABC}	15	5	- 0.0 %
α_{DEF}	10	9	- 40.0 %
α_{ABCDEF}	25	7	- 20.1 %



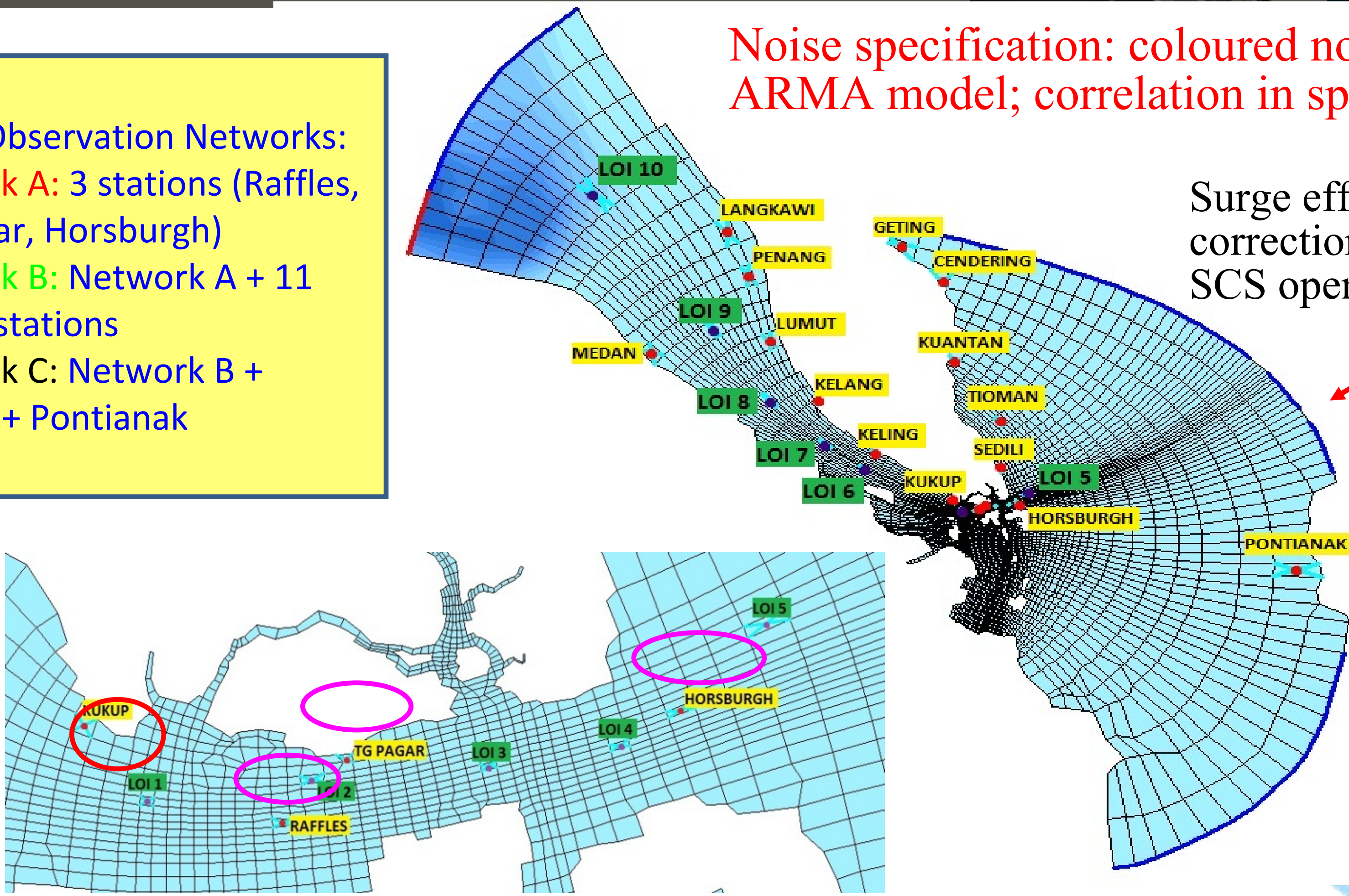
OpenDA
Deltares

22

Kalman filter Twin experiment – 3 data networks

Three Observation Networks:
Network A: 3 stations (Raffles, TG-Pagar, Horsburgh)
Network B: Network A + 11 UHSLC stations
Network C: Network B + Medan + Pontianak

Noise specification: coloured noise; ARMA model; correlation in space and time

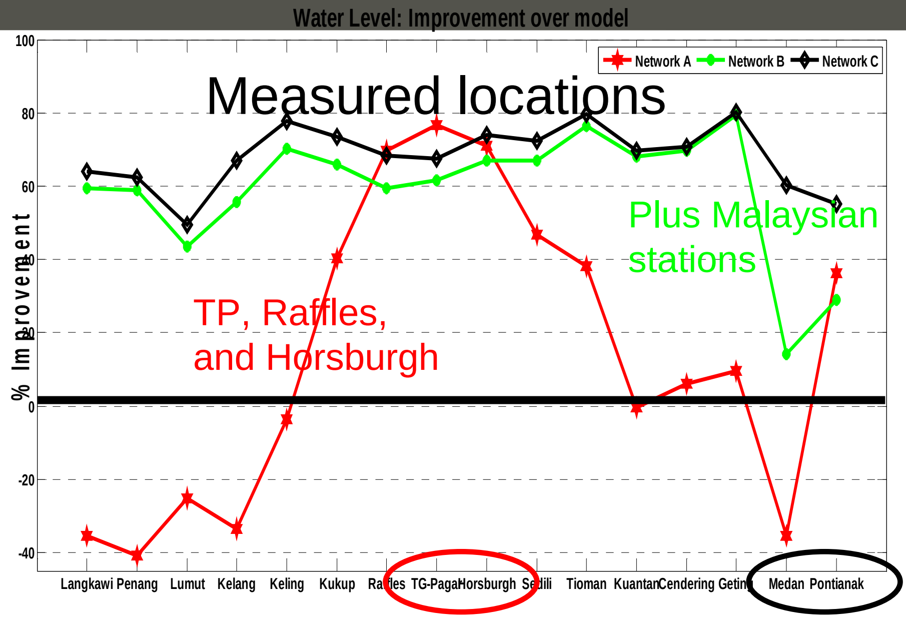


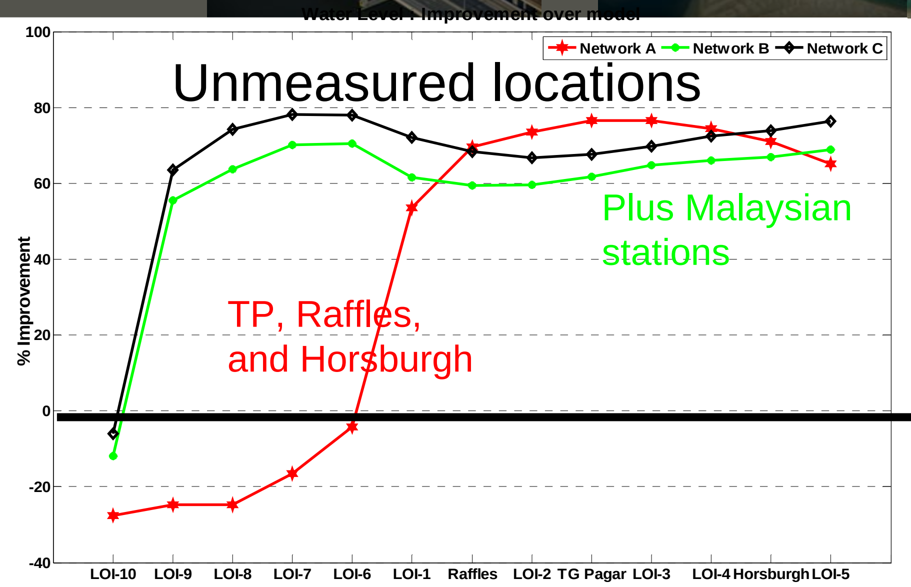
Surge effects via correction of the SCS open boundary

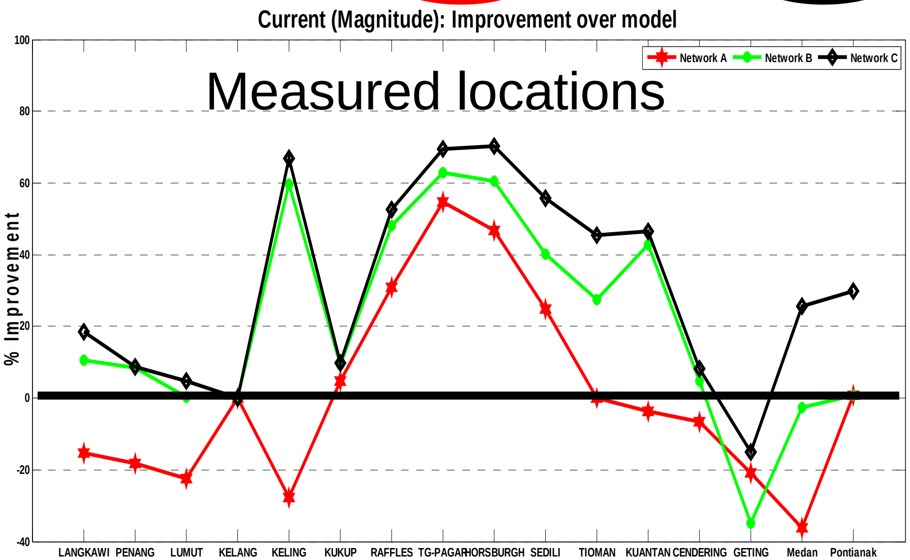
To verify quality of estimates / forecasts at additional 10 unobserved “locations of interest” (LOI).

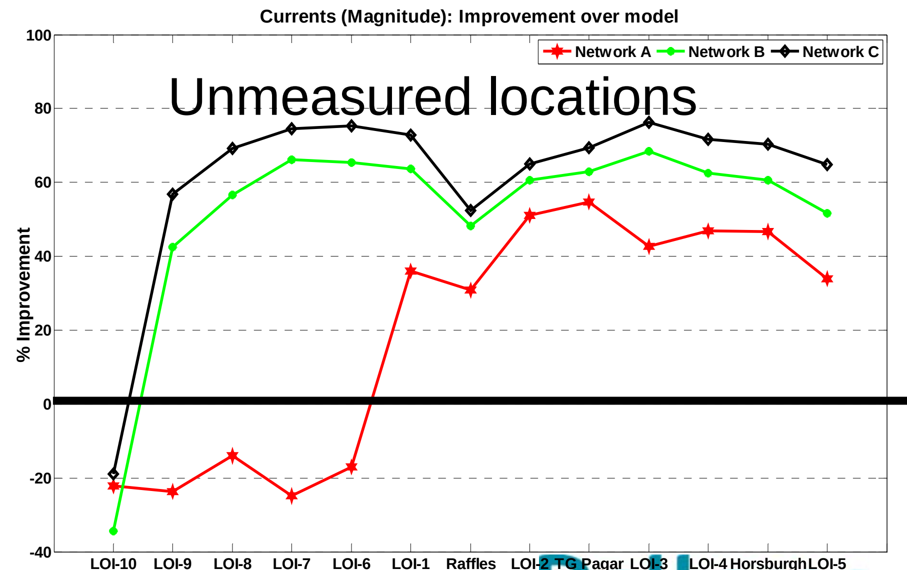
OpenDA
Deltares

% Hindcast improvement for the 3 data networks



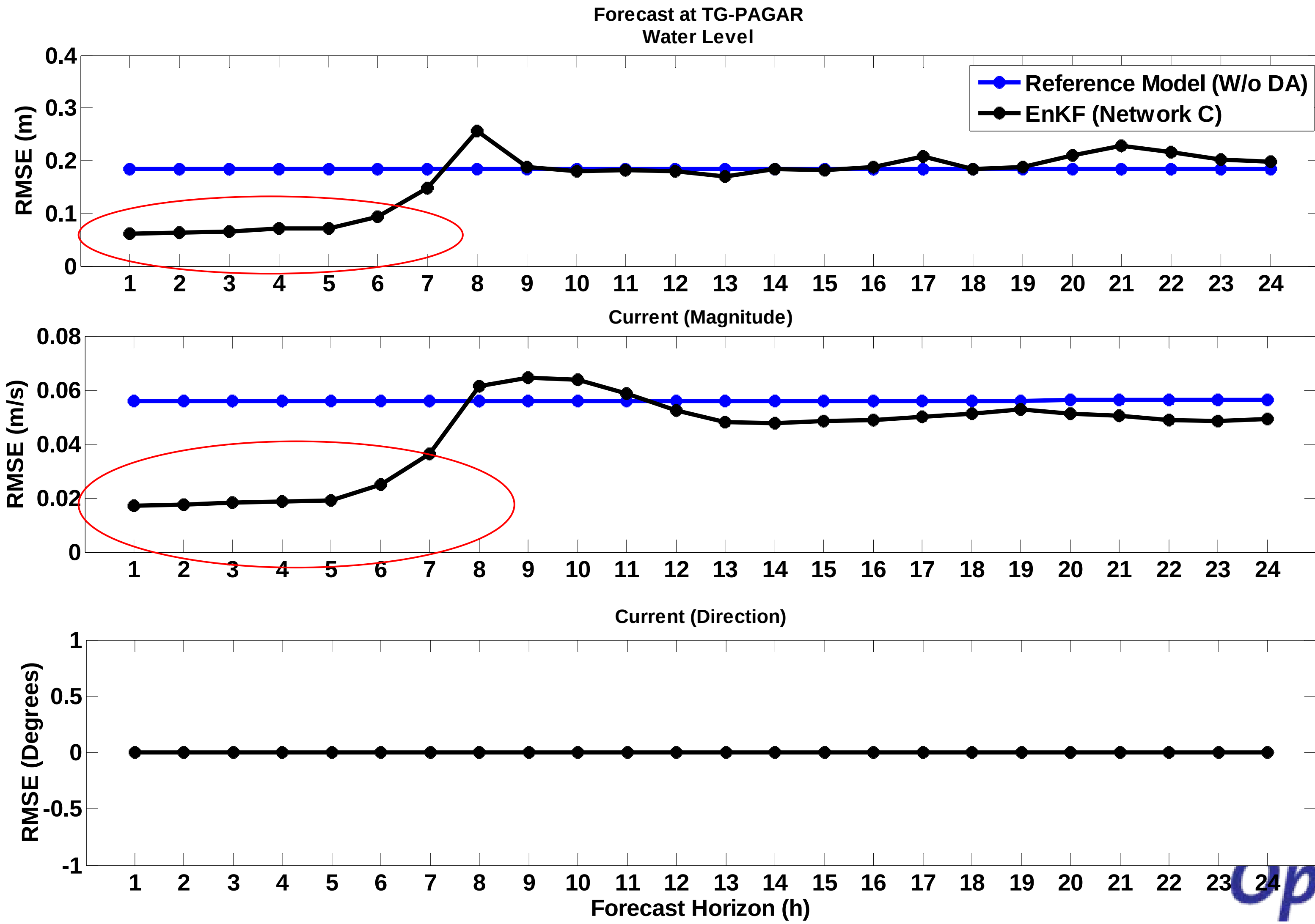




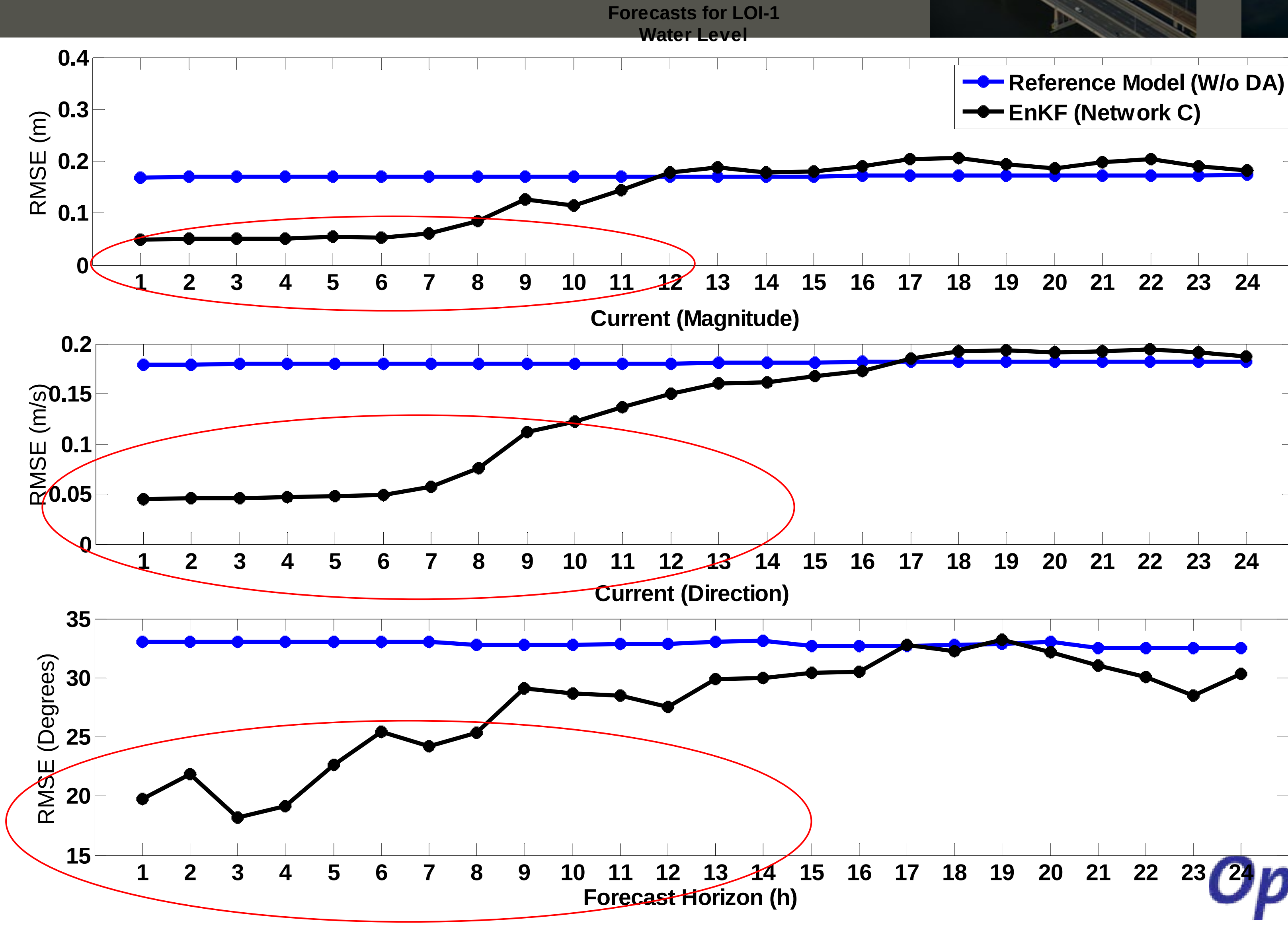


Deltares

Forecast RMSE statistics at T Pagar for T = 1, 2, .. 24 hrs



Forecast RMSE statistics at LOI-1 for T = 1, 2, .. 24 hrs




Summary of the observations

- DA provides a mathematical tool that should not be used as a black box
- Observation data need to be checked and validated for consistency, coherence, removal of outliers
- Check the information content of the data in view of the selection of uncertain parameters and assimilation potential and objective
- Define and evaluate process indicators that provide information on the physics in parallel to the mathematical optimisation of the GoF
- Information on correlation scales of observation data guides the problem specification – avoid data correlation and so minimise non-uniqueness
- Optimising only part of the uncertain model parameters implies ranking - selecting first those with largest expected impact, and iteration; this requires insight and judgment and cannot easily be automated
- Inadequate or incorrect specification of uncertainties / adjustable parameters quickly leads to “overadjustment”, that is, negative improvement or deterioration in space and / or time
- Only proper use of the software leads to good results

Questions?

Data assimilation scope

- Calibration of model parameters (the model)

- Real time updating of forecast (the information)
  - Kalman filtering
  - Variational algorithms

- Uncertainty analysis (the uncertainties)

- Study potential value of new observations/ optimisation of monitoring networks (the data)

- Estimating sources (e.g. rainfall; the forcing)