Historical covariance creation and one *in situ* data consideration to define the best geographic position for satellite data calibration

G. Pennucci, A. Alvarez and C. Trees

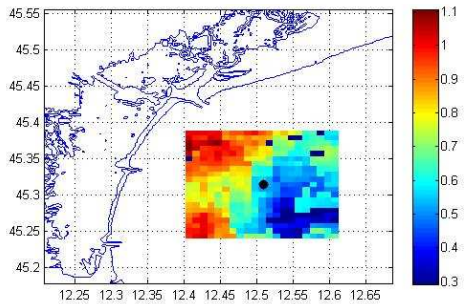
Statistic analysis of the time-series

The covariance matrix have been evaluated using the available monthly time-series MODIS acquisitions from 2005 to 2009:

Month	2005	2006	2007	2008	2009	MODIS ORIGINAL IMAGES	Sub-area selection (box of 21 x 21 Km around Venice station)
Jan	No	No	No	Yes	Yes	151 images	8 frames
Feb	No	No	No	Yes	Yes	143 images	29 frames
Mar	No	No	No	Yes	Yes	156 images	27 frames
Apr	No	No	No	Yes	Yes	144 images	30 frames
May	Yes	Yes	Yes	Yes	Yes	275 images	73 frames
Jun	No	No	No	Yes	Yes	148 images	32 frames
Jul	No	No	No	Yes	Yes	149 images	46 frames
Aug	No	No	No	Yes	Yes	154 images	52 frames
Sep	Yes	Yes	Yes	Yes	Yes	365 images	81 frames
Oct	No	No	No	Yes	Yes	150 images	29 frames
Nov	No	No	No	Yes	Yes	150 images	16 frames
Dec	No	No	No	Yes	Yes	150 images	15 frames

For each month:

+ Uncertainly map of 21km^2 (historical covariance)



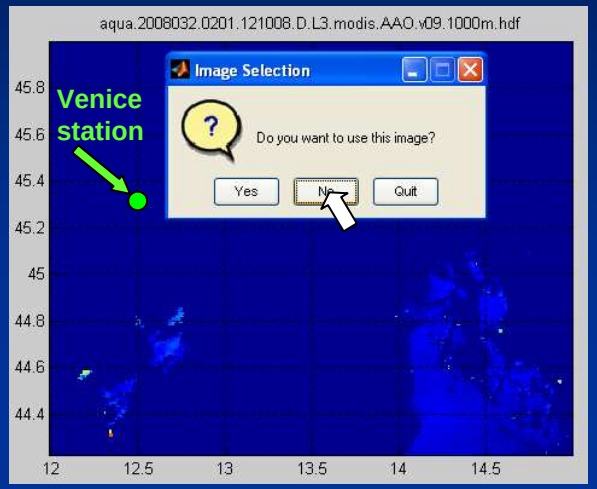
+ Covariance of the time series;

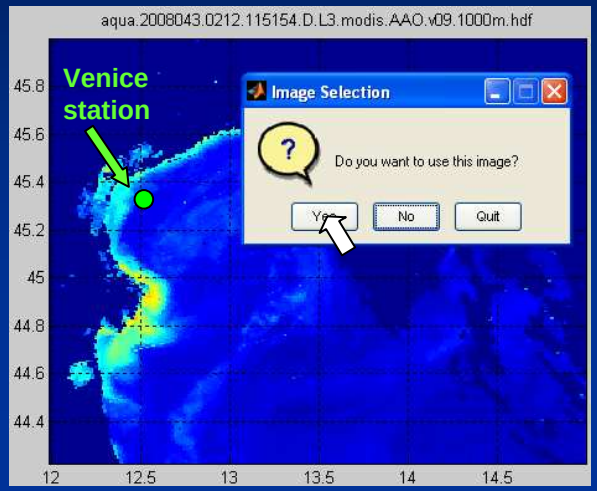
+ Mean of the time series.

The historical covariance map (C_{HIST}) represents the pixels standard deviation of considered time-series. C_{HIST} , the covariance (C) and the mean of the time-series well represent the monthly statistic analysis.

An example of implementation (1/4):
nLw 547 [nm] from February 2008-2009

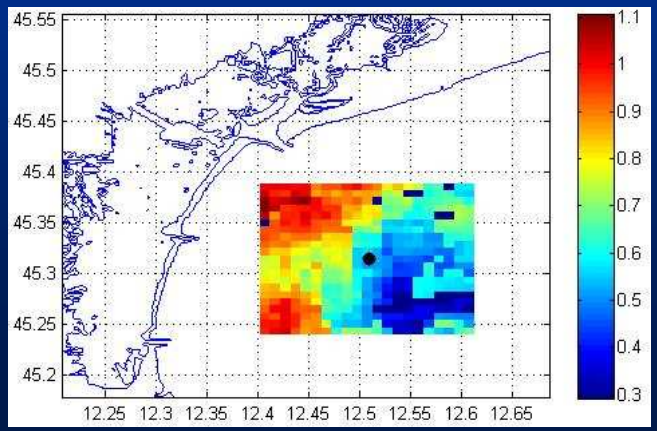
The user selects the images that can be used to implement our statistical processing.



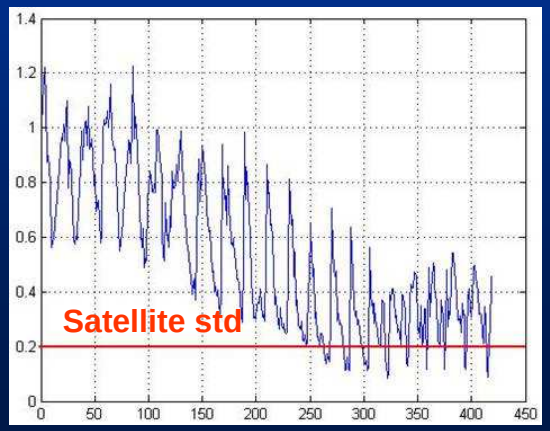


On the resulting time-series (29 frames), the user defines a bounding box around the *in situ* station (e.g. 21 by 21 km) on which implement the procedure.

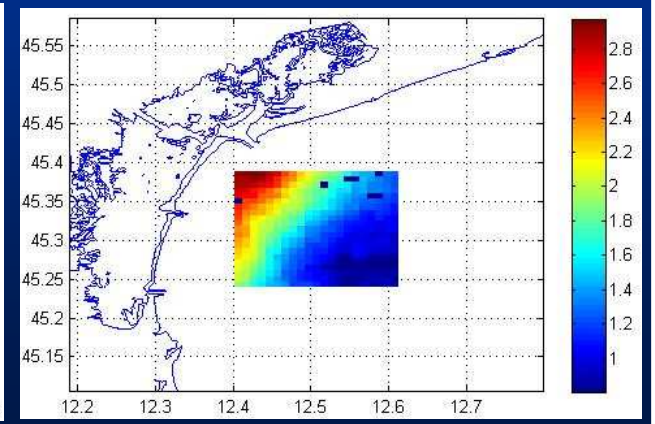
HISTORICAL COVARIANCE



COVARIANCE (diagonal)

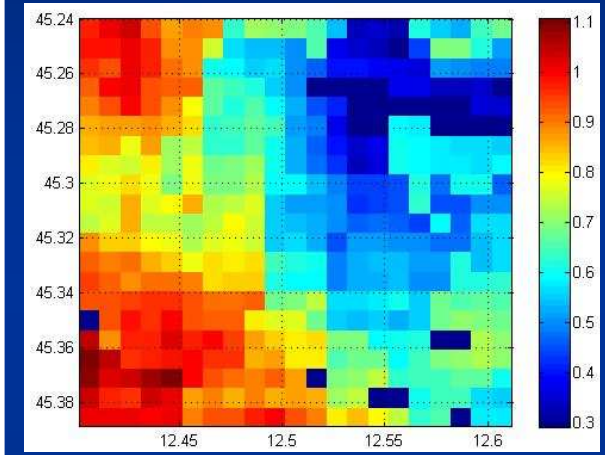


TIME-SERIES MEAN

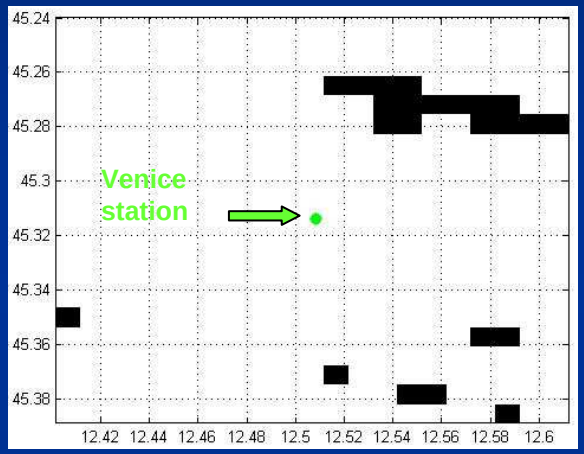


An example of implementation (2/4):
IN SITU LIST

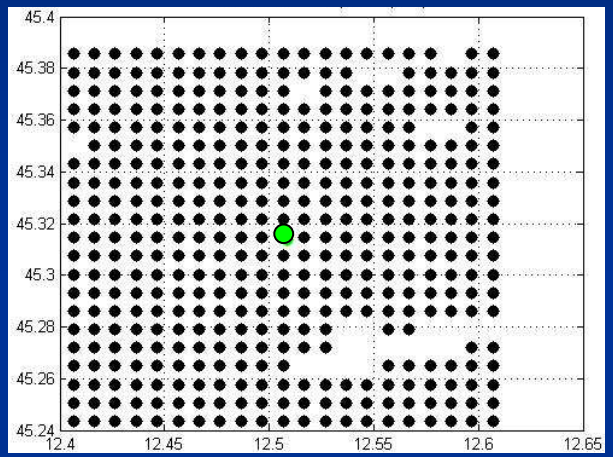
To find “the best” geographic position for an *in situ* we merge it with the historical_C map and we find “the best” result (varying its geographic position) in terms of optimal design (A-optimum and G-optimum).



HISTORICAL COVARIANCE
(UNCERTAINTLY MAP)



MASK:
0 (black) = bad-data (noise/clouds)
1 (white) = good data (sea)
Green dot = Venice station position

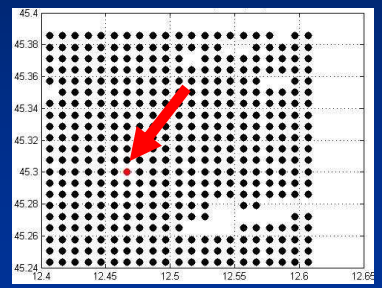


ALLOWED IN SITU POSITIONS
(center of the pixel)

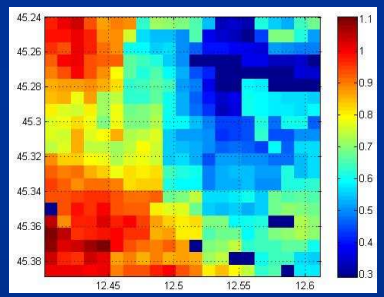
Each allowed *in situ* (i= 1:418) have to be merged with the historical map.
That merging have to be done <u>without satellite information (but just with one in situ data)</u>.

An example of implementation (3/4): MERGING

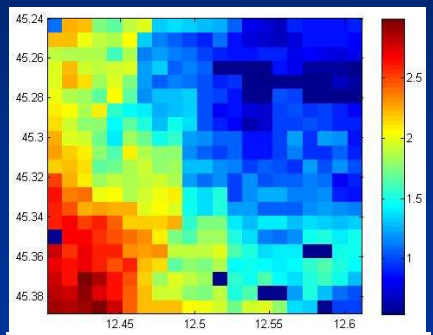
We consider the generic *in situ*$_i$, for example: Lat$_i$ = 12.4666 (Ya), Lon$_i$ = 45.3003 (Xa)
nLw$_i$ = C_hist +/- std_obs = 0.7 nm. We merge it with the retrieved historical map



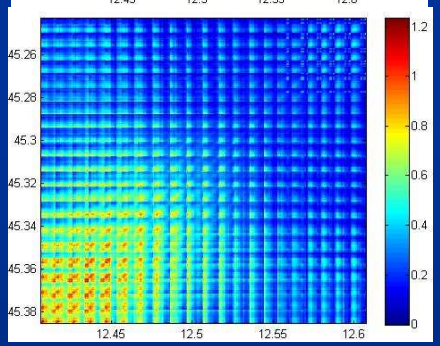
$$P(\psi_k) \propto e^{-(\psi_{obs}-H\psi_k)^T C_{obs}^{-1}(\psi_{obs}-H\psi_k)-(\psi_k-\overline{\psi})^T C_{HIST}^{-1}(\psi_k-\overline{\psi})}$$



Merging procedure



MERGED IMAGE



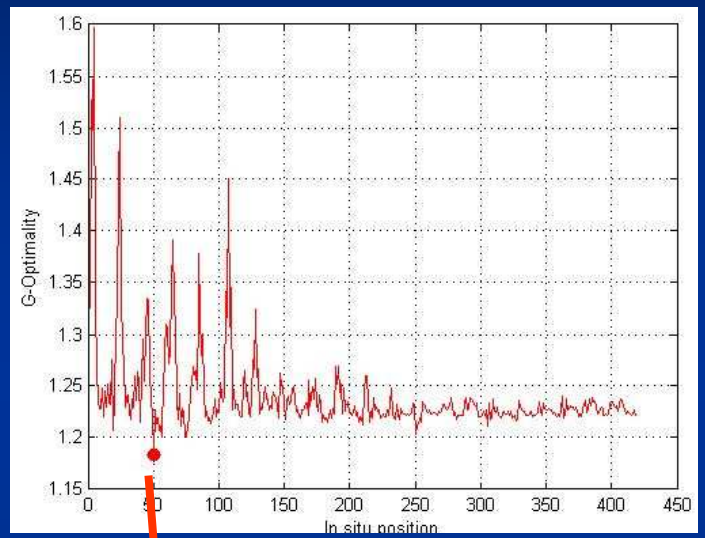
VARIANCE (Sigma_e)

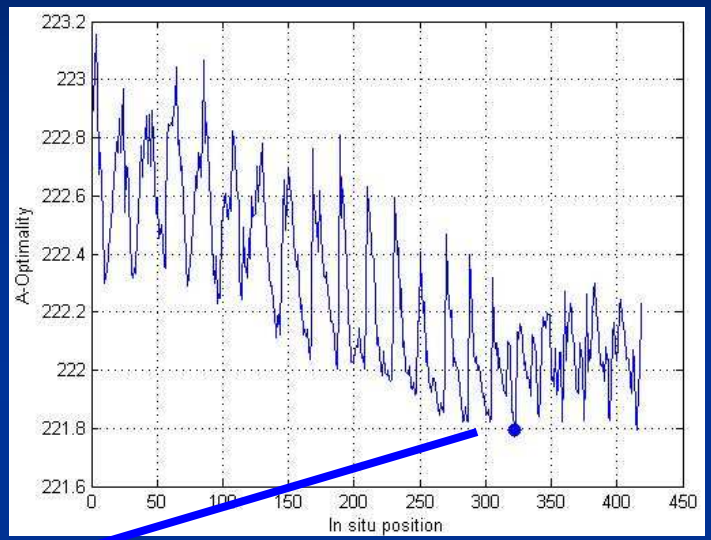
The merging procedure produces an error map (Sigma_e) and a merged image.
Our idea is to convert Sigma_e in one number in order to have an index (Uncertainly Index-UI) that can be used to identify the "best solution" (best *in situ* position) .
To convert the Sigma_e in one number we use two methodology (Experimental Design):

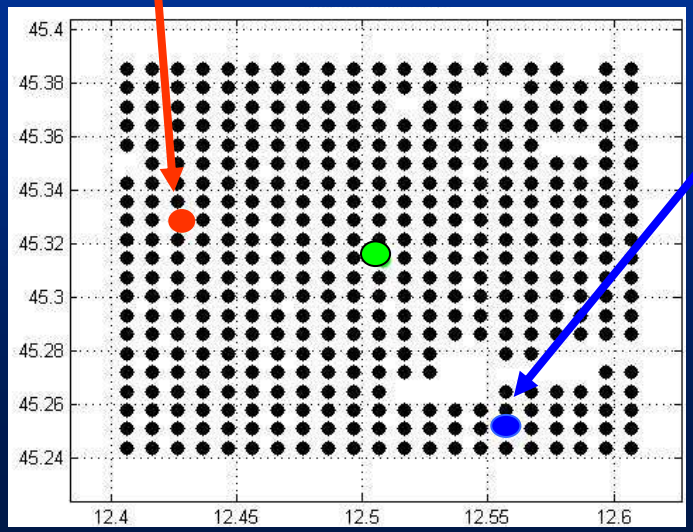
1. **A-OPTIMAL design.** It minimize the trace of Sigma_e;
2. **G-OPTIMAL design.** It minimize the maximum of the standard deviation of Sigma_e

An example of implementation (4/4):
ITERATION PROCEDURE AND RESULTS

The previous procedure was iterated for each *in situ* (i= 1:418) and the error map results were converted in index using two procedures: A-optimal design and G-optimal design.

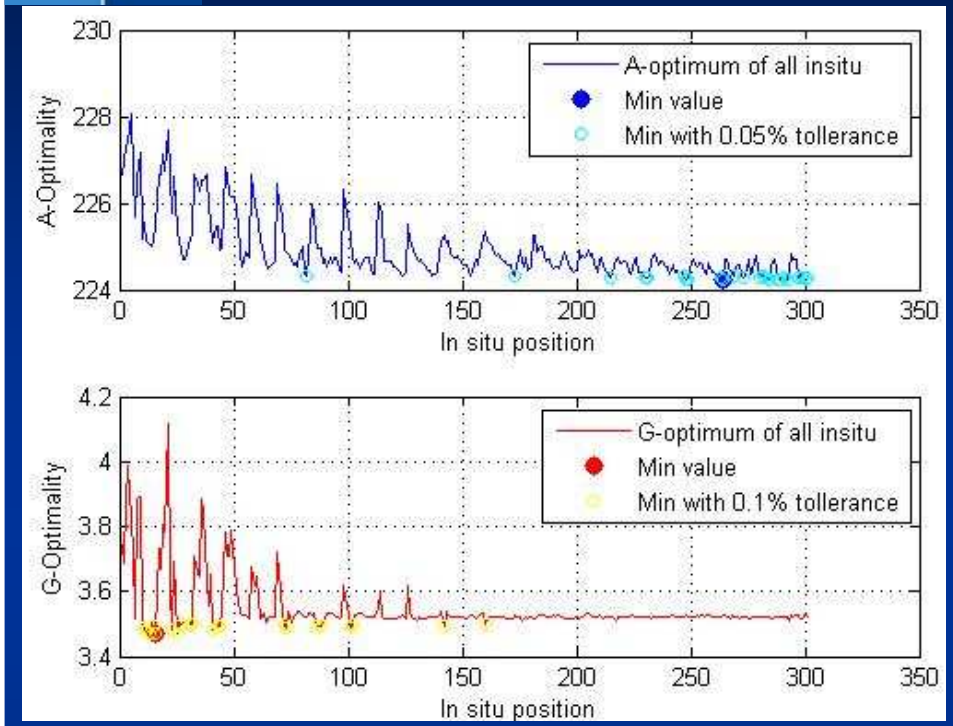


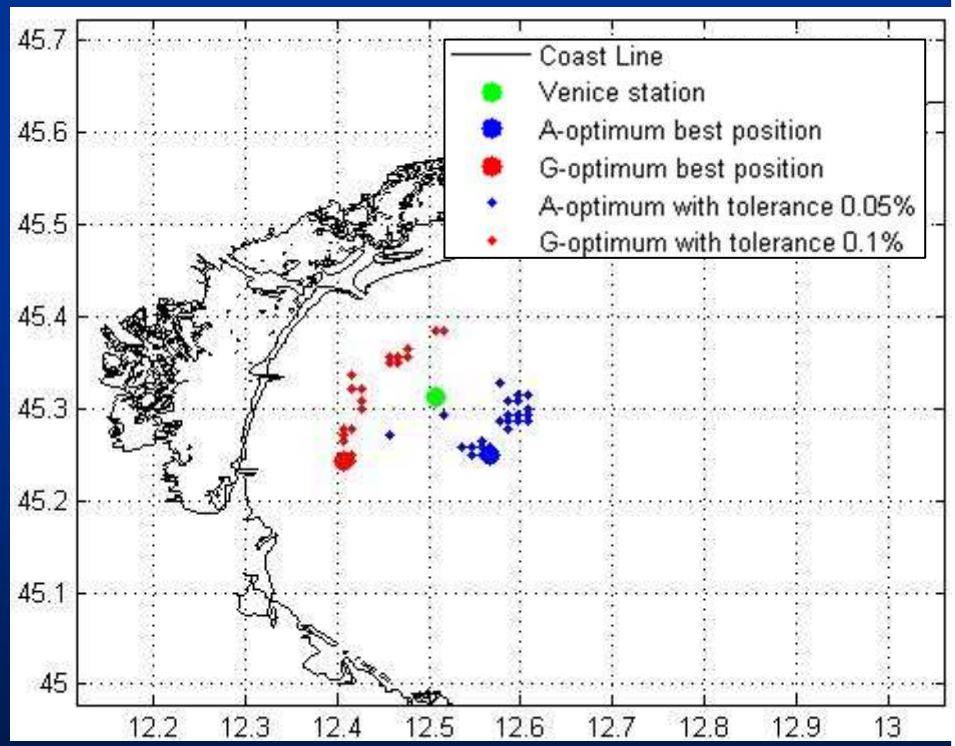




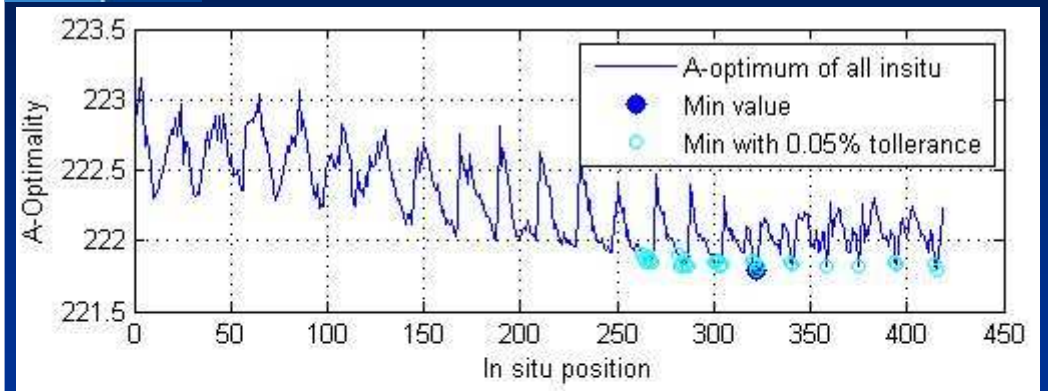
The minimum value of A (or G) represents the best *in situ* position in terms of "statistical information".
To extend those two points and provide areas, the results in next slides will consider also points closer (0.05%-0.1%) to the minimum.

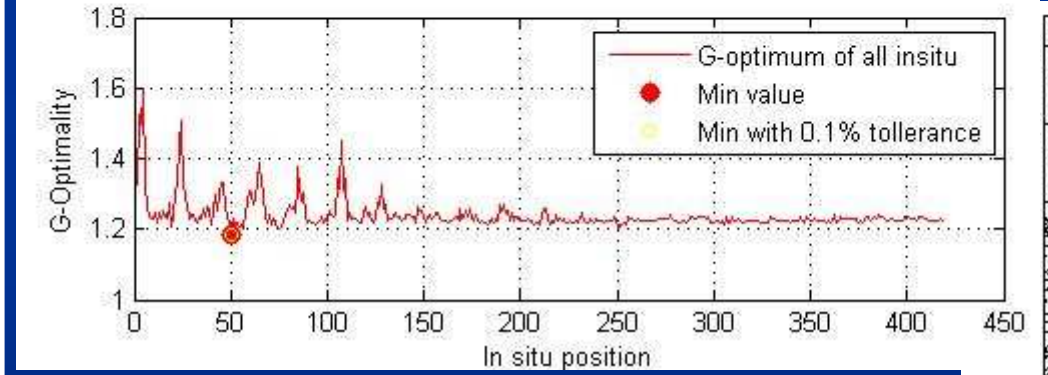
January (2008-2009) results

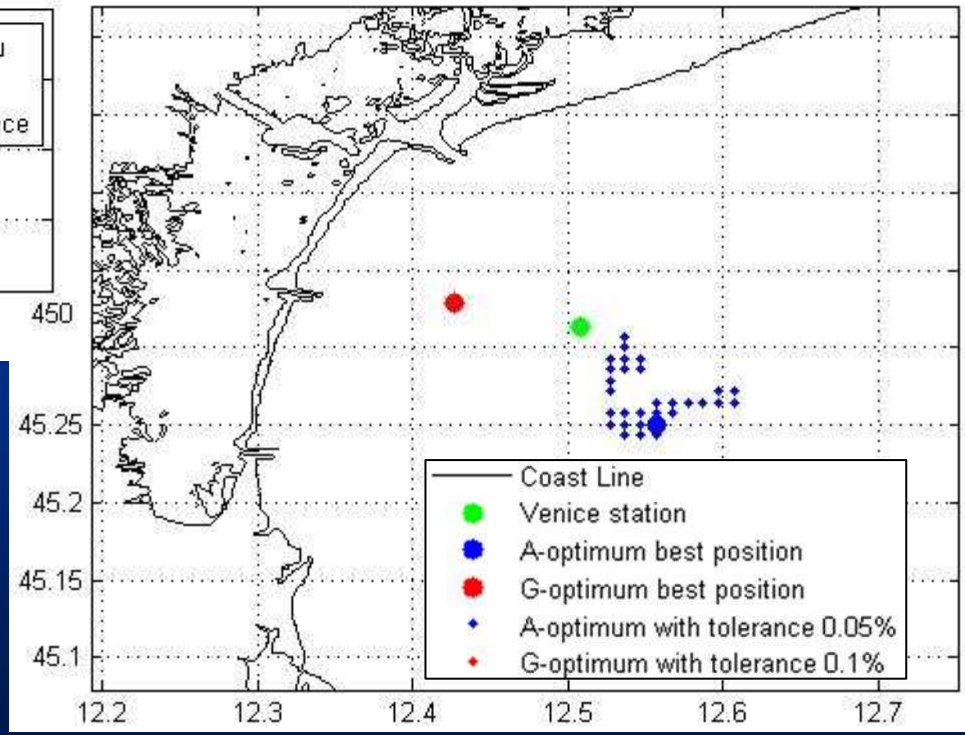




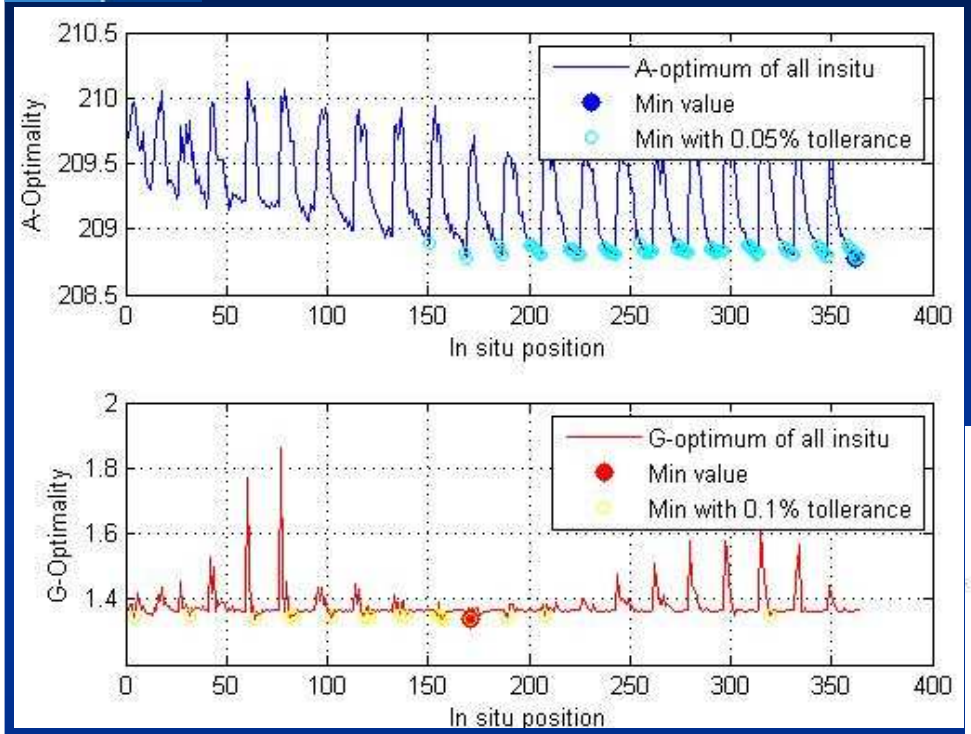
February (2008-2009) results

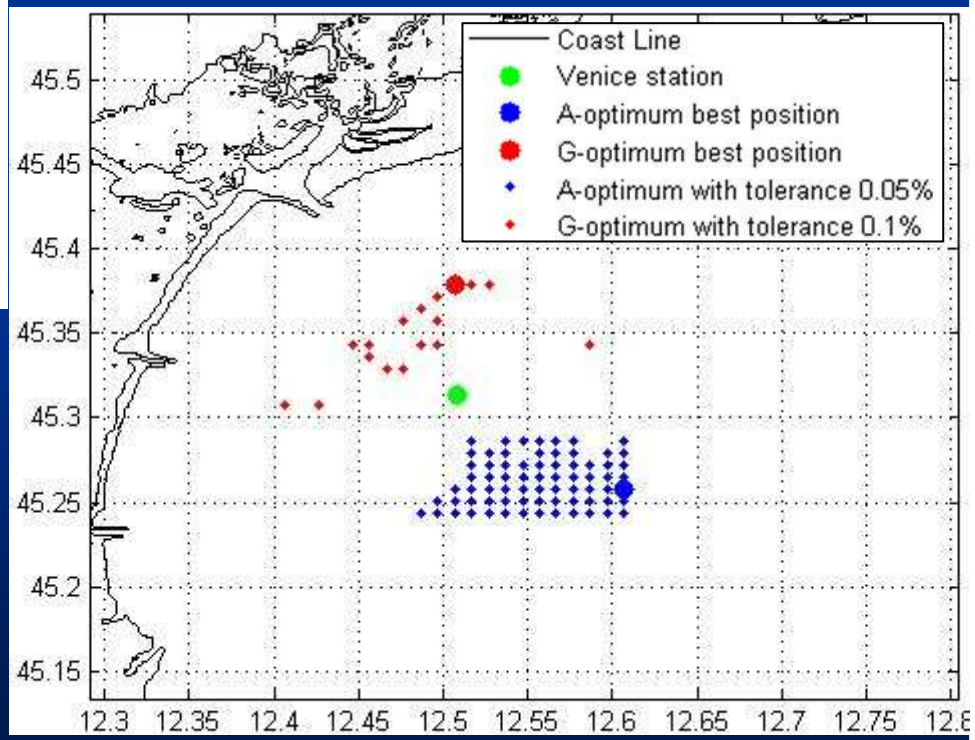




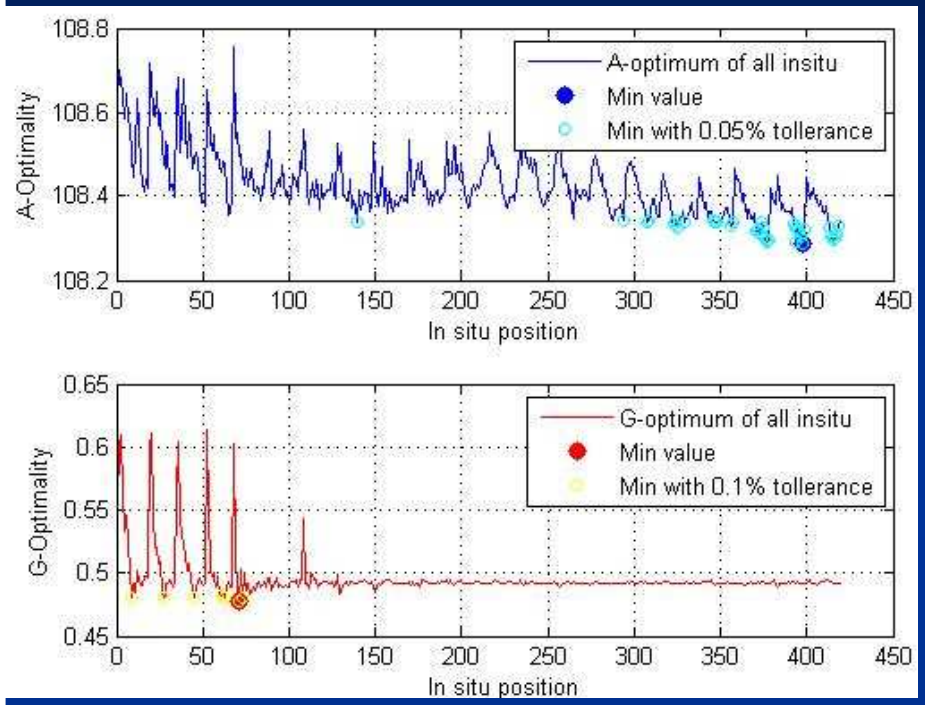


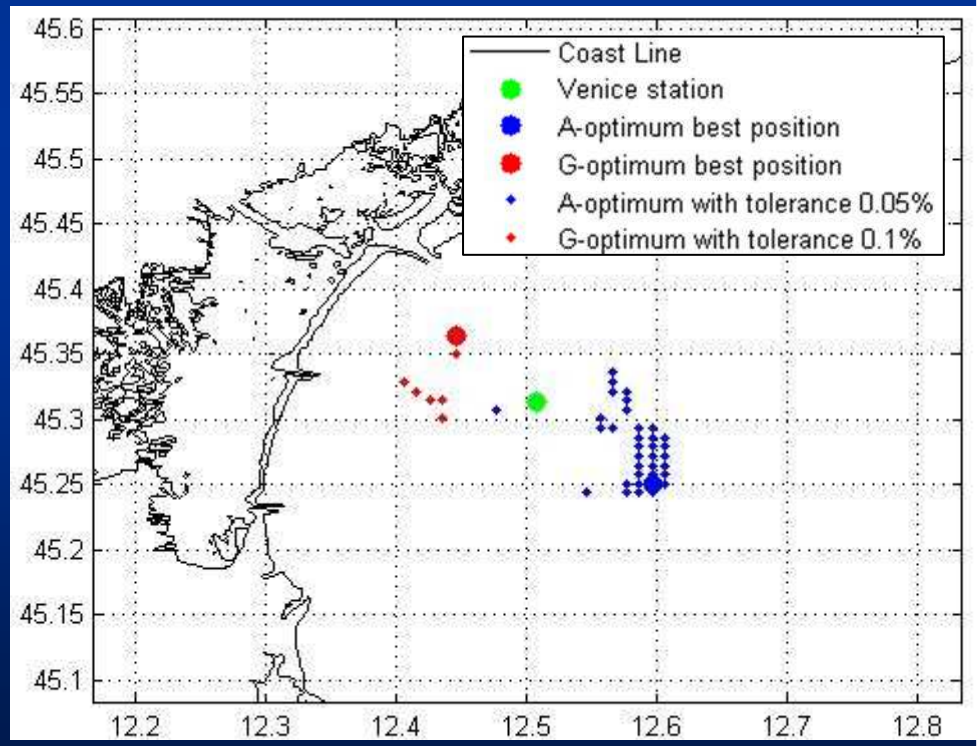
March (2008-2009) results



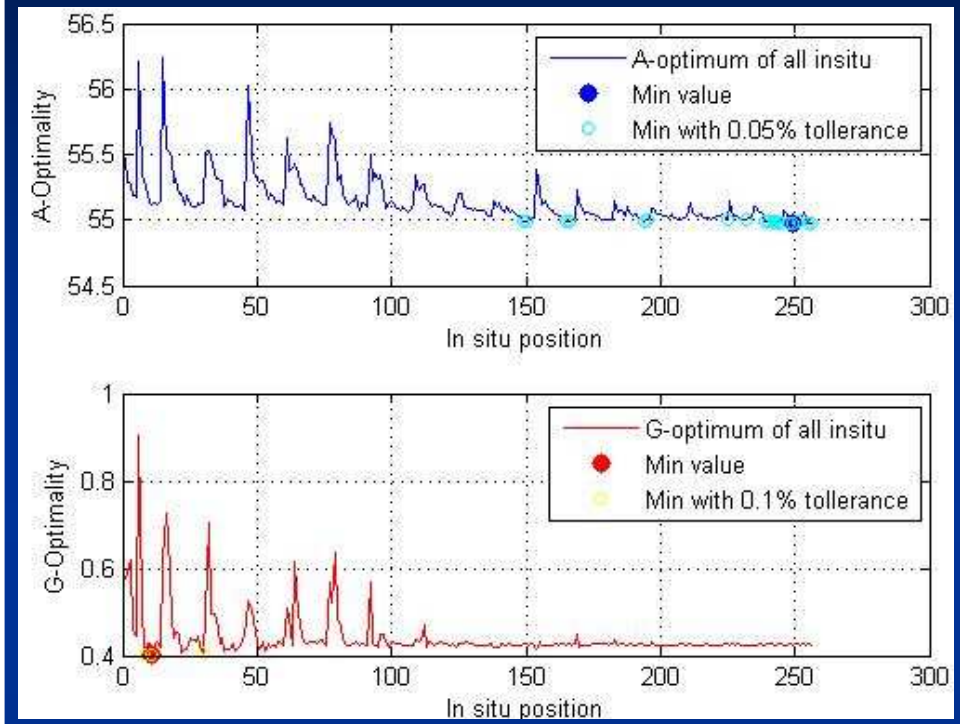


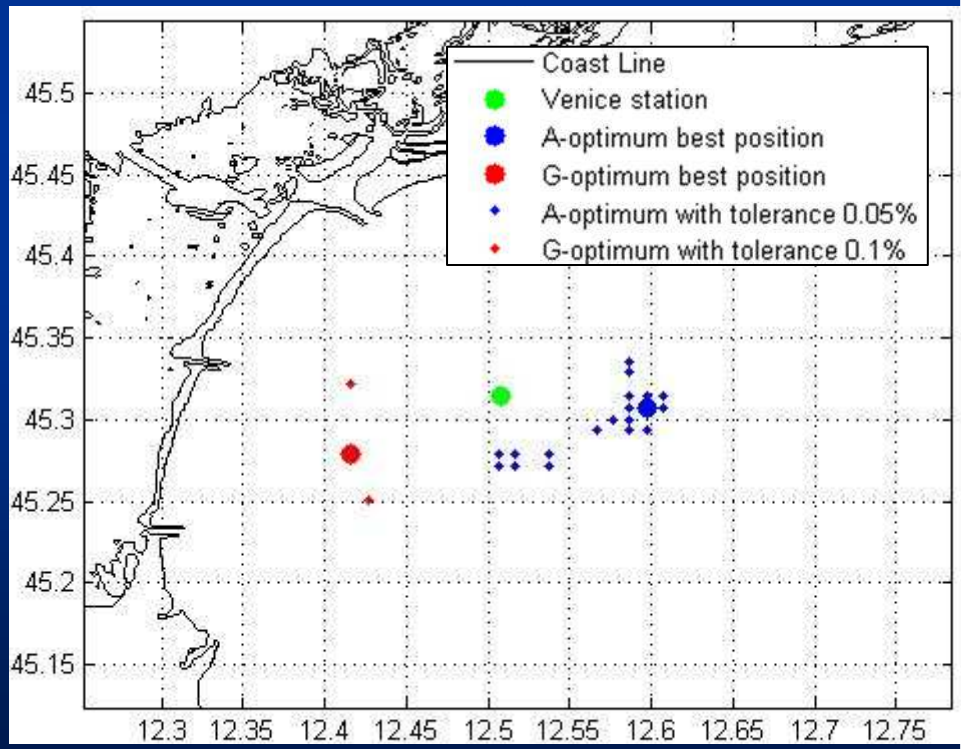
April (2008-2009) results



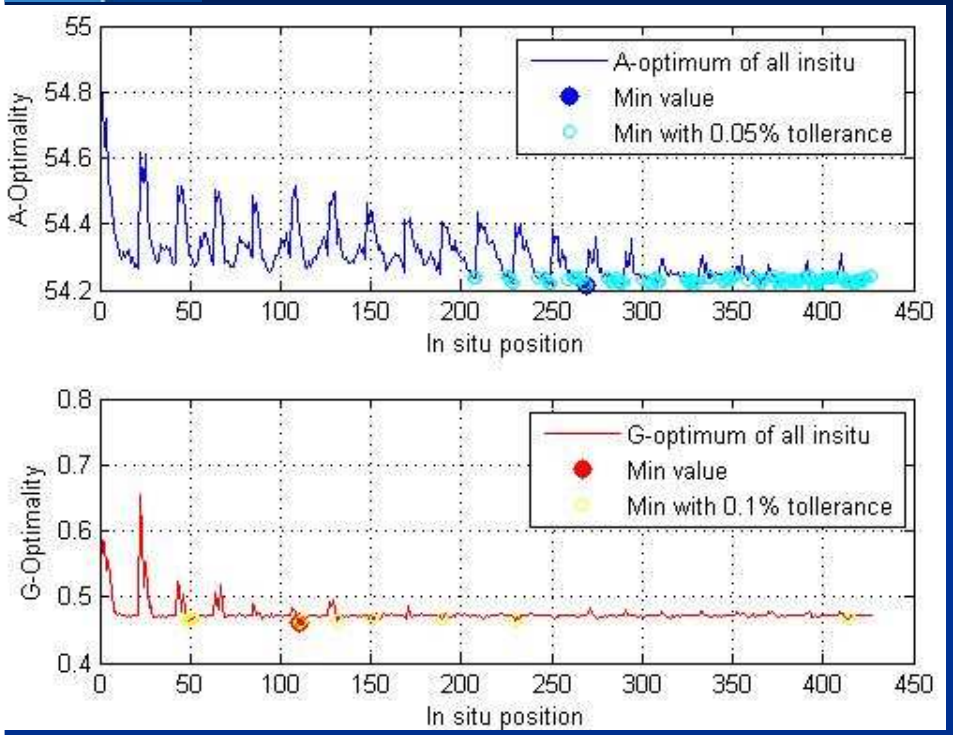


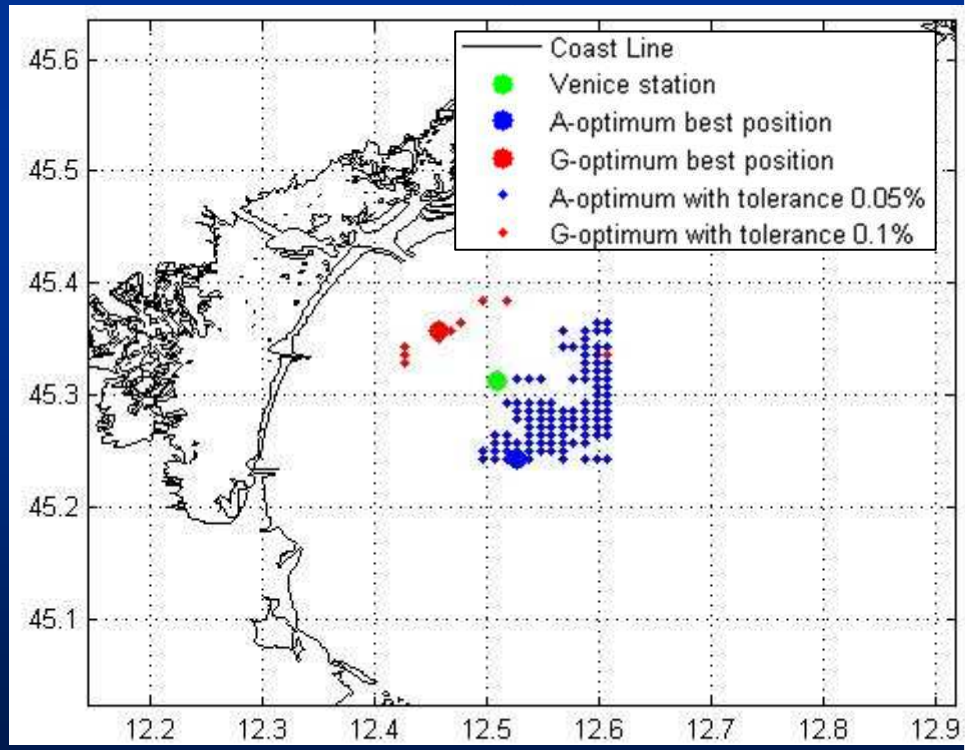
May (2005-2009) results



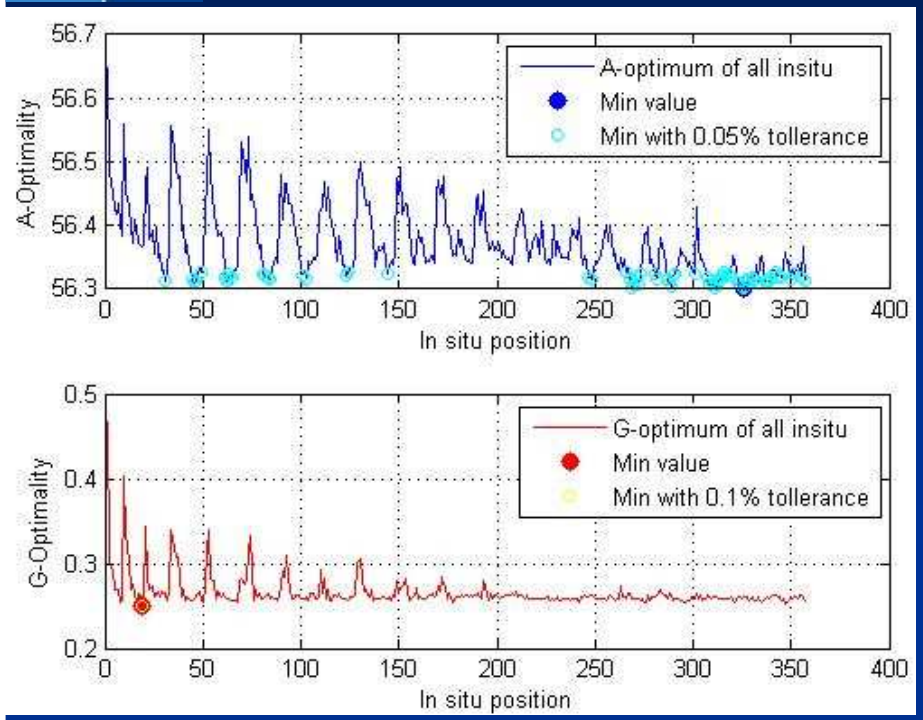


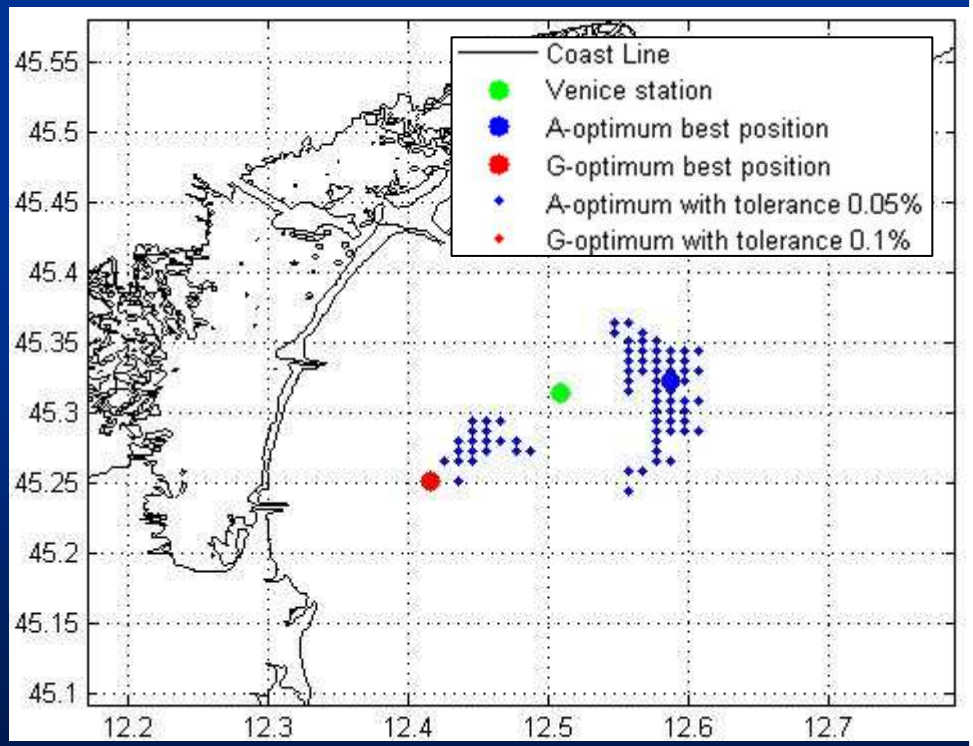
June (2008-2009) results



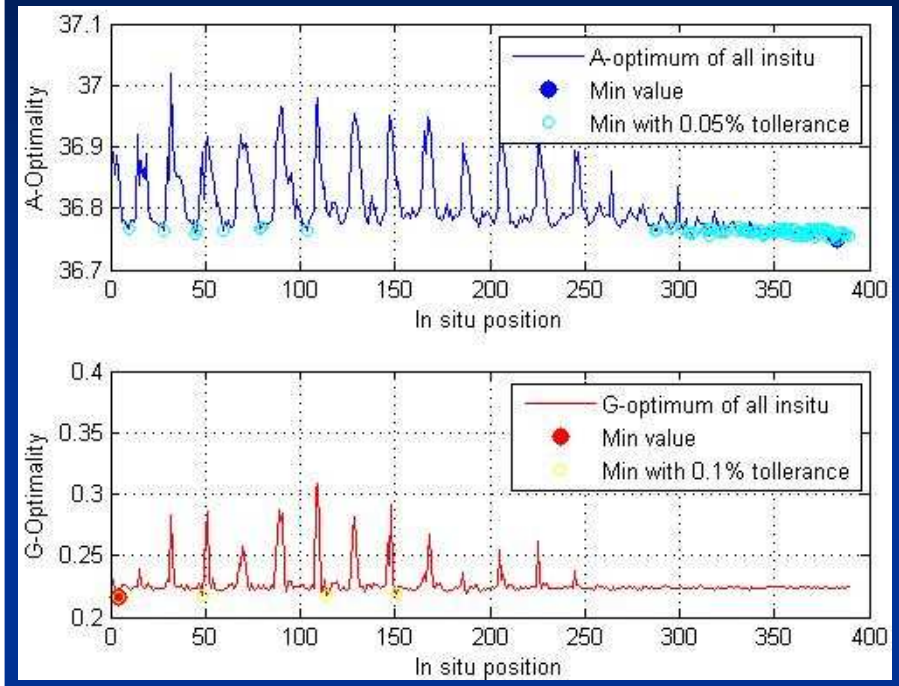


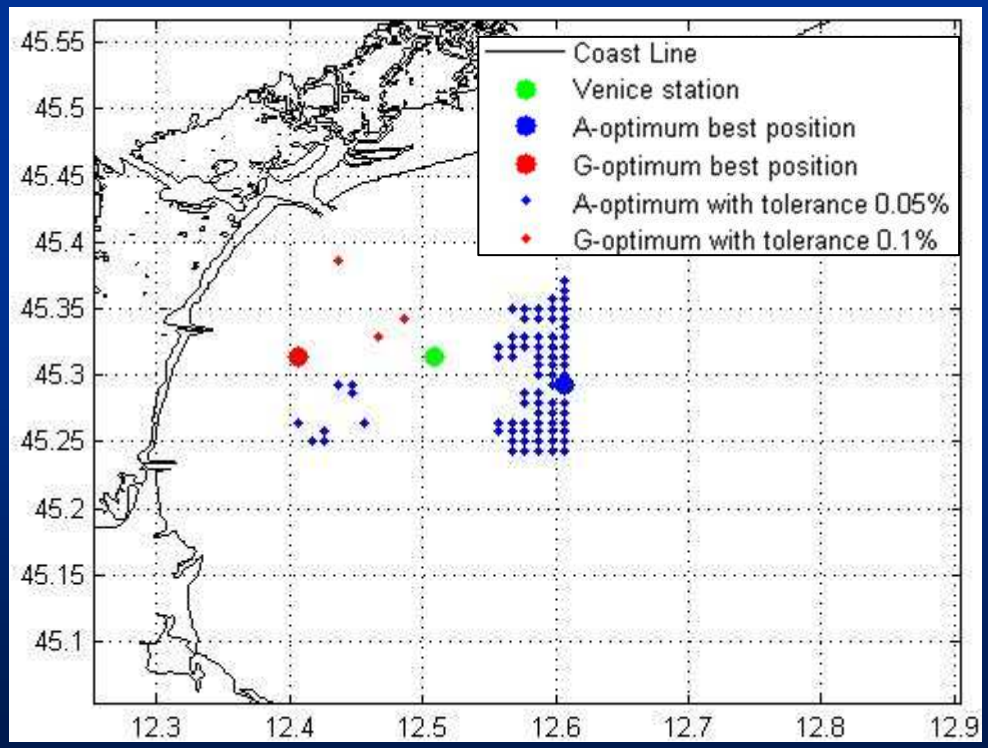
July (2008-2009) results



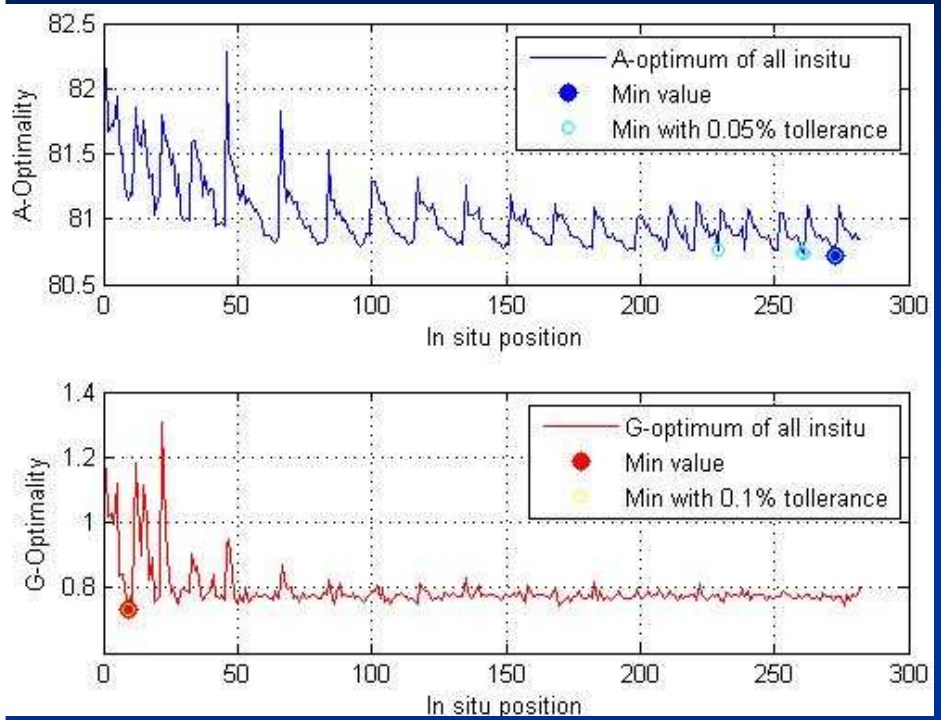


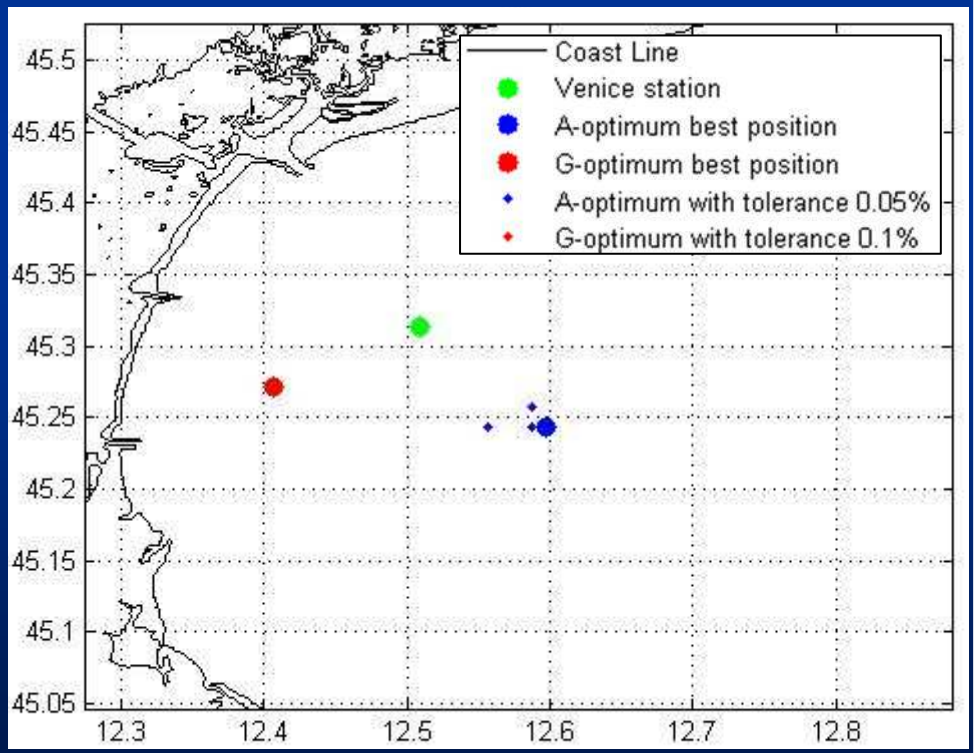
August (2008-2009) results



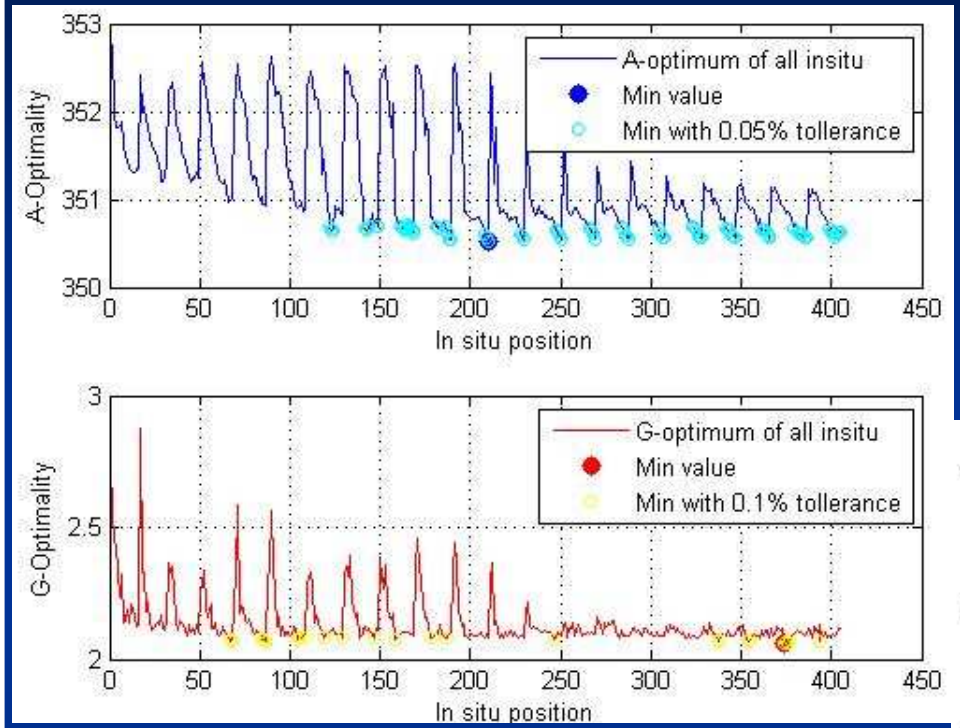


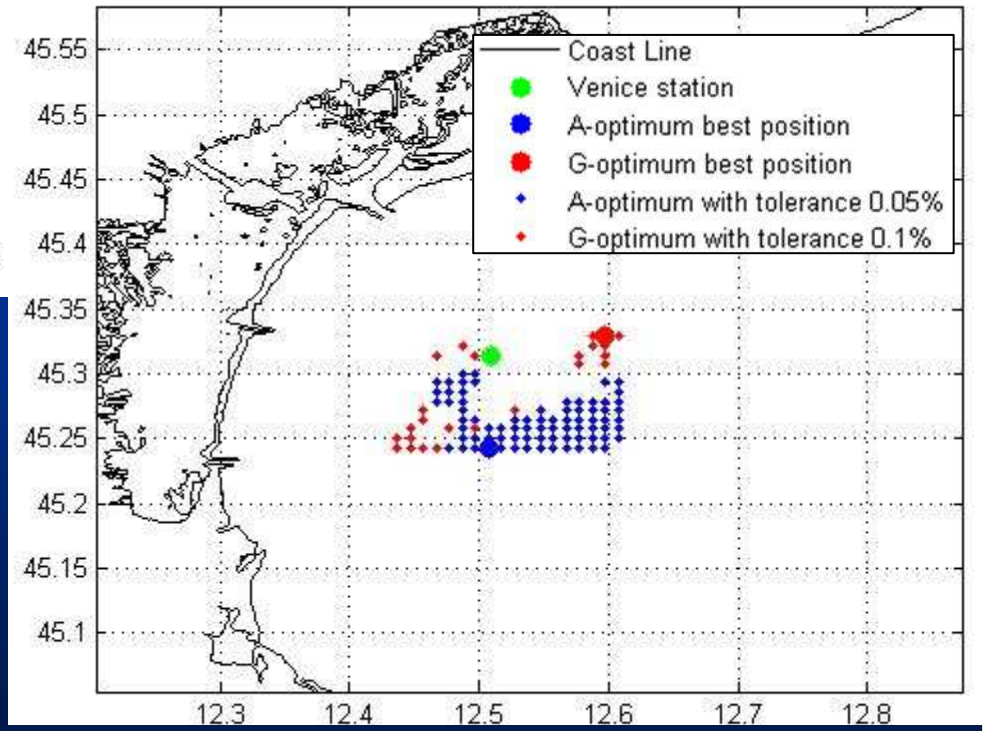
September (2005-2009) results



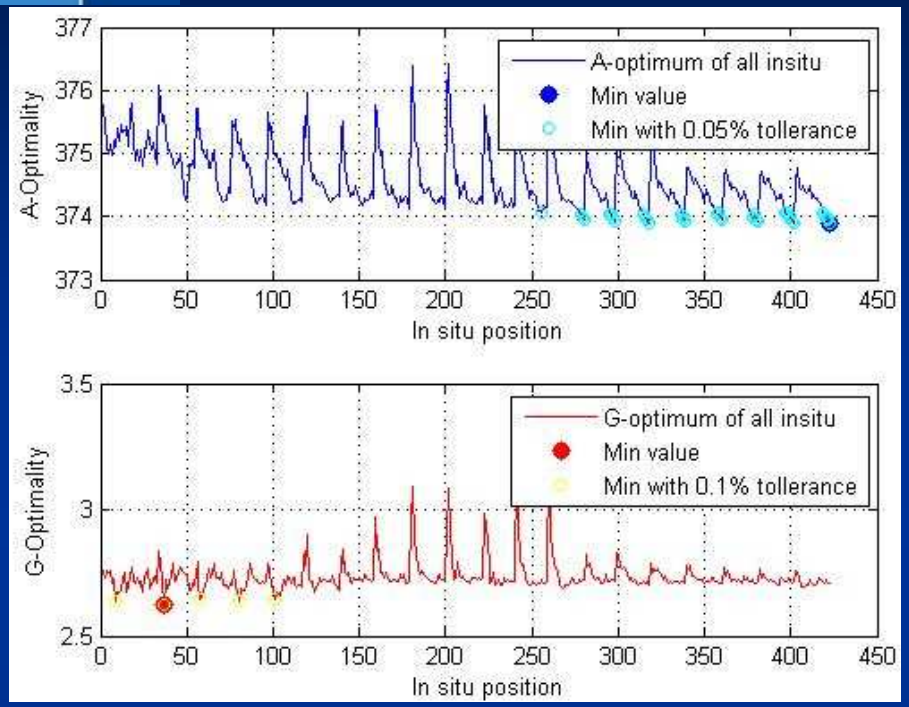


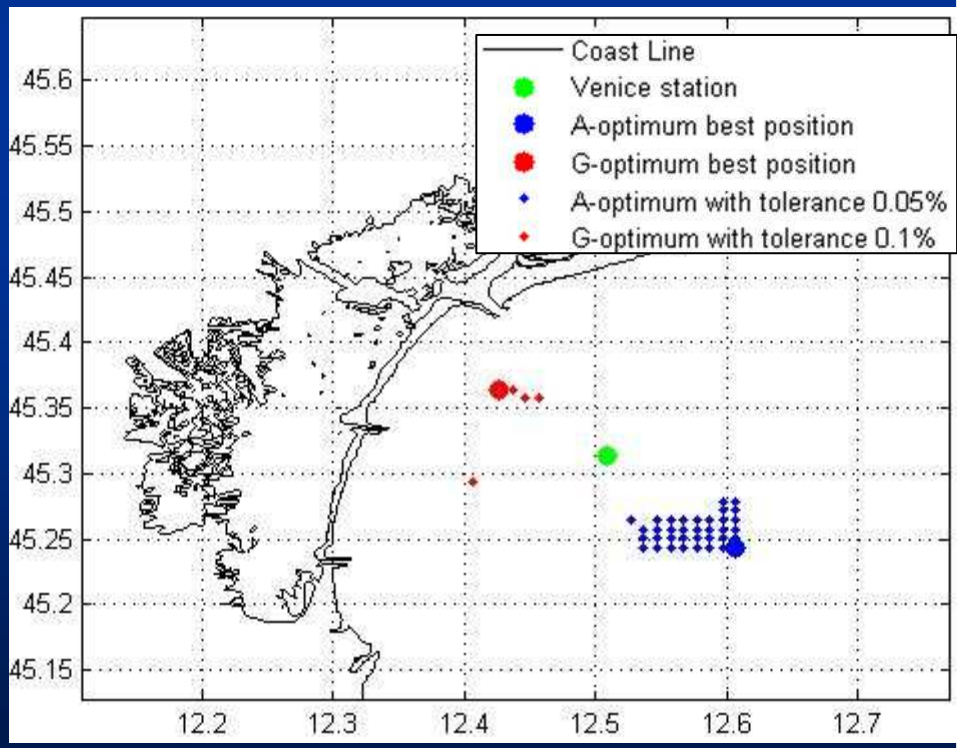
October (2008-2009) results



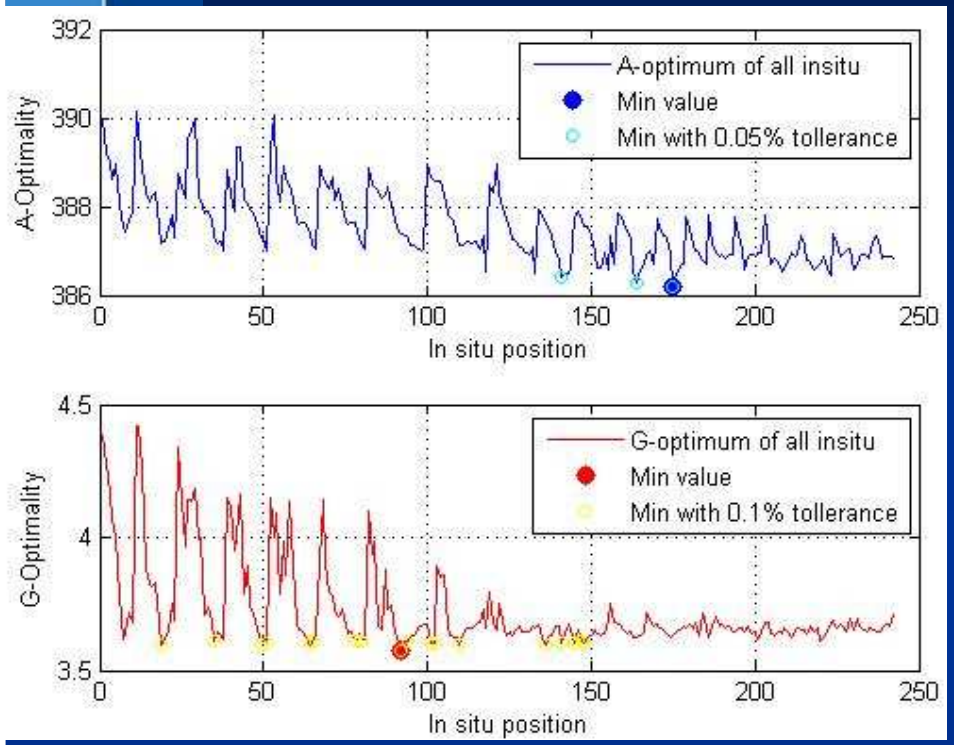


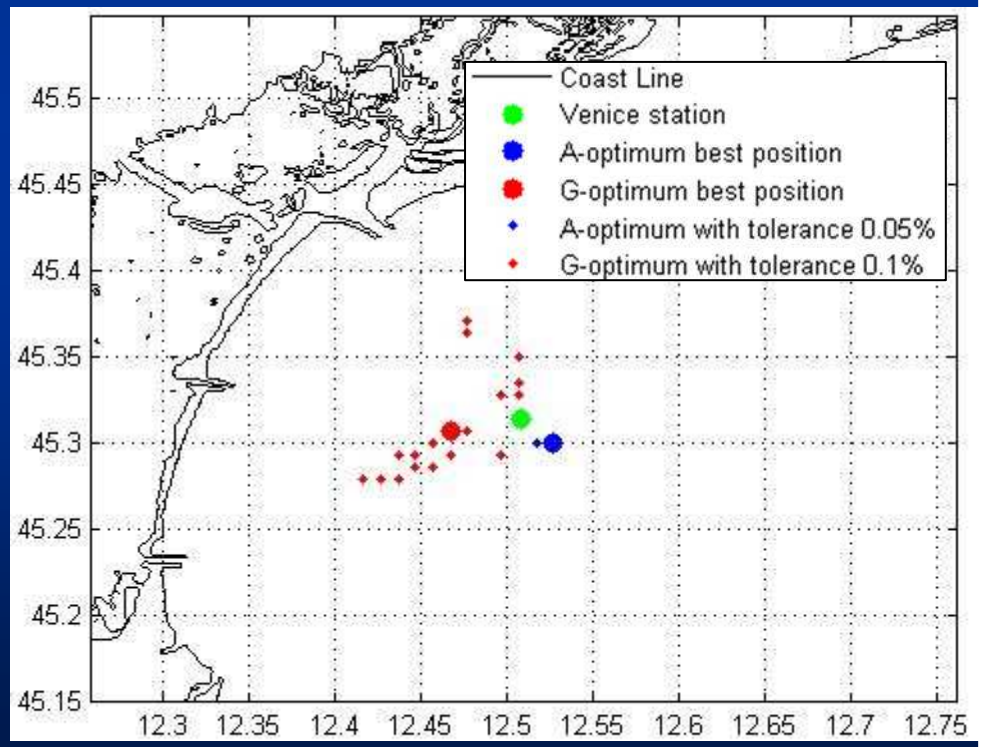
November (2008-2009) results





December (2008-2009) results



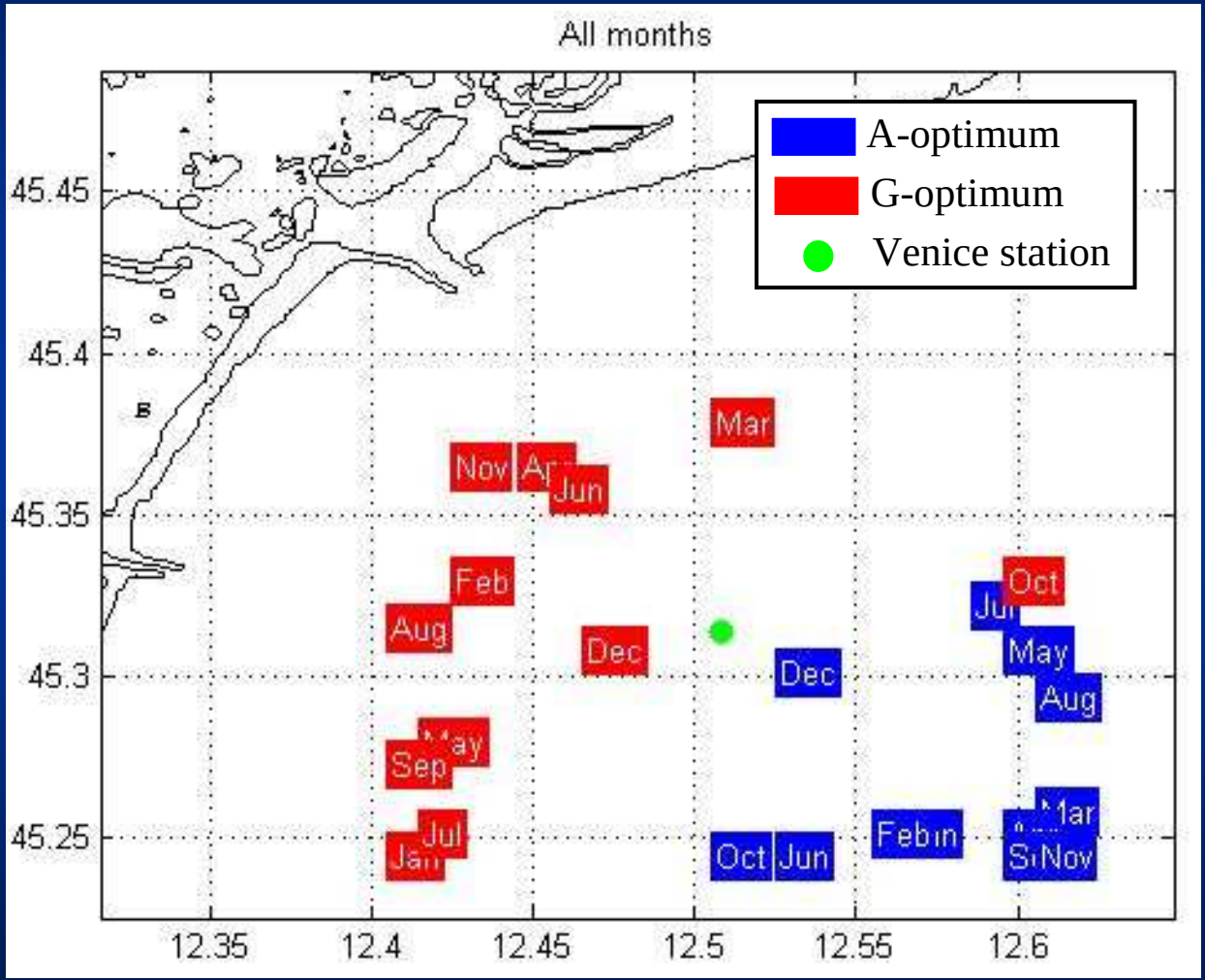


General results 1/2

(A and G optimal values for each month)

To find if there is a common place we have fuse the previous results.
The square boxes represent the best *in situ* placement month by month.

As showed each month is in different position and there isn't a common area.
However we observe that the two strategies applied have a common behavior (as presented in the next slide).



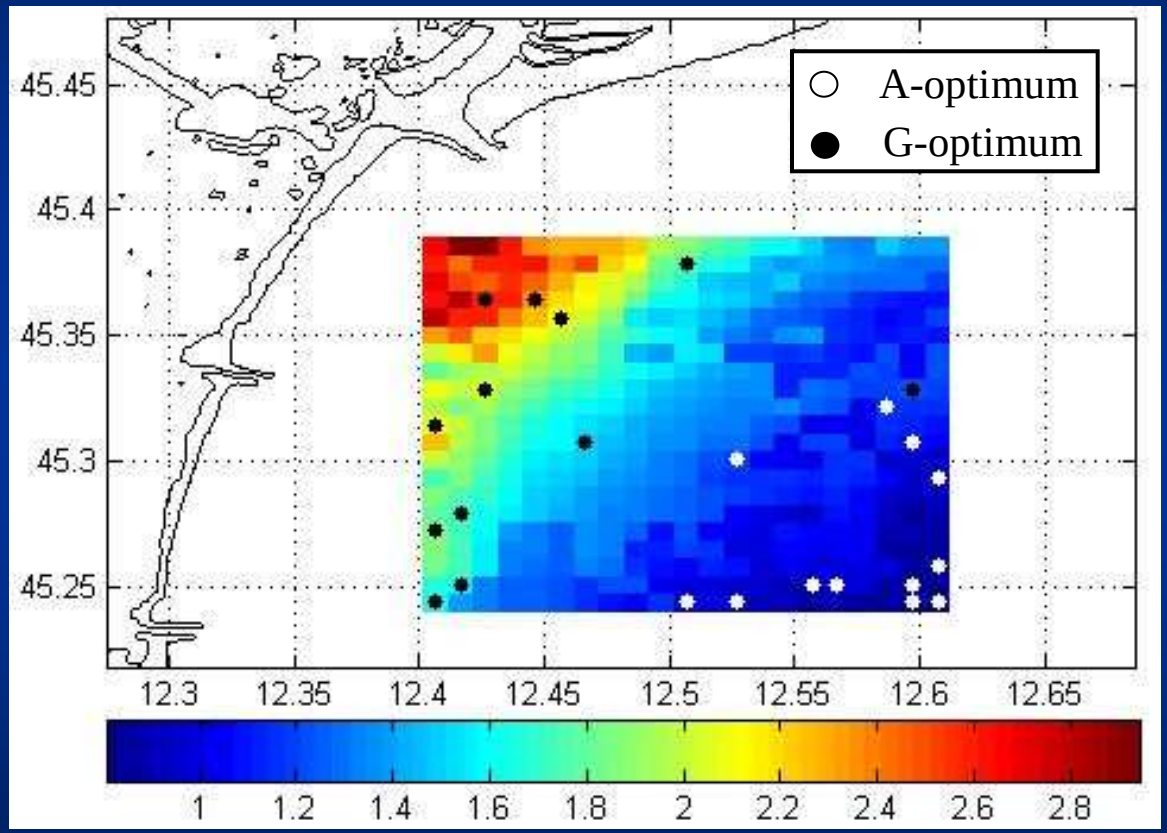
General results 2/2

(A and G behavior)

To explain the common behavior of the two strategies the previous results have been overlaid on the weighted average of the historical covariance maps.

We have defined:
A-optimal as the minimum of the trace of the variance
and G-optimal as the minimum of the maximum of the standard deviation.

A-optimal results will be on portion map where the uncertainly is bigger (higher variability)
while
G-optimal results will be on the lower variability area of that map.



As expected, the results depend on the optimizing sampling procedure and that have to be defined taking to account on the procedure to calibrate with the satellite.

Conclusion

The covariance matrix have been evaluated using the available monthly time-series MODIS acquisitions from 2005 to 2009. This technique produces 12 historical maps (COVARIANCE) that can be used to calibrate *in situ* data without satellite acquisitions but taking into account on the "satellite statistical behavior".

On the basis of those twelve-maps we have retrieved "the best" geographic position for the AERONET Venice station in terms of optimal design (A-optimum and G-optimum).

<u>It is important to note that the optimal design strategies defined depend on the satellite calibration procedure and have to be defined *a priori* taking into account on the considered optical sensor.</u>