Sanity checks of the tsunami elementary source data base

1. Introduction

In an earlier document from March 2017, we presented sanity checks on the Mediterranean set of elementary sources for a subset of coastal points. Here we present a similar set of checks for all basins (Mediterranean, Black Sea, Northeastern Atlantic) for the full final set of TSUMAPS-NEAM coastal points.

In order to perform a probabilistic tsunami hazard assessment, a large quantity of earthquake-tsunami scenarios have to be evaluated. TSUMAPS-NEAM achieves this goal by calculating tsunami propagation of Gaussian-shaped so-called elementary sources (ES) to the coastal points of interest (POI) and then making linear combinations of the mareograms at the POI according to earthquake source parameters. The ES are spaced at about 7 km with an initial wave height of 10 m and width parameter sigma=4000m (ca. 20km base width).

In this document, we analyze some key features of the mareograms (wave height time series with 30 seconds time step) in order to find possible inconsistencies. Some recommendations are given in the last section.

2. Table of contents

3. ES Description

There are 12 sets of ES: one for the Mediterranean (MED), one for the Black Sea (BLK), and 10 for the Northeastern Atlantic (NEA). The original netcdf files with about 10x larger set of POI than finally used take up several. From these files, slices containing the subset of POI used were taken and again saved as netcdf files. Since the so-called 'deterministic production chain' (taking earthquake scenario as input and delivering mareogram properties as output) works in a POI-oriented way, the ES Green's functions were also extracted and stored POI-wise (so called 'prefetched ES Green's functions). There are two files per POI (both binary), one containing the actual mareograms, and one the indices to the ES. An overview of the ES DB is given below.

```
Original netcdf DB:
@auriga:/nas/cat/TSUMAPS/
MEDITERRANEO/gaussians/med/   2.3 TB
BLACK_SEA/OUTPUT/              49 GB
ATLANTICO/OUTPUT/R*/           24 TB


Sliced netcdf DB:
MED: @auriga:/scratch/users/tonini/tsumaps/gaussians_db/
NEA and BLK: @auriga:/scratch/users/faromano/TSUMAPS_NEAM/multi_nc

Nr ES    nlon   nlat     tot   active    GB
NEA_R1    113     98   11074     8508    63
NEA_R2e    74     98    7252     7224    29
NEA_R2w    87     83    7221     7197    29
NEA_R3     70     60    4200     3734    15
NEA_R4    154     38    5852     5386    22
NEA_R5    164    113   18532    14534    58
NEA_R6    107     38    4066     4065    16
NEA_R7    143     75   10725     6657    27
NEA_R8    107     45    4815     4815    19
NEA_R9     34     38    1292     1187     5
MED       558    233  130014    51101   201
BLK       187     72   13464     8850     7
tot                           123258   491


Prefetched POI-oriented ES data base
@auriga:/scratch/projects/cat/SPTH/TSUMAPS/results/step2/prefetched
@auriga:/scratch/projects/cat/SPTH/TSUMAPS/results/step2_ext/prefetched

basin  #POI    GB
MED    1107   122
BLK     137     2
NEA    1092   250
tot    2336   374
```

Note: …*step2_ext* folder contains prefetched Green's functions for POI in NEA West of Gibraltar, which were beforehand falsely related to sources in MED, so NEA contains: 1071(step2)+23(step2_ext)=1094 POI of which 1092 were finally used.

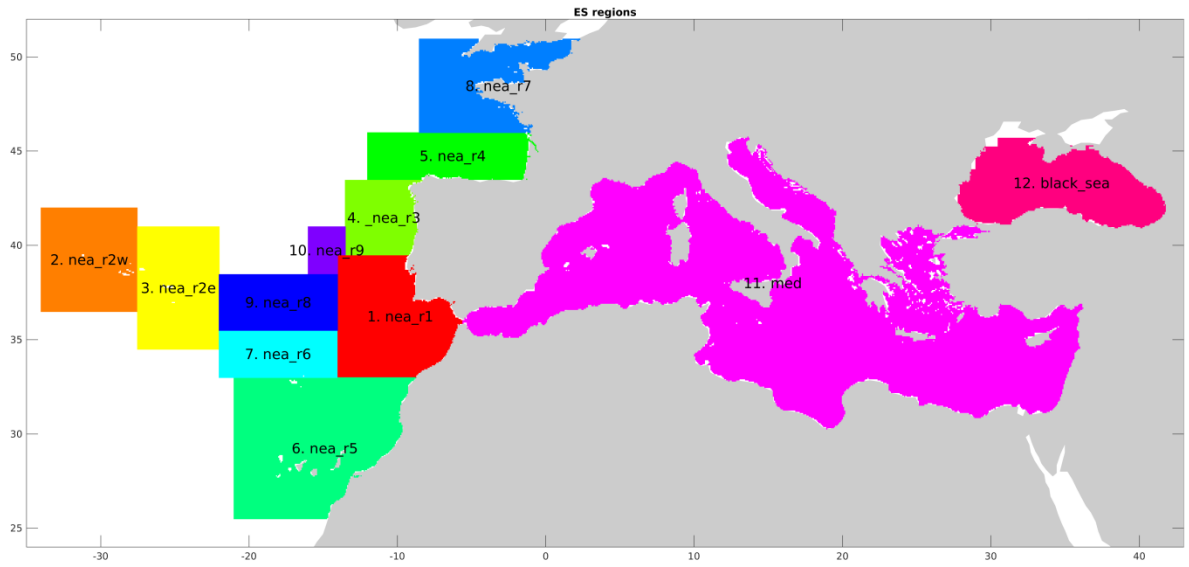

Figure 1: Elementary source (ES) data base (DB) regions.

4. Sanity / plausibility checks

All ES-POI mareograms were analyzed and data below were stored row-wise as ASCII files named 'check_prefetched_basxxxxx.stat', where bas is basin and xxxxx is the numerical ID of the POI.

The abbreviations and definitions are as follows:

iGS	ID of the ES (NEA:1..63307, MED:63308..114408, BLK:114409..123258)
sshMin	min. of mareogram in m
sshMax	max. of mareogram in m
t1st	time in s of first arrival at POI
nmins	number of local minima
nmaxs	number of local maxima
nzeros	number of zero crossings

Additionally, in the files: 'check_prefetched_basxxxxx.fail', it is reported where the numbers are outside the defined range.

Files are located on:
@auriga:/scratch/projects/cat/SPTH/TSUMAPS/results/step2/check_prefetched
@auriga:/scratch/projects/cat/SPTH/TSUMAPS/results/step2_ext/run

A more detailed description will be provided together with the code documentation.

5. Maximum wave height

Maximum wave height (sshMax) is the primary feature used for the calculation of PTHA. In real application, it is extracted from the superposition of weighted ES mareograms at a POI. Here, we perform no superposition but use the mareogram of one single ES directly.

Since it is hard to visualize ca. 120000x2300=276e6 values, we plot the maxima, either taken over all the POI for each ES or vice versa. This test could identify POI in bathymetric conditions which cause numerical instabilities resulting in unrealistically high values, a problem which is encountered for some codes.

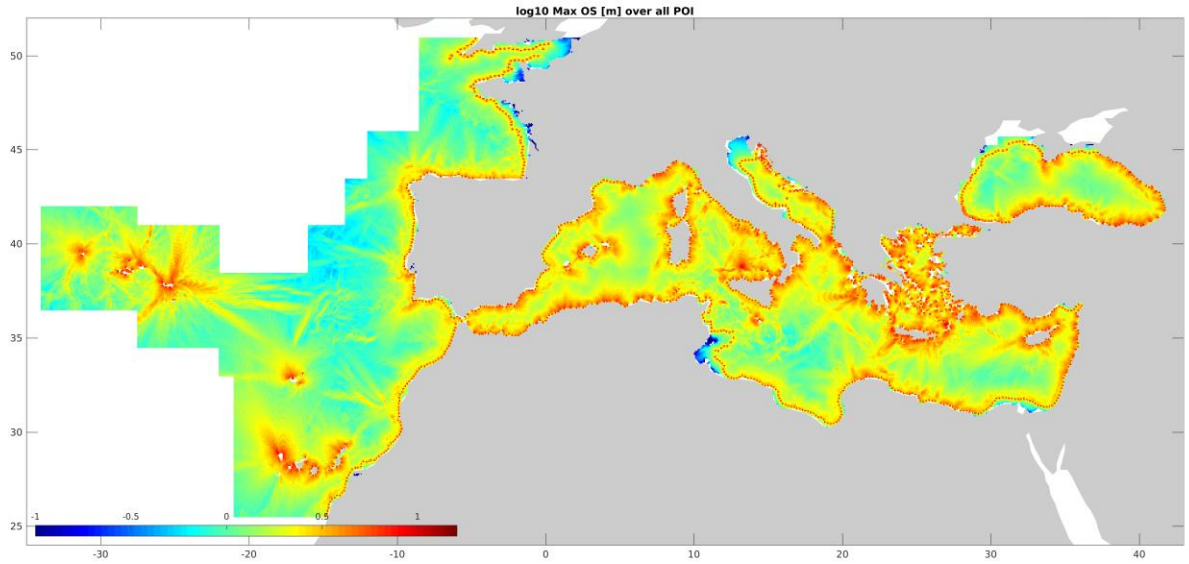

Figure 2: log10 of max of sshMax caused by ES over all POI.

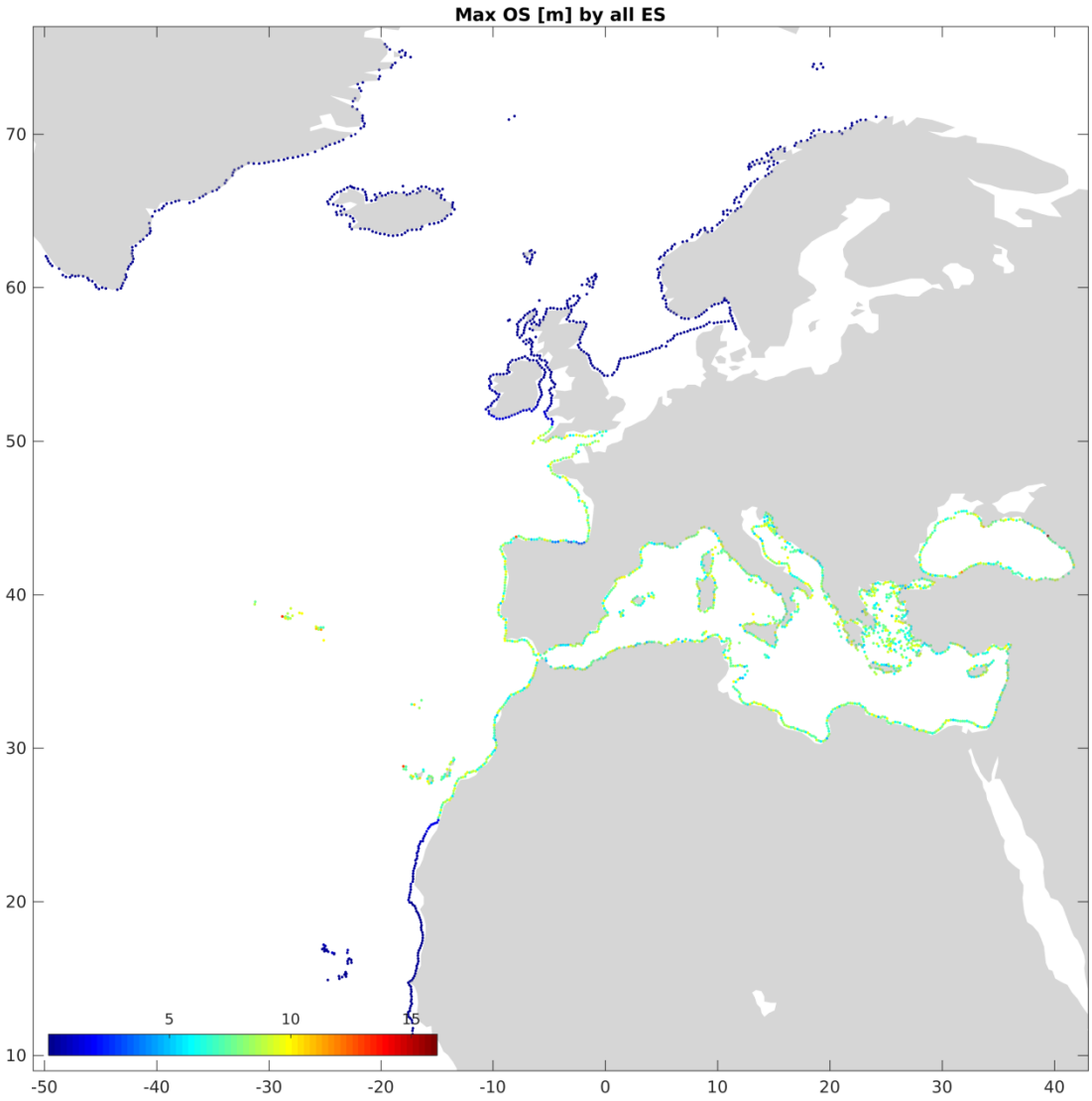

Figure 3: Max of sshMax at POI caused by all ES.

6. Cluster analysis with respect to sshMax

We perform a hierarchical cluster analysis with respect to sshMax using Matlab linkage and cluster functions. This should reveal potential indexing problems. The clustering is based on weighted average distance of correlation between rows or columns of the sshMax matrix.

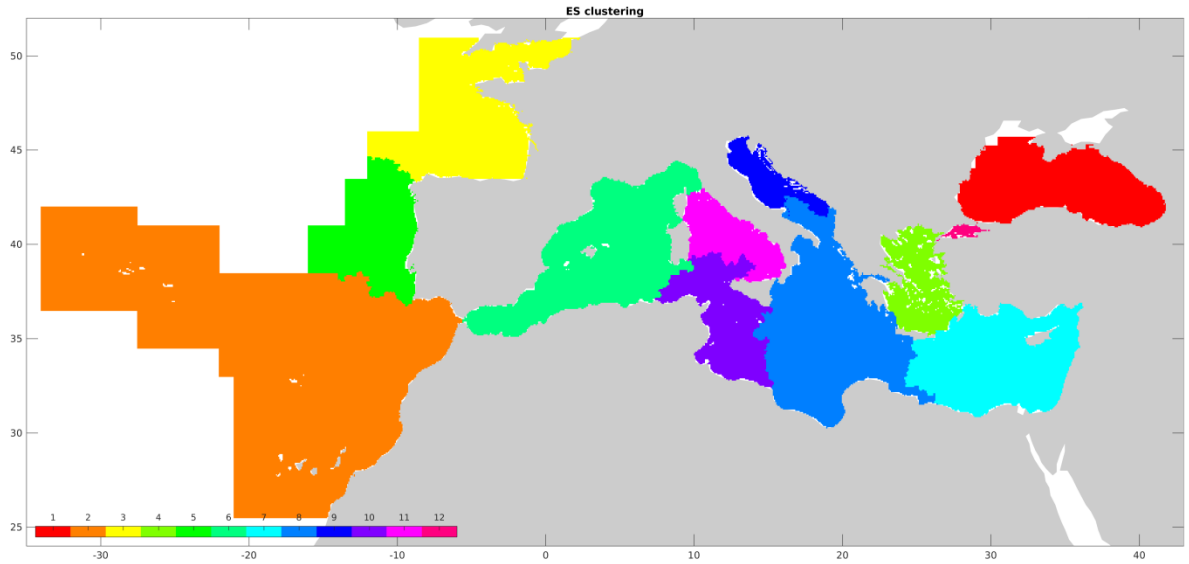

Figure 4: Clustering of ES (similar impact on POI).

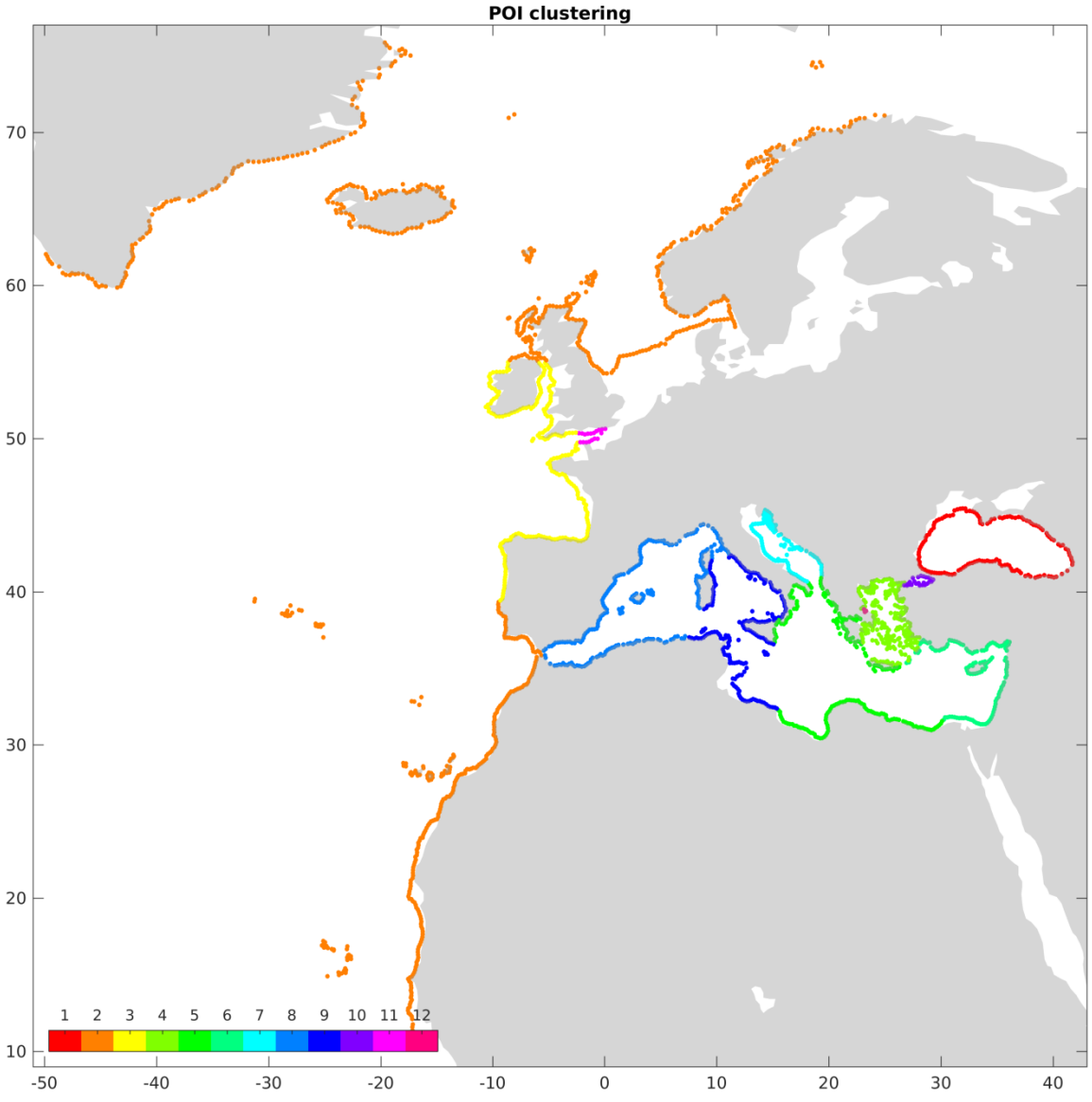

Figure 5: Clustering of POI (affected by similar ES).

7. Arrival times

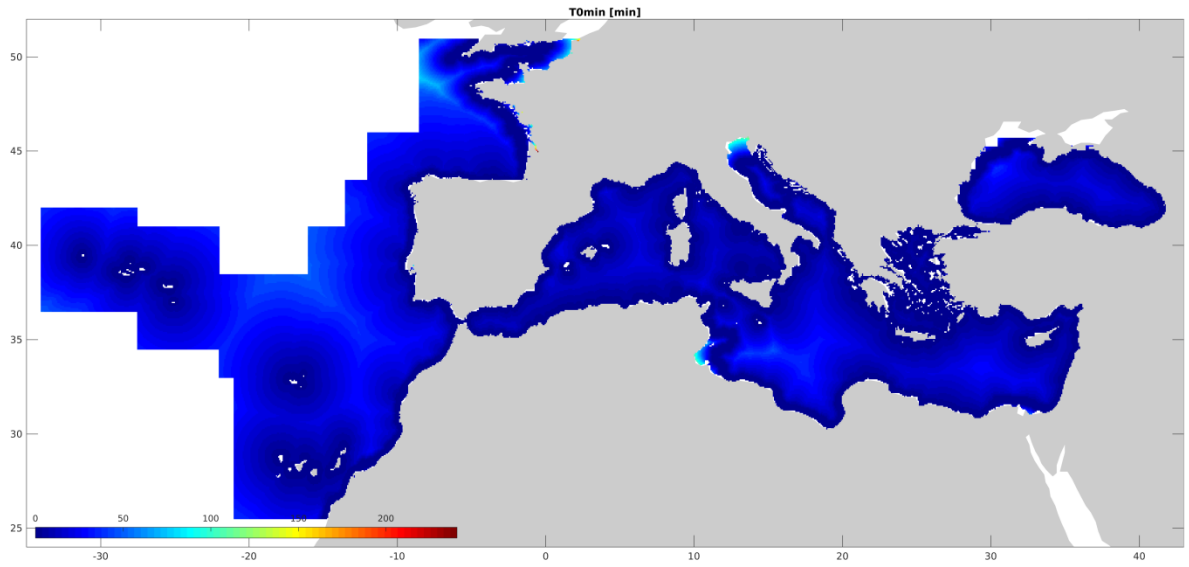

Figure 6: min(t1st) in minutes.

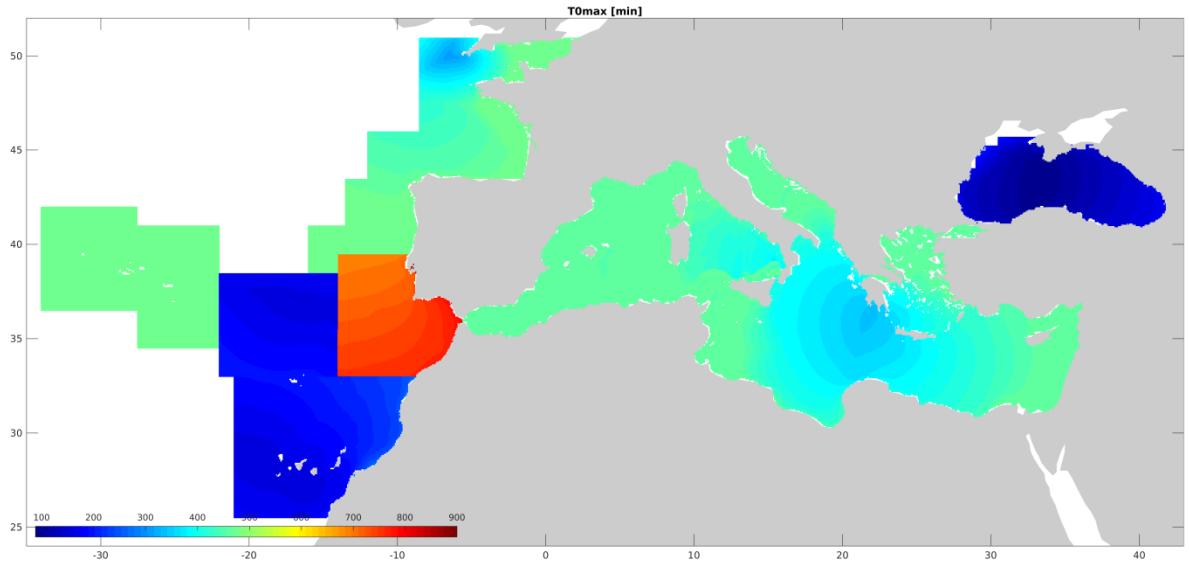

Figure 7: max(t1st) in minutes.

8. Analysis Error Codes

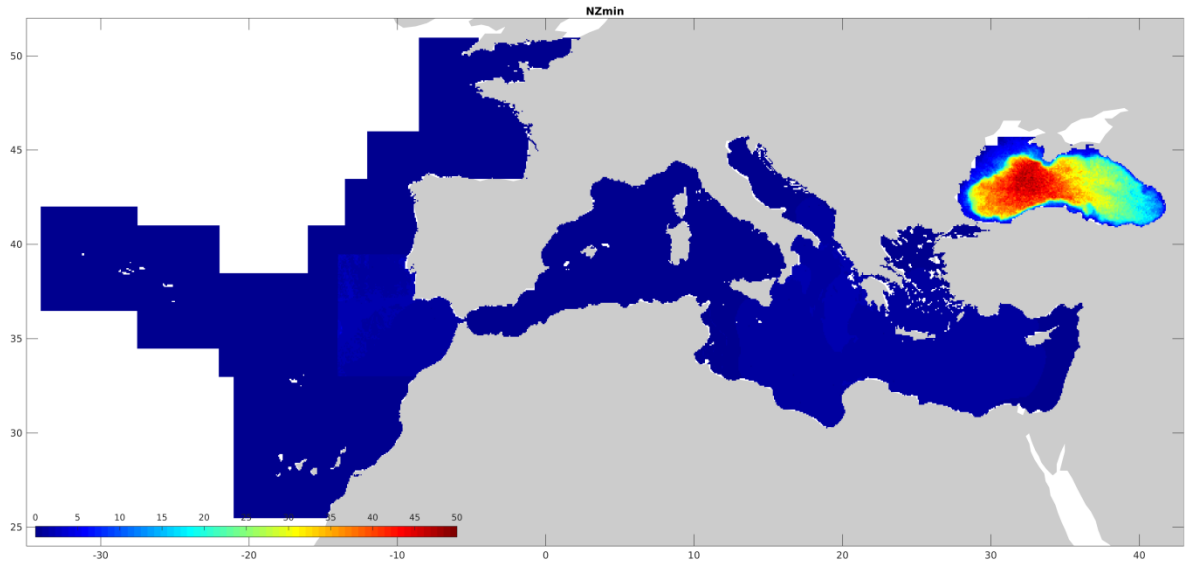

Figure 8: min(nmaxs).

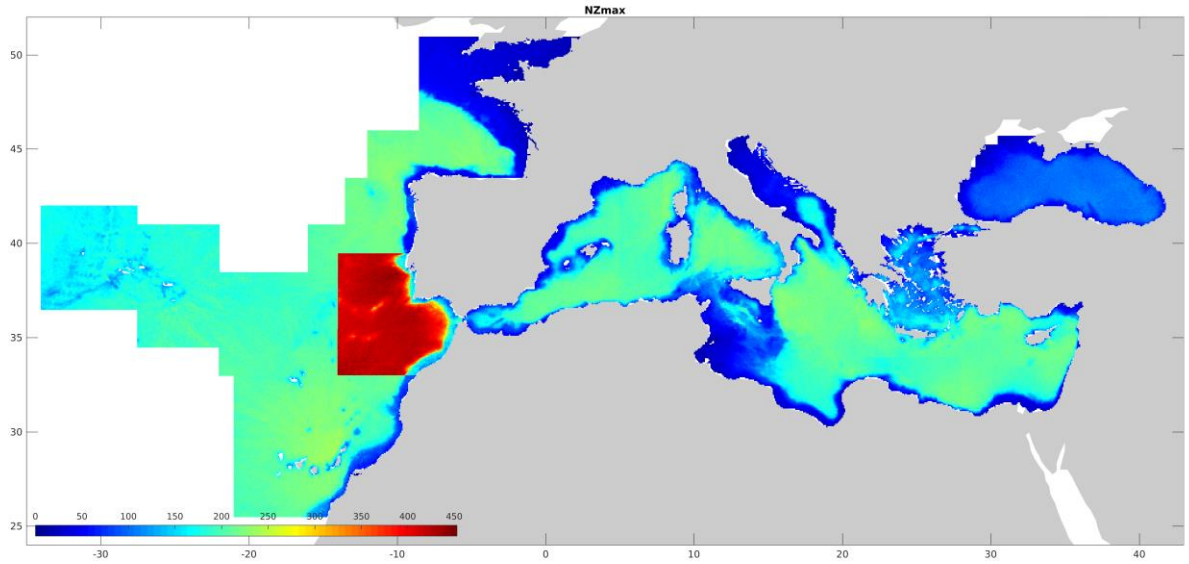

Figure 9: max(nmaxs).

9. Summary and recommendations

Maximum wave height

All ES-POI combinations behave as expected. No problems as in previous version apparent ('faulty Libya ES').

It is clear that the highest values at the POI are generated if an ES with its 10 m initial wave height is centered very close to the POI. This is sub-optimal, as in this case, the POI should be uplifted together with the initial condition, and wave height would be 0. It should be considered to either filter these cases out or to correct the mareograms for initial uplift or find a different solution as a recommendation for a future project. However, only few scenarios and POI should be affected.

Some of the lon/lat grids of the NEA sources are slightly shifted against each other. This is not optimal but could be compensated in the ES weighting scheme. As a recommendation for future

projects, a hexagonal grid should be considered instead of regular lon/lat spacing, which would be able to better cover large latitude extension.

Clustering

Both the ES and POI cluster in very ‘natural’ groups of bathymetric connectedness.

Arrival times

Minimum first arrival times are larger for ES distant of coast. Exception: few regions (e.g. Veneto, S Tunisia) where POI are far offshore due to shallow bathymetry. Ok.

Maximum first arrival times reflect simulation time of the ES set. ES at edges of a basin have generally large values. Ok.