Comments on "Sensitivity study on the main tidal constituents of the Gulf of Tonkin by using the frequency-domain tidal solver in T-UGOm."

I General comments

The objective of the paper – as presented in the introduction – is "to propose a robust and simple approach that allows to improve the tidal representation in the Gulf of Tonkin and to complement the previous studies in tidal modelling of the area." However, one of the objective is also to improve the Gulf of Tonkin configuration – bathymetry and bottom friction parameterization – using TUGO hydrodynamic code (that runs very fast), in order to use this optimized configuration for the coupled hydrodynamic/sediment model SYMPHONIE-MUSTANG (that runs more slowly). It could probably be mentionned at the beginning of the paper.

The paper presents many interesting results, particularly the sensitivity to bathymetry and bottom stress parameterization. We totally agree with the importance of the bathymetry and bottom stress in hydrodynamic modelling, and this paper is really welcome. However, the results are presented in a linear way, probably as the authors conducted the study and found them – this is particularly true for the section 2.1 Shorelines and bathymetry construction. It would be more valuable for the paper to present more analysed results and clearer conclusions. A discussion could include a reflexion around how significant is the choice of the bottom stress parameterization (e.g. drag coefficient constant Cd=cst instead of varying with z0 and H Cd=f(z0,H)).

Bottom stress sensitivity

The main point is that the choice of a parameterization rather than another is not justified, and finally, we do not really understand why the authors choosed to vary spatially z0 instead of Cd, as the choice of Cd constant all over the domain (SET 1) gives results very similar to Cd varying with z0 (SET2), i.e. cumulative errors are very close to each other (respectively 11.50 cm and 10.96 cm). Moreover, the fact of varying spatially the bottom friction (here z0) is not very convincing. Indeed, the choice of three areas (SET4) finally leads to results very close to one area (SET2) as two of the three areas as the same z0 than the SET2 optimized (z0 $=1.5\ 10^{-5}$ m) and without surprise, the cumulative errors are also very similar, even if slightly better with SET3 (10.43 cm instead of 10.96 cm for SET2). More disturbing is the fact that increasing the number of areas, the results are worse (SET5 with a cumulative error of 12.29 cm instead of 10.43 cm for SET4). This raises the question of the robustness of the method. The last point is that choosing for the mud area a linear expression instead of a quadratic one has no significative influence (SET 3). Finally, the sensitivity study allows to optimize bottom parameters (e.g. r, Cd or z0), but the sensitivity to a parameterization rather than another (linear, quadratic with Cd constant, quadratic with Cd=f(z0) with z0 constant, or even quadratic with Cd=f(z0) and z0 varying spatially) is not well demonstrated (e.g. similar cumulative errors). As a consequence, the reasons of choosing one parameterization rather than another is not clearly justified in the paper (except a cumulative error slightly lower). A discussion would be welcome to clarify this point: How significant is the choice of the parameterization between Cd=cst or Cd=f(z0) or … ? Why is it here not so significant ?

Bathymetry sensitivity

Results show a clear improvement between GEBCO and improved bathymetry. However, the use of bathymetry from nautical charts could reduce the depth (because charts are made for navigation purpose, see explanation in the specific comments) and lead to an overestimation of the tide. The improvement between the results and existing atlas FES2014b-hydro (without assimilation) is not so significative. For example, if we look at M2 (Figure 6), results show lower errors in the South but greater in the North, with significant differences between TKN model and FES2014bhydrodynamics. The overestimation of the amplitude nearshore could be partly due to the use of the bathymetry from the nautical charts. Note that the part of the FES2014b outside the Gulf of Tonkin could be masked, as it is not modelled in TKN. Otherwise, the scale of the figure is not appropriate, e.g. in Figure 7 for S2 the scale ranges from 0 to 0.33 m (note that there is no unit on the figures) whereas the maximum in the Gulf of Tonkin is only 0.19 m. Another point is that comparisons were made only between the model and altimetry, and comparison with **tides gauges** would be of great interest for the paper. For example, on Figure 6 authors could add dots at tide gauges locations colored with corresponding M2 amplitude. Finally, the TKN model has clearly errors greater than FES2014b-synthesis (with assimilation) which shows that the first objective (improve tidal model) is not really reached.

Maybe the paper would be clearer if the part 3.2 “Sensitivity to the bathymetry and assessment of tidal solution” could be separated into two parts, as here sensitivity to the bathymetry and assessment of tidal solutions are mixed, which is quite confusing. Results of the paper could be reorganized into three parts: 1) Sensitivity to bottom stress parameterization, that would lead to parameterization Cd=f(z0) with z0 varying spatially 2) Sensitivity to the bathymetry, that would lead to TKN bathymetry 3) Assesment of tidal solutions, that would show tidal improvement compared with Minh et al. (2014) and Chen et al. (2009), but not so clear improvement with FES2014b-hydrodynamics (e.g. Figure 6) and clearly, no improvement compared with FES2014b-synthesis (with data assimilation). This underlines the importance of data assimilation, and the need to go on developping **satellite missions** and **in-situ campaigns**, despite the great improvements of numerical models in the last decades.

In the following, we detail the specific comments.

II Specific comments

- line 45: the “strong improvement (compared to pre-existing tidal atlases)” of tidal solution is not so clear compared with FES2014b-hydro (without assimilation, see general comment for Figure 6). Moreover, the model errors are really greater than FES2014b-synthesis (not surprising, as this last one is with data assimilation).
- line 101 and following: add on Figure 1 geographic elements quoted in the text as for example Gulf of Tonkin, Hainan Strait, Hainan Island, Zhanjiang Peninsula, Qinjjiang, Nanliu, Yingzai rivers, Hai Phong harbour…
- line 153 and following: precise the expression of the tidal form factor, and the existence of four regimes.
- lines 153-162: maybe a map of F would be useful, as the variation of F is here described spatially.
- line 181: the objectives are not clear. It is mentionned to improve the tidal representation, whereas the TKN tidal model is finally not improved compared to FES2014b-synthesis. Probably introduce here as an objective the idea of using the tide as a response to calibrate the bottom friction parameterization. Objectives are also to improve model configuration (bathymetry and bottom friction) with TUGO (running fast) with the final goal to run the configuration on SYMPHONIE-MUSTANG (that runs more slowly).
- line 189: “in poorly sampled regions”, precise in terms of what, bathymetry? sea level? tidal currents? tide gauges?
- line 193: “in situ data and soundings are consequently rare and extremely valuable”, is it possible to list these data in the area? Particularly tide gauges may be of great interest for this study.
- line 230: why did you choose Bing as “the reference”? Is there is a paper reference for this?
- line 235: “OpenStreetMap shoreline is most of the time shifted”, is there is an explanation for that?

- line 236: “The GSHHG dataset suffers from the same problem but shifted by up to 500m eastwards.”, the shift is here very significant, is there is an explanation for that?
- line 238: “matching the reality”, what is considered as the “reality”, and why? Bing maps?
- line 240: shorelines from POCViP, is there is a reference paper for this software? What are the data behind this software?
- line 257: the GECBCO resolution is of approximately 1 km. There is an important difference between the grid resolution and the data resolution, as the data resolution could be lower than the grid resolution (and interpolated). Of which resolution are we talking here? Grid resolution ? Do we know what is the data resolution ?
- line 261: “digitalized nautical charts” Charts are made for navigation, and near the coast, the hydrographer generally chooses the lower soundings (shallow waters) for security purpose. As a consequence, the bathymetry from nautical charts in coastal areas gives shallower waters than “real” bathymetry. This should be mentionned, as it could partially explain the overestimation of tidal amplitude with TKN compared with FES2014b-hydrodynamics.
- line 285: “TONKIN_bathymetry dataset is not considered as the truth”, rather say that due to sampling, there are still uncertainties on the bathymetry. Also mention problems linked with nautical charts (shallower waters, see comment above).
- line 300 and following, no reference of TUGO for “storm surges simulations”?
- line 319: version of the code is a little bit complicated… “2616:78a276dd7882 of 2018-07-22
320 13:17 +0200”
- line 320: TKN is the name of the code or of the configuration? Not clear here.
- line 332: “The quadratic parameterization may be obsolete and a linear parameterization more adequate”. Is there is a justification or a reference for this sentence? The results will show the contrary.
- line 340: the “final goal”, i.e. hydrodynamic-sediment transport with SYMPHONIE-MUSTANG could be introduced earlier.
- line 377: the names FES2014b-hydrodynamics and FES2014b-synthesis are not really explicit. Choose for example FES2014b-without-assimilation and FES2014b-with-assimilation or something else, but more explicit.
- line 383: is there is a reference for FES2014b?
- line 402 and following: 2.3.1.1 Bottom stress parameterization, this section is not clear enough. The three parameterizations (2) (3) (4) are finally two paramaterizations (1a) and (1b), the first one with Cd=constant or Cd=f(z0,H) and the second one with r=constant. Particularly, the sentence 421 is not clear “In this study, we test three commonly used parameterizations: a constant drag coefficient Cd assuming a constant speed profile or a linear speed profile, and a drag coefficient Cd depending upon the roughness height z0”. This paragraph could be rewritten to be clearer.
- line 447: “In presence of fluid mud,...” repetition, yet said before, line 415
- line 457: (2) and (4) are linked with (1a) parameterization, whereas (3) is linked with (1b) parameterization. The way the parameterizations are presented is confusing.
- line 464: “two of the parameterizations described above: a quadratic bottom stress with a uniform drag coefficient Cd (Eqs. 1a and 2) and a logarithmic variation of Cd depending on a uniform bottom roughness height z0 (Eq. 4)” is not very clear. It would be more appropriate to talk about a quadratic bottom stress with a drag coefficient constant (Cd=cst) or varying with the roughness length (Cd=f(z0,H)).
- lines 521: “Spatially varying uniform friction parameters induce the best results on the tidal solutions rather than uniform parameters.” Is the improvement significative? Moreover, 12 areas give worse results than three areas, and three areas correspond finally to only two. Is the method robust enough ?
- line 522: “However, prescribing a linear parameterization in supposed fluid mud areas does not allow to significantly improve the solutions” How to explain this?

- line 530 and following “3.1.1 Sensitivity to the value of spatially uniform parameters”

To be clearer, this title could be “Sensitivity to a constant or varying Cd”, because SET1 corresponds to Cd=cst, and SET2 corresponds to Cd=f(z0,H). Optimisation conducts to Cd=0.9 10^{-3} m, and z0=1.5 10^{-5} m. A map of Cd for z0=1.5 10^{-5} m would help to see the differences in term of Cd between the two parameterizations. The cumulative errors between Cd=cst or varing with z0 are very similar (11.5 and 10.96 cm). Is there is a significative improvement with Cd=f(z0,H)? It is not justified why the authors choosed Cd=f(z0) rather than Cd=cst for the following. Indeed, SET4 could have been made with Cd varying spatially, instead of z0 varying spatially.

- line 585 and following: “Sensitivity to the value of spatially varying roughness length (SET3, SET4, SET5)”

This part could be separate in two parts “Sensitivity to a quadratic or linear stress” (which correspond to SET3 compared to SET2) and “Sensitivity to a roughness length varying spatially” (which corresponds to SET4 and SET5 compared to SET2).

SET3 has been conducted with a constant value of r and varying value of z0. Without surprise, optimization leads to z0=1.5 10^{-5} m, which is the optimized value of SET2. Why not fixing the z0 value (z0=1.5 10^{-5} m) and make vary the r value? This could lead to an optimized r value, probably different from 1.18 10^{-4} m from Le Bars et al. (2010), and perhaps results would show more sensitivity to a quadratic or linear stress (depending on the r value).

SET4: the three areas are finally only two areas (Regions 1 and 2 with the same z0), and the z0 in areas 1 and 2 is the one corresponding to SET2 optimized (z0=1.5 10^{-5} m). Finally, SET4 is not so different from SET2, with very similar cumulative error (10.43 cm instead of 10.96 cm). Is the improvement significative?

SET5: how to explain that it is worse with 12 areas? Why don’t we converge also to z0=1.5 10^{-5} m? It could be interesting to include these results and analyse them, otherwise, we don’t have enough elements to understand, and we can wonder if the method is robust enough.

- line 622: “3.2 Sensitivity to the bathymetry and assessment of tidal solution”, this section could be split in two parts, see comment in general comments.

- line 649: it is not so clear why TKN and TKN-gebco show bigger errors than FES2014b-hydrodynamics. We understand than K1 is less sensitive to bathymetric variations, but this is probably not the only explanation.

- line 665: TKN is better than Minh et al. (2014) and Chen et al. (2009). Why? This could be explained.

- line 686 and following: the acronym SLA is not detailed, are we talking about Sea Level Anomaly? This term is generally corrected from tide. It is not clear here.

- line 756 and following: the fact that TUGO is used to prepare a best configuration for SYMPHONIE-MUSTANG could appear earlier in the paper (e.g. in the objectives). Otherwise, it is not clear if the objective of the paper is to improve the tidal model or improve the configuration (bathymetry) and parameterization (bottom stress) for SYMPHONIE-MUSTANG.

- line 763: it is clear that the new bathymetry improves the results compared with GEBCO, but it is not clear if the final configuration TKN is improved compared with FES2014b.

- line 784: “the use of a constant Cd parameterization or the use of a Cd depending on the roughness length led to **fairly similar results**” We totally agree with this conclusion, that could appear earlier in the paper (in the results). As a consequence, why choose a Cd depending on the roughness length instead of a constant Cd for SET3/SET4/SET5?

- line 791: “the regionalisation of the roughness length into three regions, for addressing the issue of representing the complexity of seabed composition and morphology, **moderately improved the accuracy of our simulation**, with a lowest cumulative error for all four waves of 10.43 cm”, instead of 10.96 cm. We totally agree with this conclusion, is the improvement significative enough?

- line 794: “Finer local adjustments of the roughness length or the choice of a linear velocity profile in the area of fine mud, **did not improve** the accuracy of our simulations.” We totally agree with this conclusion, how to explain that there is no improvement ?
Finally all the SETs - once optimized individually - are very close to each other, in term of cumulative error.
- line 800 : “results therefore quantitatively showed the importance of the bathymetry and shoreline dataset and of the choice of bottom friction parameters for the representation of tidal simulations over a shallow area like the GoT”. This could be clarified. The results show that the choice of the bottom stress parameterization is not so important (e.g. SET2 optimized with Cd=cst gives similar results than SET3 optimized with Cd=f(z0,H), in terms of cumulative error), but the value of the bottom parameter (e.g. Cd=0.9 10-3 m for SET1 or z0=1.5 10-5m for SET2) is important, as it impacts clearly the cumulative error (e.g. Figure 8).
- line 806: “Our resulting configuration brought a clear improvement in the tidal solutions compared to previous 3D simulations from the literature and to the tidal atlas FES2014b (without data assimilation) for the semi-diurnal waves.” The improvement compared to FES2014b is not so clear if we look at Figure 6 for example. The addition of tide gauges data should greatly help to qualify the results.
- line 813: “Using bathymetry data available from digitalized navigation charts was a relatively simple way (compared to performing additional in-situ measurements) to significantly improve the representation of topography in the coastal and estuarine areas of the GoT”. We agree, however, as mentionned above, bathymetry could be underestimated (shallower waters) because charts are made for navigation and the shorter soundings are choosen for security reasons. The use of nautical charts could then lead to an overestimation of the tidal amplitude in some coastal areas.
- Table 2, FES2014b-hydrodynamics and FES2014b-synthesis are clearly missing.