National Oceanography Centre
NATURAL ENVIRONMENT RESEARCH COUNCIL

National Oceanography Centre

Research & Consultancy Report No. 62

A reassessment of the UK operational surge forecasting procedure

Joanne Williams[1], Andy Saulter[2], Clare O'Neill[2], Jenny Brown[1] & Kevin Horsburgh[1]

2018

[1]National Oceanography Centre, Liverpool
6 Brownlow Street
Liverpool
L3 5DA
UK

[2]Meteorological Office
FitzRoy Road
Exeter
EX1 3PB
UK

Tel: +44 (0)151 795 4856
Email: joll@noc.ac.uk

DOCUMENT DATA SHEET

AUTHOR
WILLIAMS, JOANNE, SAULTER, ANDY, O'NEILL, CLARE, BROWN, JENNY & HORSBURGH, KEVIN

PUBLICATION DATE 2018

TITLE
A reassessment of the UK operational surge forecasting procedure.

REFERENCE
Southampton, UK: National Oceanography Centre, 28pp.
(National Oceanography Centre Research and Consultancy Report, No. 62)

ABSTRACT

This report is a summary of the Met Office surge forecasting procedure for the UK, and some investigations into possible sources of error. The forecast is based on the "non-tidal residual", the difference of two model runs with and without weather effects, linearly added to the "astronomical prediction" from local tide gauge harmonics. This method is exposed to several errors. Here we do not attempt to quantify errors in the model or weather forcing, but we show how errors can arise in the harmonic analysis and due to the double counting of weather-related tides. The executive summary, validation guidelines and recommendations have been prepared jointly with the Met Office.

KEYWORDS:

ISSUING ORGANISATION **National Oceanography Centre**
University of Southampton Waterfront Campus
European Way
Southampton SO14 3ZH
UK

Not generally distributed - please refer to author

(This page intentionally left blank)

Contents

1 Executive Summary

Presently, the UK system for forecasting coastal (still) water levels is based on a combination of a harmonic tidal prediction and a model derived forecast of the meteorologically induced storm surge component.

In general forecasters will attribute errors in these forecasts to the modelled surge component. However, the harmonic predictions can also contribute significant errors [e.g. Flowerdew et al., 2010]. With the introduction of the NEMO-based surge model at the Met Office [O'Neill and Saulter, 2017], and in order to enable an informed plan for further model development, the National Oceanography Centre (NOC) have undertaken a review of harmonic contributions to errors.

Observed total (still) water levels include contributions from the following components:

1 Astronomic tide
2 Meteorologically induced surge
3 Steric (density) based variations in sea surface height
4 Long term variations in mean sea level, which may include contributions from sea level rise and land movements
5 Wave setup

In the present forecasting system it is assumed that component 5 makes only a limited contribution and/or is accounted for elsewhere in the prediction framework; that components 1, 3 and 4 are fully captured by harmonic analysis of the observed water level record; and that the surge (component 2) is forecast by the model and can be linearly summed with the harmonic predictions. Implicit in this method is an assumption that the "surge residual" (model total water level minus model tide) is an adequate proxy for the "non-tidal residual" (observed total water level minus harmonic prediction; see Pugh and Woodworth [2014]), which is the quantity that actually needs to be forecast. However this and previous reports find the following effects that can cause differentials between surge and non-tidal residuals:

a Annual mean sea level can vary sufficiently that the harmonic predictions mean sea level (averaged over 19 years) may be significantly offset from that underlying present day observations (e.g. Byrne et al., 2017; see also `http://www.ntslf.org/products/sea-level-trend-charts`)

b The harmonic prediction includes low frequency (annual and semi-annual) components that are dominated by a combination of steric effects, and the effect of weather on the sea surface (averaged over the harmonic analysis period). Since the surge model forecasts the weather effects specific to the time when the forecast is issued, some double counting occurs. Whilst not affecting model verification statistics specifically, these effects could introduce a seasonal variation in errors which may influence how the surge model is tuned in order to successfully estimate the non-tidal residuals observed during a model calibration period (comparing Figures1 and 8). The issue here is that events occurring at different times of the year may not be well represented if the surge model has been trained on a specific period, season, or set of events.

c The effect of storm surges and other effects of atmospheric forcing also influence the estimate of phase in many tidal harmonics in the observational record (particularly S2 around the coast of the UK). This means that assumption allowing a linear addition of the model surge to harmonic tide is not correct. Whereas the low frequency components can be safely removed from any analysis, these effects cannot, which is likely to introduce some level of error into forecasts of non-tidal residuals, particularly when comparing residuals at high frequencies (e.g. 15 minute or hourly residual data, as opposed to skew surge).

d Reconstruction of time-series from harmonic predictions with and without the effect of surge (Figures 3-5, 11 and 12) demonstrates that, in addition to the above, the harmonic predictions contain a level of error. These errors can be particularly high due to phase differences in the original and reconstructed time-series in shallow-water estuarine locations with high tidal amplitudes (Bristol Channel, Liverpool Bay, Humber, Thames Approaches).

From these results we can identify two problems that need to be addressed in our present evaluation of surge model errors and subsequent use of this information in operational forecasting. First, surge errors potentially include systematic biases and long term variations in water level that should not be expected to be forecast by the surge model. These errors can be removed or accounted for in both verification and forecast post processing (e.g. via astronomical tide re-engineering corrections). Second, there are a more complex set of errors contained in harmonic predictions of water level that are not easily decomposed from the dominant tidal components. These errors may particularly affect forecasts of non-tidal residuals (we speculate a reduced sensitivity for skew surge as this has no tidal phase dependency). The best way to mitigate these errors will be to better quantify and attribute them within the verification and forecasting process.

2 The current procedure

The current procedure for forecasting total water level is as follows:

a Run the CS3X and/or NEMOSurge model in surge and tide mode, forced by an ensemble of wind and pressure from the current weather forecast M_s.
b Run the CS3X and/or NEMOSurge model in tide-only mode M_t
c Get the residual from these models $M_r = M_s - M_t$
d Find the amplitude and phase of harmonics at an individual gauge location from past tide records.
e Derive a tide prediction from the gauge harmonics H_g.
f Forecast the total water level W_f as model residual plus gauge harmonics $W_f = M_r + H_g$.
g Compare the forecast to the observed level W_g at the gauge over the previous few tidal cycles.
h Apply various "empirical corrections" to nudge the forecast based on the mismatch of the peak tide over the last few days. [Hibbert et al., 2015]

2.1 Possible errors

There are several possible sources of error in this procedure.

a i The model may not capture all the necessary physics, or be at high enough resolution. A full discussion of the model accuracy is beyond the scope of this project - and it has been discussed at length by model developers. We note only the similarity of CS3X and NEMO-Surge. Baroclinic and wave effects are not captured by the model.
  ii The weather forecast or ensemble selection may not be correct. This is beyond the scope of this project.
b i See a(i).
  ii It is assumed that the model in tide-only mode correctly models the astronomical ocean tide alone. In practice it may have been tuned by testing against measurements from the tide gauges, and so include some atmospheric effects. This needs checking by the model designers - eg it may be wise to tune the model only against median weather conditions. If it is tuned using M2, this has minimal sensitivity to weather.
  iii How is the tidal model forced? Which constituents does it include? Are there errors from the boundary forcing? The model is forced by 26 constituents, not including S1, or any constituents longer than 1 week.
c The purpose of C3X or NEMO-Surge is to capture the well-documented non-linear interactions of the tide and surge. [e.g. Proudman, 1955]. Yet this step and [f] disregard some of the non-linear effects. We cannot find previously documented discussion of this aspect of the forecast and will discuss in more detail below.
d i The assumption here is that the astronomical tide is better estimated from the tide gauge than from the model.
  ii There may be observational errors, and gaps in the record.
  iii The record may be too short to derive the required harmonics
  iv There may have been changes to the tide since the gauge record from which the harmonics are derived, for example due to harbour works or dredging. However unless there is specific reason to believe that this is the case locally, then a long (>18 yr) record is preferable for deriving the harmonics. Further study could assess how many of the records show harmonic changes over time which would make it preferable to only use current data for harmonic analysis.
  v The set of harmonics may not capture all of the complex behaviour of the tide, particularly in shallow estuaries. This is considered below.
e The gauge is measuring the total water level, and hence the harmonics H_g include all wave, steric and surge effects. This is not therefore a prediction of the astronomical tide alone. This is a key effect and will be discussed below.
f See (c). Also

  i The gauge is not directly comparable to the model because it is a point measurement and the model is an average over a grid cell. There may also be timing difference between the gauge and model.

g See (d). Also, there may be drifts on the gauge, because of vertical land movement and sea-level rise. Such effects that apply over a large area should probably remain in the forecast (since it affects flood risk) but those which apply only to the gauge (eg subsidence of the pier, siltation of the equipment) should have been removed before comparison. This must be examined at individual sites, by experts in tide gauge maintenance, vertical land movement and data quality control.
h Ideally, better understanding of all effects should reduce this empirical step. The problems arising from (g) mean this step is unlikely to be eliminated.

2.2 Error in the current procedure.

With the NEMOSurge data (see appendix B) and corresponding tide-gauge data we can reproduce the forecast and plot the error from the observations, figure 1. This is $W_g - W_f = W_g - (M_r + H_g)$, the difference between observations and the sum of model residual and gauge harmonics. A similar annual cycle is seen as in figure 11, but the monthly cycle is less. This lends some support to the hypothesis about the double-counting of radiational tides, but evidently it is not the only issue.

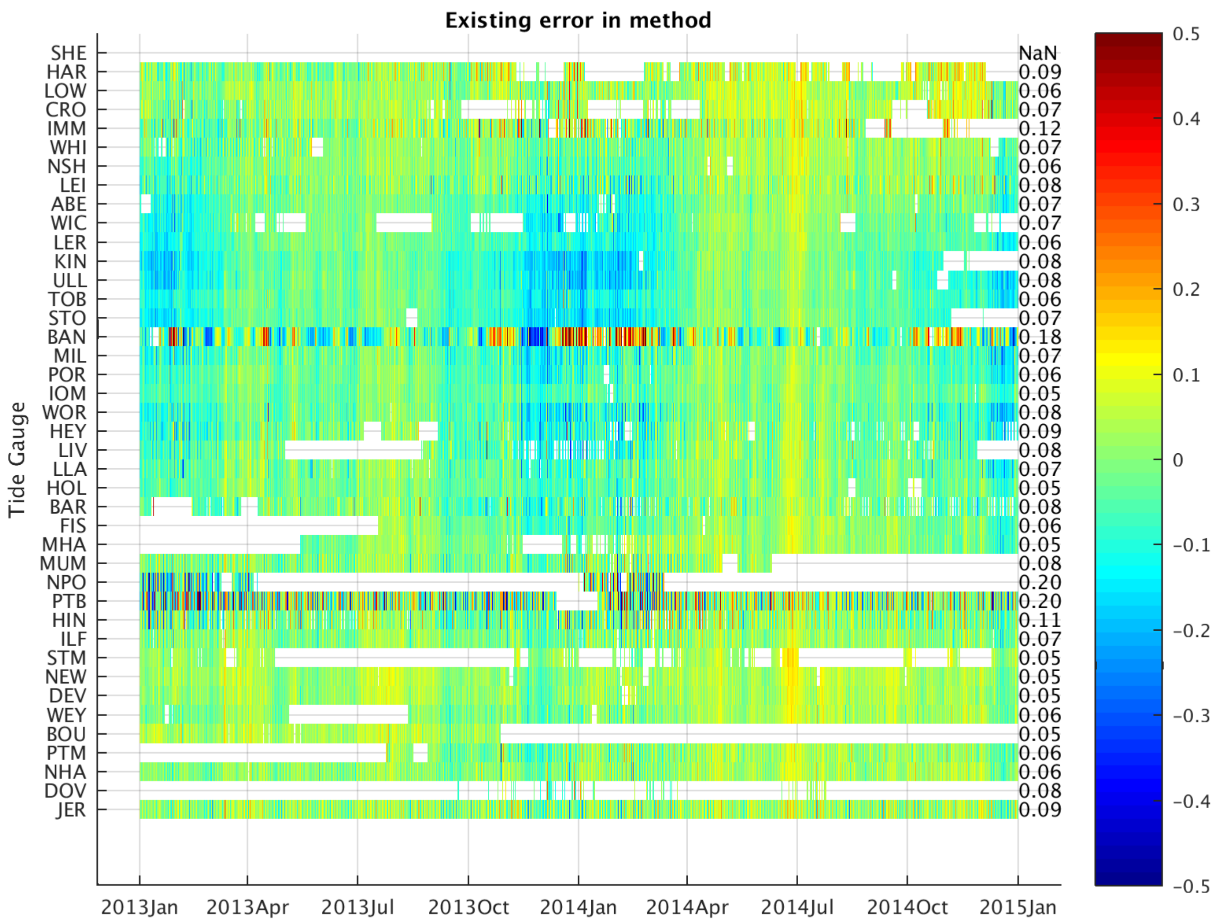

Figure 1: Difference (m) between observations and forecast water level (before empirical correction). Gauges are UK only, anti-clockwise from Sheerness to Dover and Jersey. Gaps in the plot represent flagged tide gauge data. Standard deviation for each gauge indicated. Data flagged by BODC is blocked out, and omitted from the standard deviations.

3 Double-counting of atmospheric effects

As noted in point [e] above, there is potential to double-count the effect of the weather in the forecast procedure. The assumption that is made at step [e] is that the harmonics derived from the tide gauge can be used to predict the astronomical ocean tide, and that they are equivalent to the model run with no-tide.

Earth tides are included in the gauge, but not in the model, so there is no double counting.

Atmospheric tides will be seen at the gauge and in the weather pressure forecast (since they are in pressure measurements), so will be double counted.

Seasonal effects on pressure will be double counted as tides. This has been briefly discussed in previous documentation [e.g. O'Neill and Saulter, 2017, Flather et al., 2000] but not quantified. It appears to be handled via the final "empirical correction" step.

Surge effects of storms may tend to introduce a consistent bias through non-linear effects. For example, suppose a low pressure tends to advance the tide somewhat, reducing phase, and a high pressure increases phase. This then manifests in phase change to all constituents, but especially to the large M2 and S2. M2 is decorrelated from radiational effects, but S2 is exactly semi-diurnal so could correlate to atmospheric pressure cycles. This example is a reminder that changes in phase as well as amplitude of constituents are potentially important.

Comparing the GSTM model with/without surge (see Appendix A), amplitude change of S2 in the Bristol Channel is about 0.04 m, however there is a phase change of around 3° (7 min).

M2 has a period of 12.42 hours and S2 exactly 12 hours. Through a lunar month they gradually move in and out of phase with each other, resulting in a spring-neap cycle. A small change in phase to S2 harmonic would result in a change of which days it is in phase with M2, and hence a substantial change in total tidal amplitude. This regular cycle also appears to be handled via the final "empirical correction" step. This cycle could account for that seen by Byrne et al. [2017], and that in figure 2 of Flowerdew et al. [2010]. See figure 2.

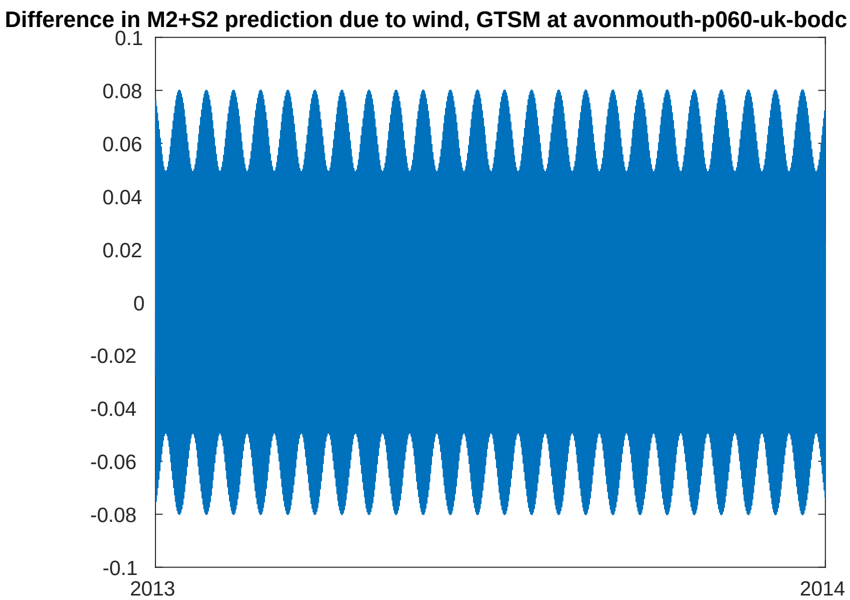

Figure 2: Cycle of prediction change (metres) due to small changes in constituents M2+S2 alone, between model with/without surge. Avonmouth S2 amplitude change = 3.5 cm, phase change 3.5°, M2 amplitude change 1 cm, phase change 0.2°.

A note on S2 from Pugh and Woodworth [2014]: "The semidiurnal term S2 also has slightly anomalous phases and amplitudes in the data from all parts of the world; response analysis separates the gravitational part from the radiational part of S2. Physically, the regular variations of tropical atmospheric pressure with a period of 12 hours and maximum values near 10:00 and 22:00 local time, the atmospheric tides, are the driving force. These pressure variations have amplitudes that vary approximately as $1.25 \cos^3 \phi$ millibars (where ϕ is the latitude)."

Dobslaw and Thomas [2005] showed that the ocean tide that results from this atmospheric pattern is quite different depending on the frequency of forcing, and the 6-hourly ECMWF reanalysis used in the GTSM model here may be inadequate to capture it.

S1 is also a known radiational component but is smaller at these sites (< 0.02 m amplitude everywhere in UK, < 1 mm everywhere in UK when fitted to the tide-only model).

MA2 and MB2 are also radiational [Flather et al., 2000], representing modulation of M2 due to seasonal changes in the weather. They have amplitudes of 5 mm–2 cm in the UK (GTSM model including wind), $< 5\ mm$ when fitted to the tide-only model. They may be difficult to separate from M2.

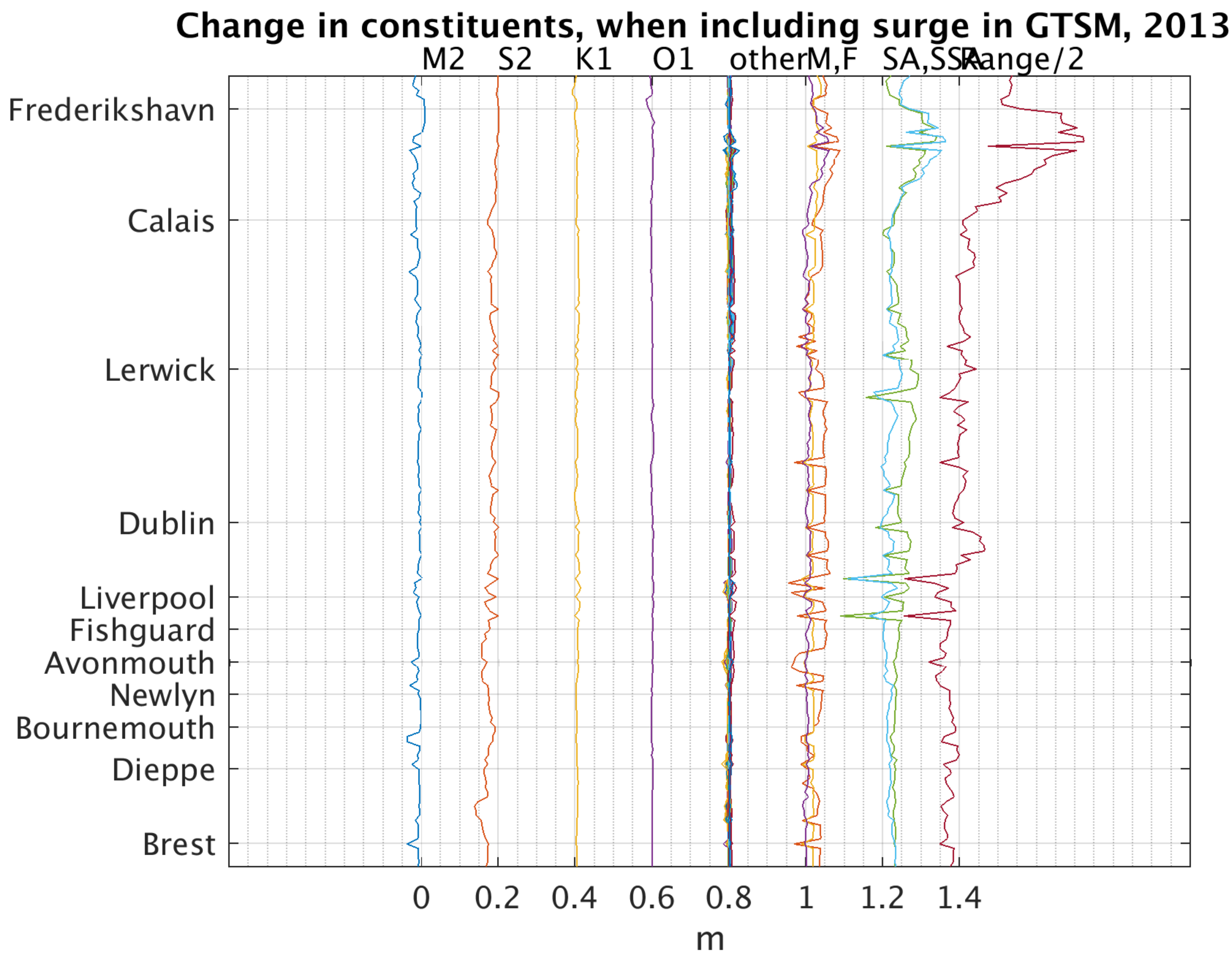

Figure 3: Difference in amplitude (m, offset) between coefficients fitted to model including surge or tide only, GTSM 2013 only.

3.1 Quantifying the double-counting of atmospheric effects

We can test a minimum effect of this double-counting purely within the model, by using the harmonics of the model *including surge*, as a proxy for the harmonics of the observations at gauges. Writing:

H_s	Tide prediction from model including surge
H_t	Tide prediction from tide-only

Then the forecast procedure can be estimated as $M_r + H_s$.
To estimate the error in this model forecast we can once again use the model, assuming $M_s \approx W_g$
Then the error is given by

$$\text{error} = M_s - (M_r + H_s) = M_t - H_s.$$

That is, the minimum error from the current forecast procedure is equal to the error in the harmonic prediction *including surge* at estimating the *tide-only* model.

This can be seen in figure 4. There is a monthly cycle of about 0.1m at nearly all UK gauges, although in the Bristol Channel it is larger and noisier - up to 0.5m. This is a similar result to that found by Flowerdew et al. [2010].

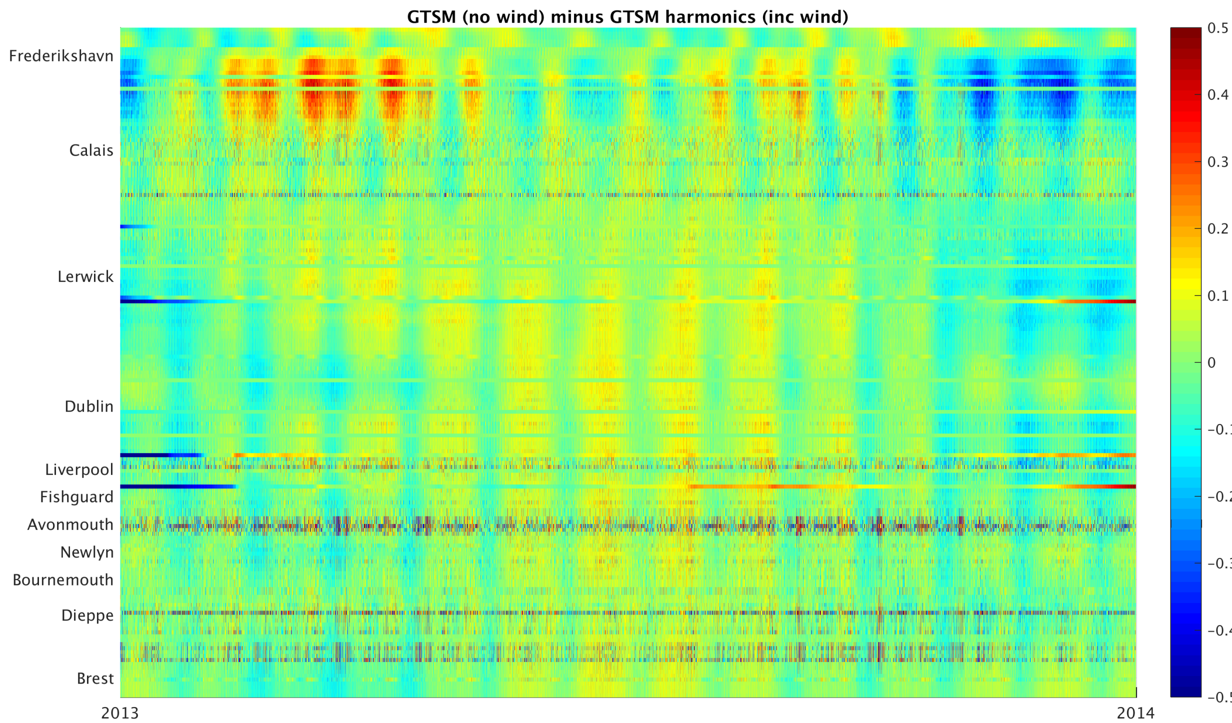

Figure 4: Error (m) in Harmonic prediction (62 constituents) *including surge* at estimating the *tide-only* model, GTSM 2013 only. See appendix A for explanation of model and coastal axis.

Figure 5 is equivalent to figure 4 (although note change in time axis and limited selection of gauges). Once again we see that $M_t - H_s$ has annual and monthly cycles, possibly due to the radiational constituents occurring in the surge and missing from the tide-only model.

3.1.1 What if we could provide astronomical tidal harmonics for the observations?

If it were possible to avoid the double-counting, and provide *astronomical* tidal harmonics for the observations, this would instead be equivalent to $M_r + H_t$, and the error would become $M_s - (M_r + H_t) = M_t - H_t$, as shown in figure 6. Since we're using the model as proxy for observations, if the harmonic prediction is an exact reproduction of the tide-only model then this would be exact. It is less than 5 cm at most UK sites and the monthly cycle has gone, but in the Bristol Channel there is still an error of up to around 0.5 m. This indicates that the selection of 62 harmonic constituents are not capturing all of the model tide. We will return to this below.

3.2 Can we avoid the double-counting?

One possibility is to omit the harmonics of the tide gauge data associated with including atmospheric effects from the tidal prediction. If the atmospheric effects were purely constrained to constituents which didn't include astronomical tide this may be possible, but in practice there are constituents, notably S2, which contain both. There is an astronomical atmospheric S2 tide of amplitude $1.25 \cos^3 \phi$ at latitude ϕ (approximately 1–3 cm in the UK) but there are other greater variations in S2 that arise locally.

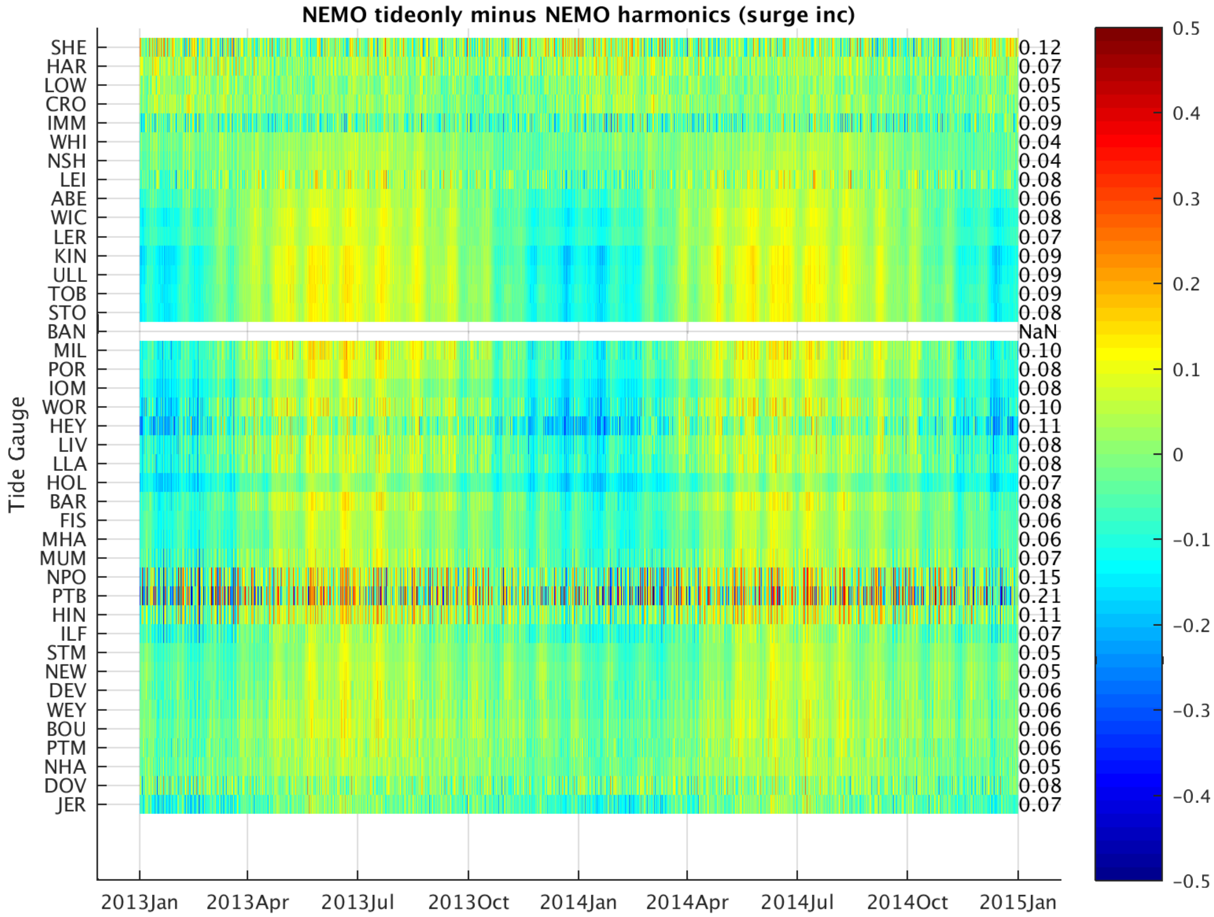

Figure 5: Error (m) in Harmonic prediction (62 constituents) *including surge* at estimating the *tide-only* model, in NEMOSurge 2013–2014. Gauges are UK only, anti-clockwise from Sheerness to Dover and Jersey. Standard deviation for each gauge indicated.

There is a clearer case to be made for Sa and Ssa, and Flather et al. [2000] recommended excluding these constituents from the harmonic tide predictions use for the operational procedure.

Flather et al. [2000] also notes that there are differences between the modelled and observed constituents Sa and Ssa in C3S.

The problem then comes back to modelling locally which of the harmonics arise from the radiational or atmospheric tides, which brings us back to the tide and surge model. Whilst the model may not be considered accurate enough to capture all of the local ocean processes at a gauge, it is possible it still includes sufficient information on the atmospherically generated tides to provide a correction factor. The greater part of the signal will be in a few constituents, including Sa, Ssa, MSf, S2 and S1, and there will be little contribution from the higher harmonics.

Suggestion for amended procedure: At stage [f]

- Derive harmonic constituents of the radiational tides using the model *residual*. This will necessarily be from a hindcast of the model, of at least 1 year, not from the live forecast.
- Produce a forecast of the radiational tide from harmonics, H_a.
- Remove this from the model forecast residual, before adding to the tide gauge harmonics: $W_{f2} = M_r - H_a + H_g$.

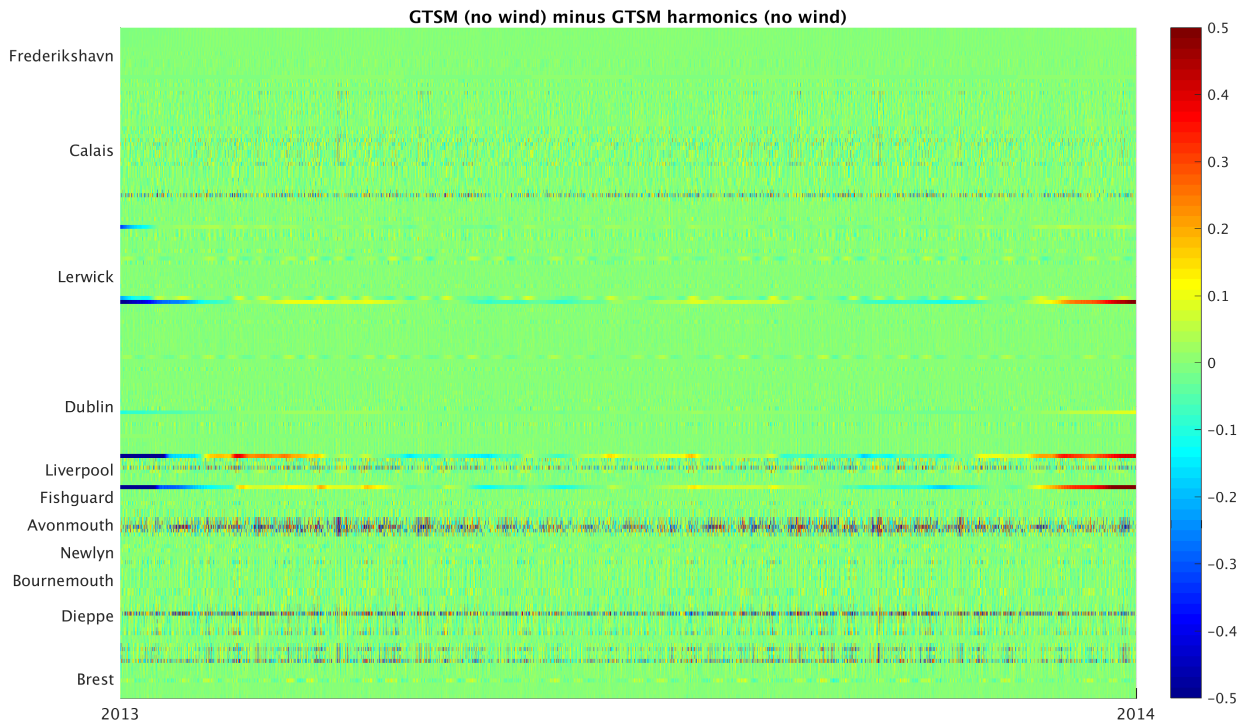

Figure 6: Error (m) in Harmonic prediction (62 constituents) *tide-only* at estimating the *tide-only* model, GTSM 2013 only. See appendix A for explanation of model and coastal axis.

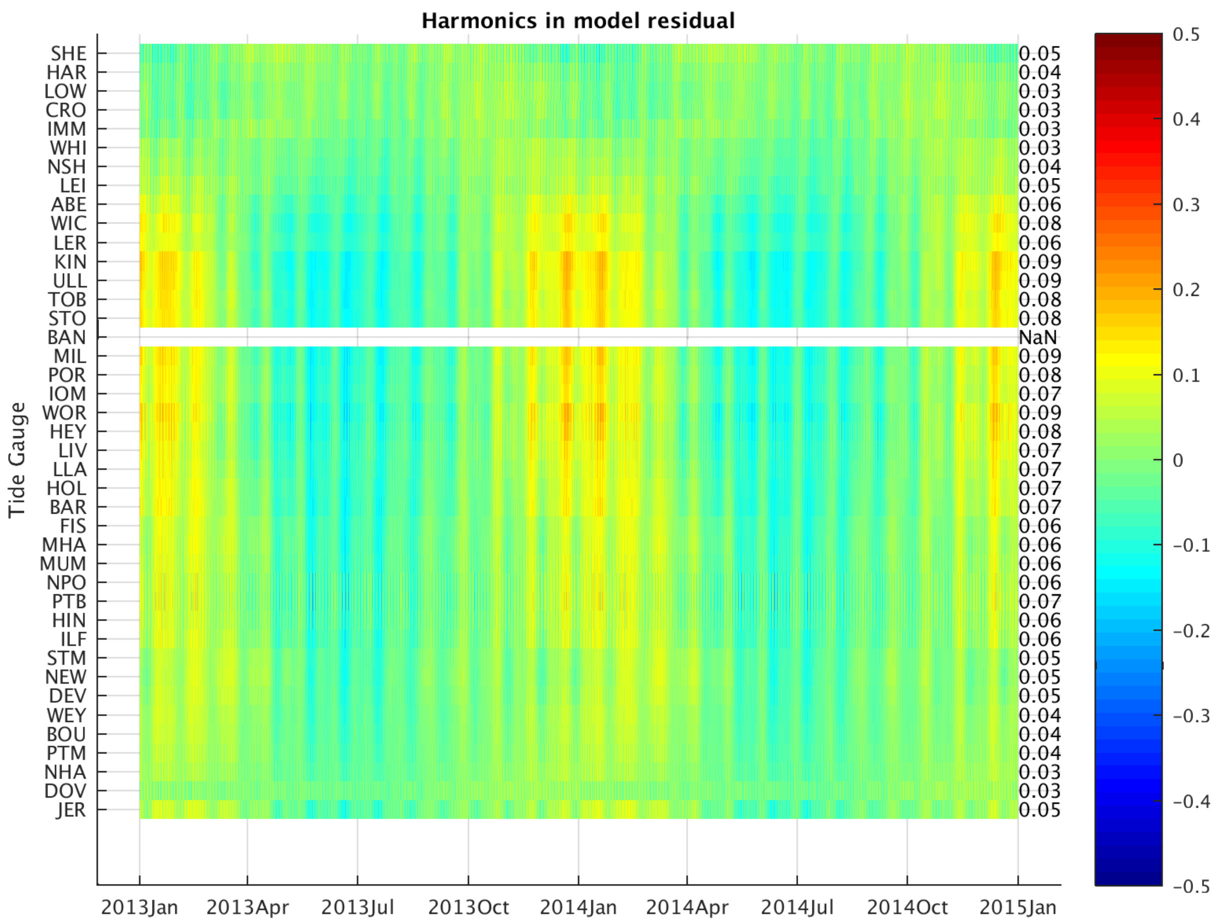

Figure 7: Prediction (m) from harmonics in the NEMO-Surge model residual, that may be double-counted. Gauges are UK only, anti-clockwise from Sheerness to Dover and Jersey. Standard deviation for each gauge indicated.

3.3 Testing removal of double-counted residual harmonics on the 2013–14 NEMO-Surge model

We can test the suggestion of amending the procedure to remove one instance of the double-counted harmonics H_a, replacing W_f with $W_{f2} = M_r - H_a + H_g$. The harmonics H_a are derived from a tidal analysis performed on the model residual, see figure 7. Some of these do not appear to be in the forecast error, and if all of them are removed (figure 8) then new errors arise, with periodic underprediction in the summer. It seems to be better to only remove harmonics known to be long term radiational, Sa, Ssa, MSf, as shown in figure 9.

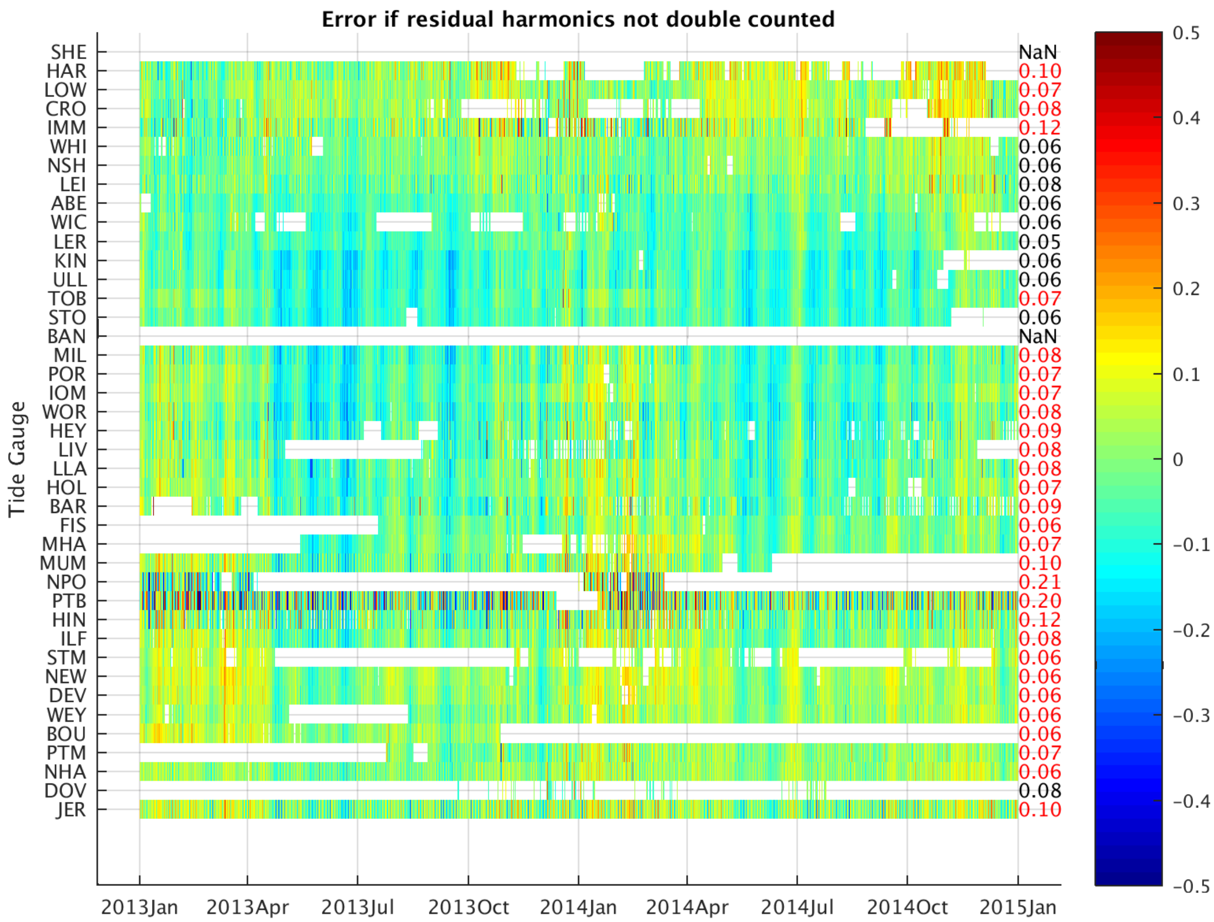

Figure 8: Difference (m) between observations and forecast water level (before empirical correction), avoiding double counting of all harmonics that are in the model residual. Gauges are UK only, anti-clockwise from Sheerness to Dover and Jersey. Standard deviation for each gauge indicated, in red if worse than figure 1. Data flagged by BODC is blocked out, and omitted from the standard deviations.

This correction does not address the problems at Avonmouth (PTB) or its neighbours, although standard deviations of error indicate some improvement at sites on the north-west coast.

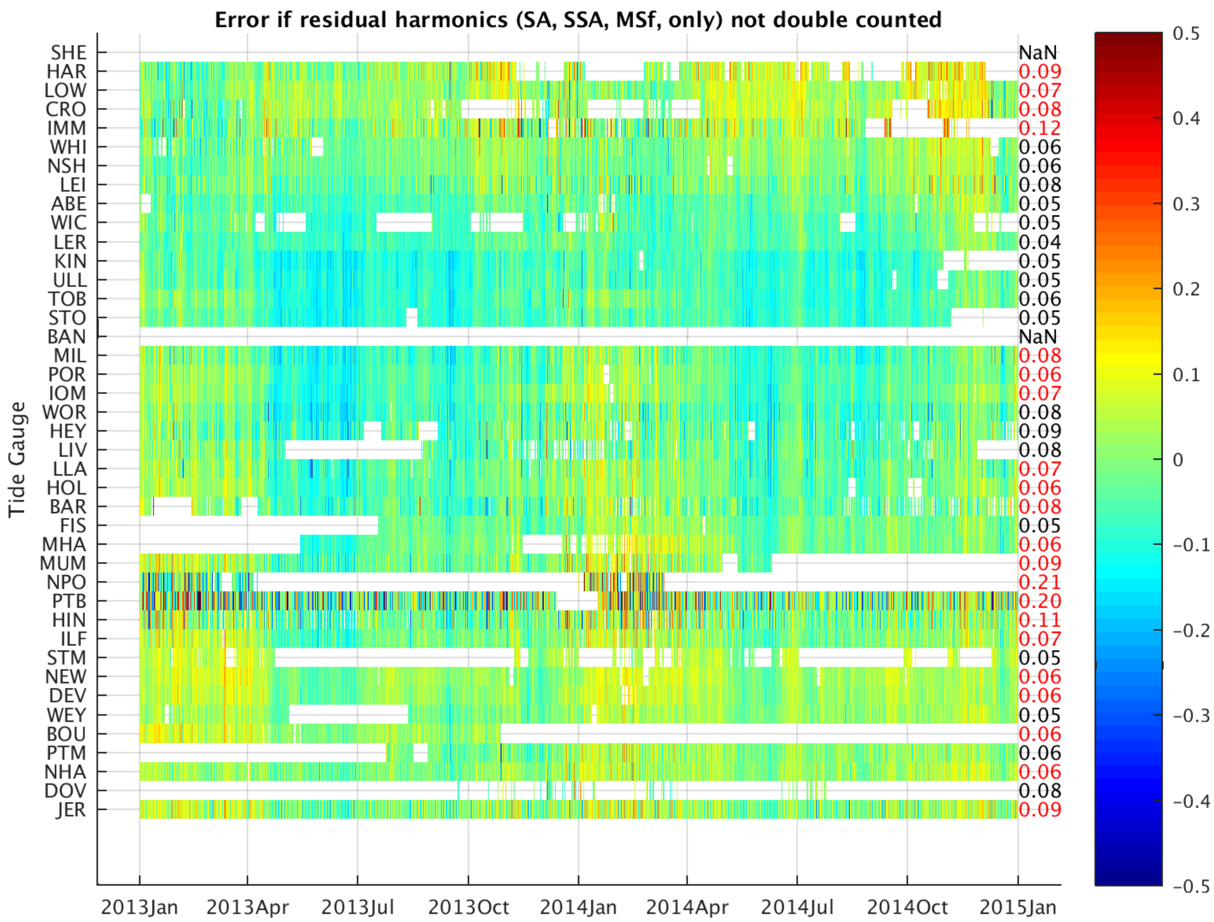

Figure 9: Difference (m) between observations and forecast water level (before empirical correction), avoiding double counting of Sa, Ssa and MSf harmonics only. Gauges are UK only, anti-clockwise from Sheerness to Dover and Jersey. Standard deviation for each gauge indicated, in red if worse than figure 1. Data flagged by BODC is blocked out, and omitted from the standard deviations.

4 Quantifying error due to disregarding non-linearity

The approach of linear addition of the harmonic prediction to the non-linear residual, $W_f = M_r + H_g$, in itself carries a risk of error, even if the harmonic prediction did not include any astronomical forcing. It is not easy to quantify, and I am not aware of any prior reports attempting to do so, although this may account for some of the error seen in figure 3 of Flowerdew et al. [2010].

Consider the following simplified example. The tide has an M2 amplitude of 3 m (ignore all other constituents). There is a surge of constant amplitude 0.2 m, which also advances the tide by a constant 30 min. The harmonics of the observed tide have the same amplitude, but are out of phase by 5 min (equivalent to 2.4° M2 phase change). As illustrated in figure 10, the residual is decreasing during High Water, so if the observed harmonics have High Water later than the model the forecast skew surge is underestimated by 3 cm. If the observed harmonics predict High Water earlier than the model, the forecast skew surge is overestimated by 3 cm.

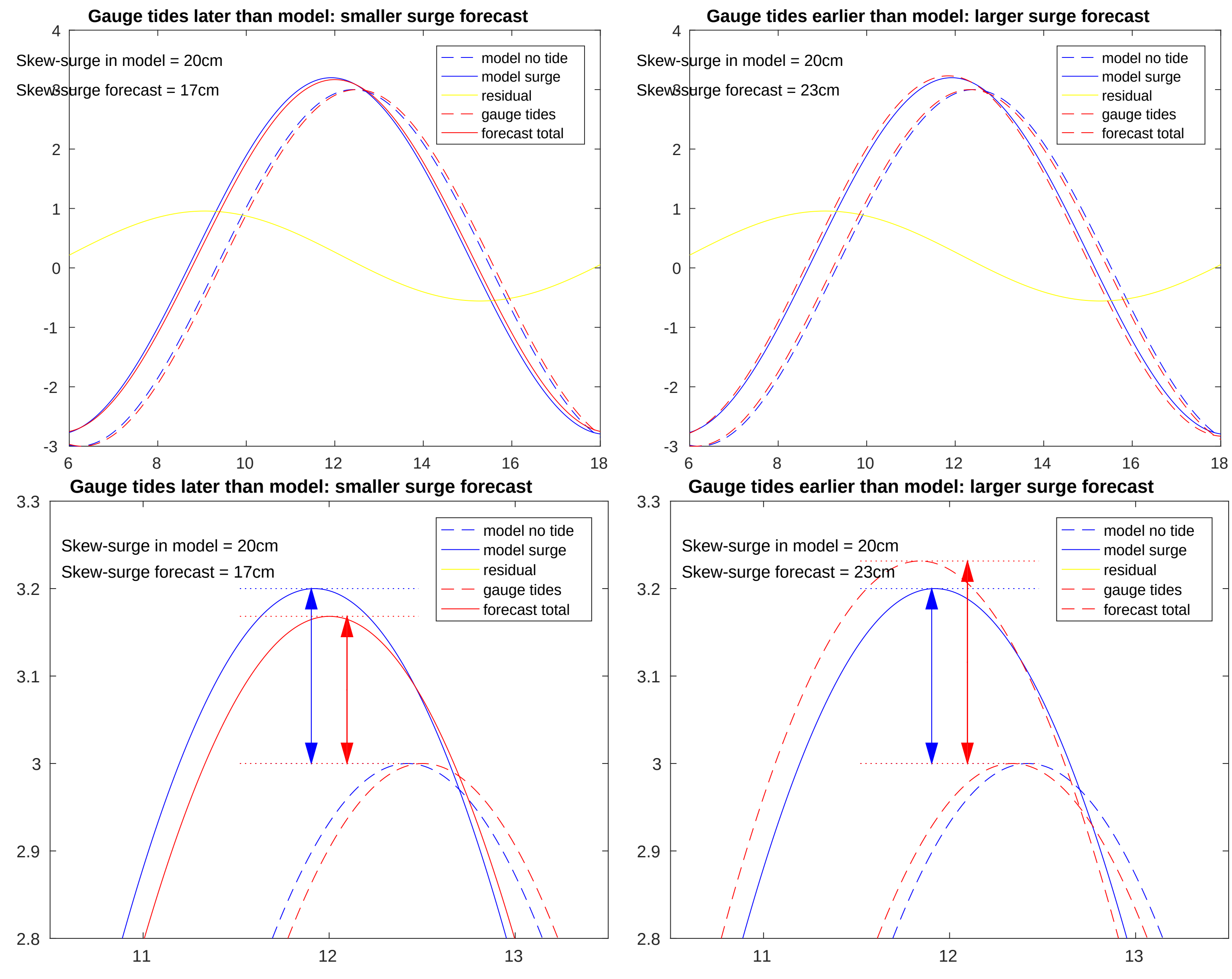

Figure 10: A surge is imposed which adds a constant amplitude of 20cm and advances the tide by a constant 30 min. If the harmonics of the observations differ in phase from the model a forecast error will result as shown.

This quick calculation indicates that surge amplitude may be affected by the non-linear calculation, but to quantify this more accurately requires an estimate of the error in the tidal prediction.

The M2 phase at Avonmouth is:

NEMOSurge, tideonly	200.43°
NEMOSurge, surge	200.87°
Colin Bell's predictions from observations	200.61°

This range results in a smaller error, but since in practice the phase of the tide is a function of the other constituents as well, such a non-linear effect should not yet be dismissed. As discussed above the harmonics of the observed total water level are a less accurate estimate of the tide-only model, and this will contribute to non-linear error.

5 Steric effects

Flather et al. [2000] also discusses steric effects - that is changes in the sea-level on the time-scale of weeks due to ocean temperature (and occasionally salinity) changes. In the UK the largest contributor of steric effects will be an annual cycle, and hence an adjustment of the observed Sa and Ssa. Flather et al. [2000] reports an effect on amplitude of 2–5cm and 1–2cm respectively. The tide-and-surge model knows nothing of this (although it could be plausibly incorporated as an additional term based on global ocean modelling). For now, the annual cycle is included in the Sa and Ssa from tide-gauge harmonics, and further adjustment for unusual seasonal temperature changes would be in the final empirical correction.

We could estimate the magnitude of these effects using for example POLCOMS UK shelf model. The following runs, for 2008 are available from Jenny Brown:

- Tide only
- Barotropic — tide, river input, surge, forced with wind and pressure
- Baroclinic — Tide and surge, forced by boundary conditions and an initial heat and salinity field and heat flux from the Met data. The river inputs also have temperature and salinity.
- Baroclinic + waves.

Comparison between the residuals Barotropic-tides and baroclinic-tides can give us an indication of the magnitude of the steric effects. Maps of the difference in maximum water level in each month indicate a 10–20 cm seasonal cycle, with greater baroclinic effects in the summer. However initial indications are not sufficient to show the increased problems in prediction in the Bristol Channel, and more work (several weeks at least) would be required to examine this data in greater detail. Also as we only have 2008 then direct comparison with the events of 2013–2014 is not possible.

6 Importance of high-order constituents in the Bristol Channel

Recall that figure 6 showed that the harmonic predictions from the tide-only model were unable to reproduce the tide-only model at Avonmouth as well as at most other UK sites. The same is seen in the NEMOSurge model, figure 11, with harmonic predictions from the tide-only model unable to reproduce the tide-only model to better than about 20 cm. Harmonic predictions from the observed data may be better (being based on more than 2 years, so more constituents can be resolved). However it is possible that there are still insufficient constituents to accurately reproduce the tidal cycle at these sites using the harmonic method. To progress further here we need to go into more detail about the tidal cycle in the Bristol Channel.

The tides at Avonmouth have contributions from tidal constituents including for example 3MS4 and 4MS6 that are not included in the 62 constituents for a year's analysis, but are in the 18 yr list of 115. They may not be fully resolved after only 2 years, but there is improvement in the harmonic fit from standard deviation 0.2 m to 0.1 m.

This can also be seen in the spectra of the tide-only model and the harmonic predictions, particularly around M4.

Further constituents could potentially be included in the fit, but this would require a careful rewrite of the code, which has the nodal correction for a maximum of 115 constituents hard-wired. It would be interesting to test this but this is going to take more work.

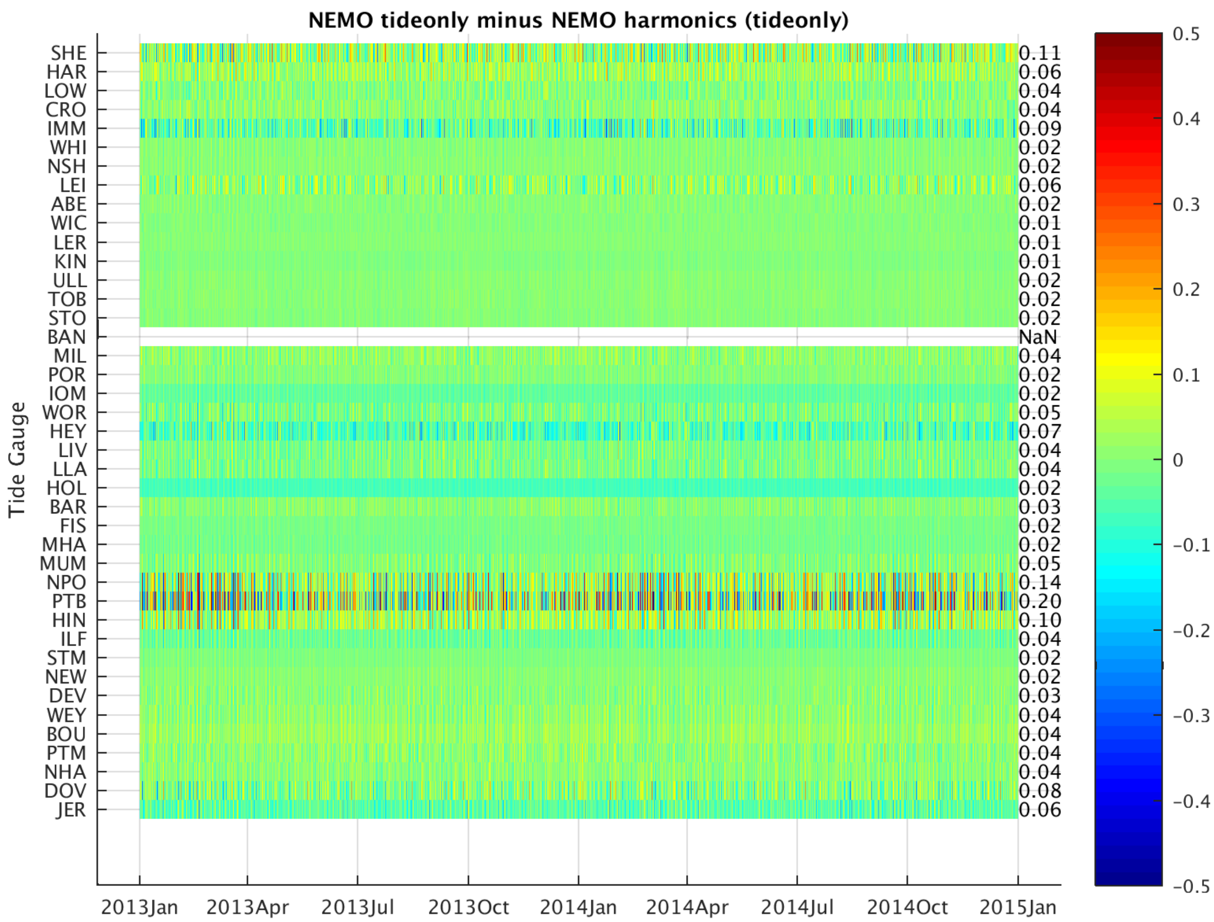

Figure 11: Error (m) in Harmonic prediction (62 constituents) of *tide only* model at estimating the *tide-only* model, in NEMOSurge 2013–2014. Gauges are UK only, anti-clockwise from Sheerness to Dover and Jersey. Standard deviation for each gauge indicated.

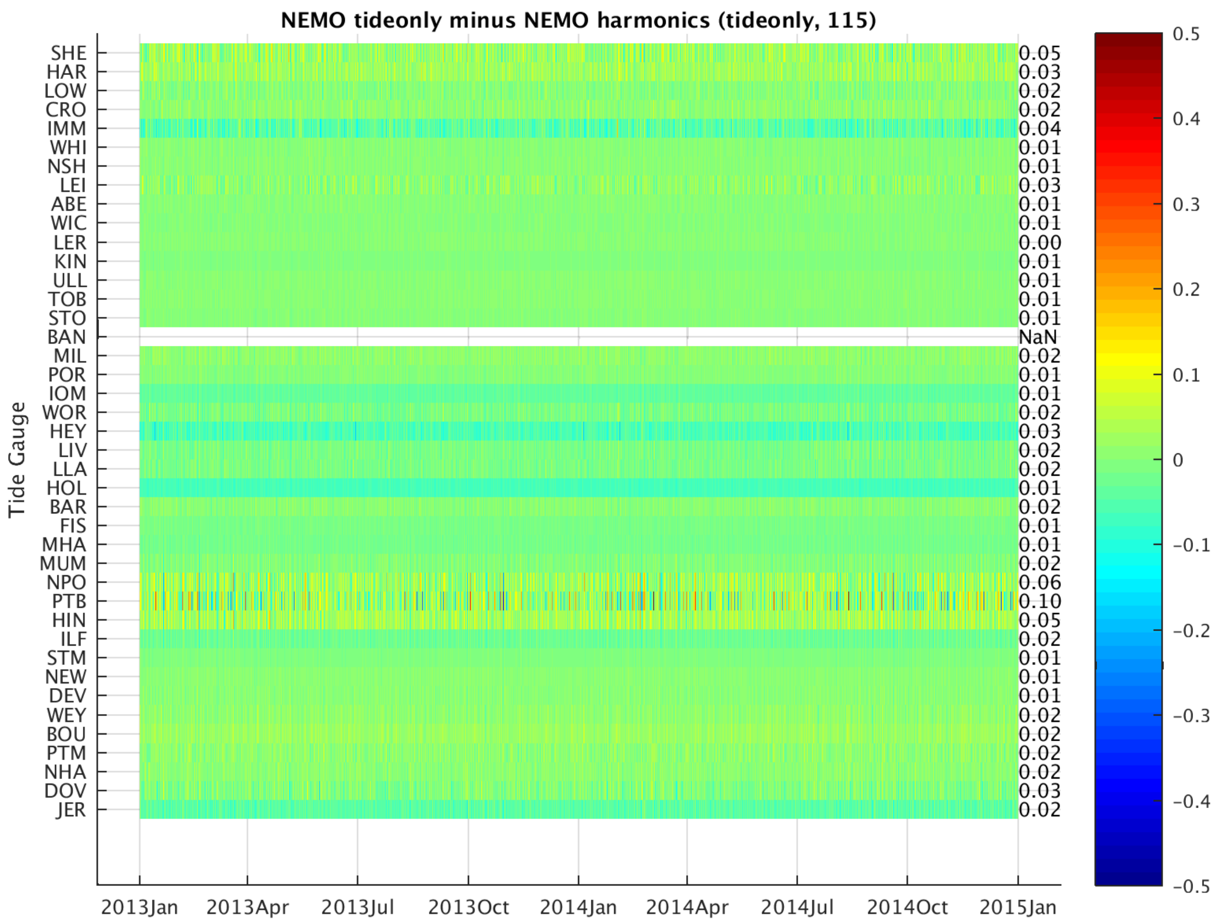

Figure 12: Error (m) in Harmonic prediction (115 constituents) of *tide only* model at estimating the *tide-only* model, in NEMOSurge 2013–2014. Gauges are UK only, anti-clockwise from Sheerness to Dover and Jersey. Standard deviation for each gauge indicated. A clear improvement from the 62 constituent harmonic prediction is seen at Avonmouth. However there remains a 10 cm error, indicative of a minimum error that must exist in the observation harmonics.

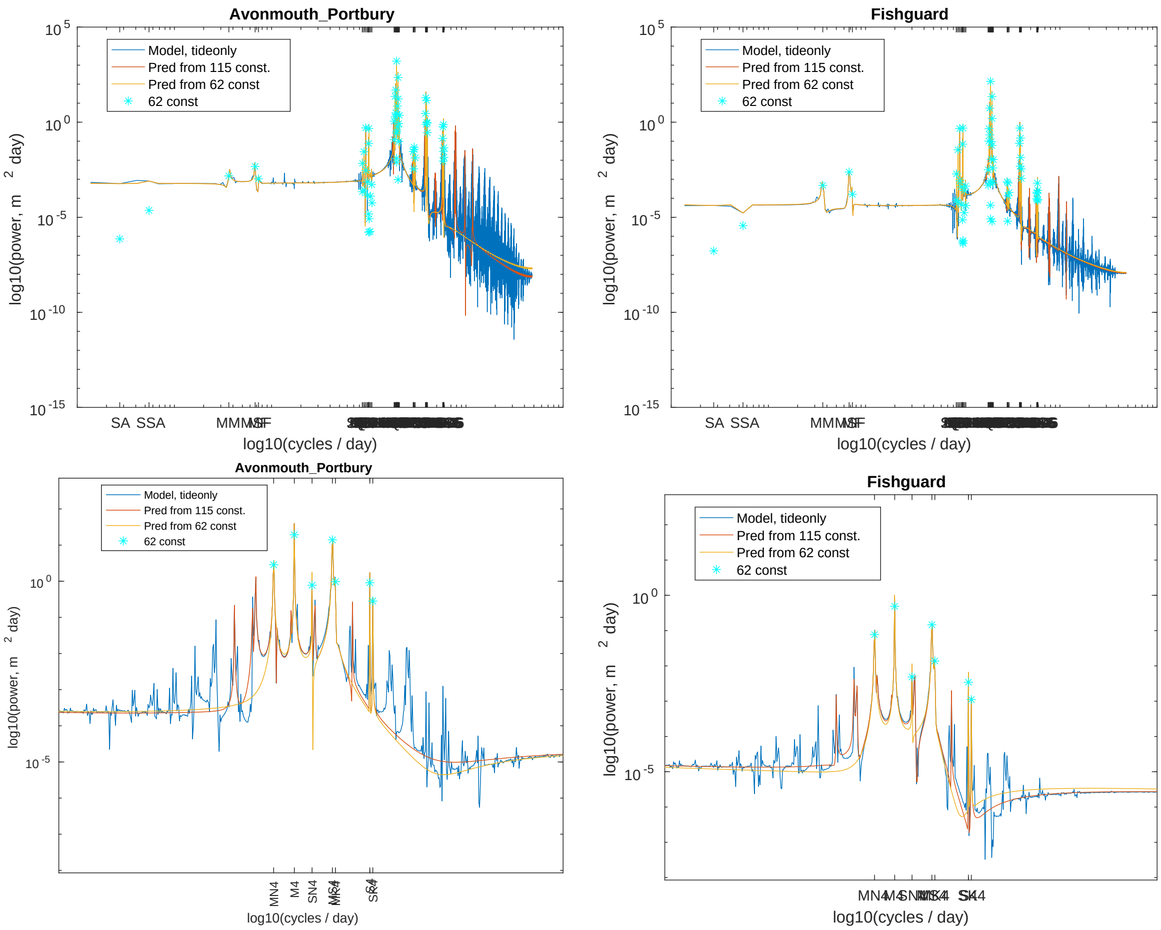

Figure 13: Spectrum of *tide only* model (blue), and the harmonic predictions (red and yellow) in NEMOSurge 2013–2014, at Avonmouth and Fishguard. Note that the spectrum is only a rough guide, since it includes no nodal corrections. At Avonmouth, there are non-linear cross-terms containing energy at frequencies other than the list of 115 (where the blue line is not covered by the red or yellow).

7 Conclusions

- Some effects of the weather on tides are double-counted in the forecast procedure. Even if the model were perfect, the minimum error from the current forecast procedure would be at least the error in the harmonic prediction *including surge* at estimating the *tide-only* model. If 62 constituents are fitted, this has a standard deviation of 20 cm at Avonmouth and 4–10 cm at most other UK gauges.
- This result is consistent across two very different tidal models, GTSM and NEMOSurge.
- It may help to amend the forecast procedure to remove a prediction time series from harmonics H_a derived from a tidal analysis performed on the model residual. At some sites this reduces the forecast error by around 1 cm. Removing only residual harmonics known to be long term radiational, Sa, Ssa, MSf, results in tiny improvements to a few more sites. But there is no improvement in the Bristol Channel.
- It is important to consider changes in phase as well as amplitude of tidal constituents. A difference in phase between observational harmonics and model tide-only harmonics leads to an exaggeration of non-linear effects, as the model residual is added to the wrong phase of the tide. Typical errors could be a few cm.
- Steric effects could contribute a significant portion of the error, with a 10–20 cm seasonal cycle on the change in maximum water level in each month, with greater baroclinic effects in the summer.
- Care must be taken comparing tidal constituents from different sources, as different lists may be used.
- The problems specific to Avonmouth may arise from the harmonic prediction. The harmonic fit reproduces the tide-only model to 20 cm if 62 constituents are used, but 10 cm if 115 constituents are used. This is because although the model is only forced at the open-ocean boundary by 26 constituents, non-linear interaction in the Channel generates further cross-terms. The spectrum of model data indicates that resolving more non-linear constituents may further improve the fit.
- Phase demodulation may provide a diagnostic tool, to highlight occurences when phase shift leads to a large model residual without necessarily a change in total water level.

7.1 Suggestions for further work

- Check the timing of forecast errors at Avonmouth. For example, if at low tide it may be due to river runoff, if at rising tide, we suspect phase problems, if at high tide then it's something else. This would be quick to check.
- Investigate effect of precise differences in tidal constituents from models and observations at Avonmouth. This could be quite quick as the necessary details are available (from Colin Bell). Check maximum error due to non-linearity due to these differences.
- Calculate the model power in higher order constituents at Avonmouth. Can the observation harmonics be improved by the addition of further constituents, and if so does this improve the observations? Is it better to include higher harmonics even with a time series that is theoretically too short to resolve them? To amend the code to include nodal correction to further constituents will take at least a week, but further investigations are recommended. It may well be that this has been previously studied.
- Examine data from baroclinic model to investigate whether missing steric height or wave effects could be behind forecast errors. (Coordination with Jenny Brown required. Also note the existing baroclinic model run is 2008, whereas NEMOSurge is 2013-2014.)

8 Model validation guidelines

Part of the motivation for this work was the necessity of a consistent scheme to validate the surge models deterministic performance. The model development team is better placed to advise on the detail of appropriate validation statistics. However the following guidelines should be taken into account:

1. Model validation should be considered as having two components: 'Trials Verification' in which the development team aims to specifically assess the surge model's ability to generate a surge residual (i.e. the short term meteorologically forced component of the non-tidal residual); and 'Operational Verification' in which the overall ability of either a) harmonic tide prediction plus modelled surge residual, or b) model total water level to predict observed water levels are measured. Since the principle purpose of the surge model is to forecast peak water levels within a given cycle, it is an absolute requirement that the verification measures the skew surge component in addition to higher frequency residuals.
2. Trials Verification should be carried out on a minimum of 1 years data in order to properly capture seasonal variations in model bias and identify any long term drifts in the observations.
3. Tide gauge observation time-series should be thoroughly reviewed against both quality flags and harmonic prediction data. This approach enables the analyst to both identify and remove bad observations, but also to identify any systematic offsets between the harmonic predictions and observations. Such offsets may resulting from e.g. long term climate change, land shifts or instrument drift over the harmonic analysis period, which is typically 19 years for UKCFF (UK Coastal Flood Forecasting) port tide tables. This step is a definite requirement for Trials Verification, but is also recommended for Operational Verification (with a follow up recommendation that the resulting quality controlled observations and offset information are stored and provided back to UKCFF as part of the surge model support service).
4. For Trials Verification, the observed non-tidal residual should be post processed to remove both any systematic offset and Sa/Ssa components of the harmonic tide prediction. This reduces double counting of long term atmospheric contributions to the residual.
5. For both Trials and Operational Verification it is recommended that long term (e.g. 3-monthly) variations in water level bias are calculated, as these can be used to estimate contributions from long-term steric effects (Trials Verification) or describe variations in long term background error for forecasters (Operational Verification).
6. The verification process needs to acknowledge that differences between surge and non-tidal residuals will be introduced as a result of uncertainties in radiational harmonics (e.g. S2) and ability of the harmonic analysis to fit asymmetric tidal curves in shallow-water estuarine locations. It is recommended that, particularly in Trials Verification, the levels of these contributions to overall error are estimated through comparison of the model surge residual (model total water level model tide only run) against the model non-tidal residual (model total water level harmonic prediction from model data). The resulting errors should be considered as a target for achievable model performance.
7. To further quantify the effects of phase uncertainties introduced through the harmonic predictions, where sufficient data is available, residual errors should be analysed at specific phases of the tide (e.g. HW, HW+1 hour, etc) and/or tidal range (phase within the spring-neap cycle). Skew surges may have limited phase dependencies, but should be assessed for any dependency on tidal range.
8. Verification reports should include, or refer to, some level of tide gauge metadata including a commentary on the likelihood of measurements at the gauge in question to include a contribution from wave set-up during high energy storms, the period over which any harmonic analyses were carried out and a link to sea level trend data at http://www.ntslf.org/products/sea-level-trend-charts .

9 Recommendations

Based on the above we propose the following research and development activities to improve our understanding and attribution of harmonic errors for total water level forecasting and the comparison between modelled surge and observed non-tidal residuals. The underlying principle is that, for the present, the surge model should be developed to predict only the meteorologically induced surge component. Accuracy of the modelled surge residual requires good tide prediction within the model. Other methods should then be used to predict the differentials between surge and non-tidal residuals, arising from the harmonic prediction method. For example applying systematic and seasonal bias corrections via post-processing systems, or applying an estimate of harmonic prediction errors within the probabilistic forecasting framework used by UKCFF.

1. Future verification of the surge model and assessment of development needs should be based on a comparison with (at least) a full year of observed data, for which mean sea level bias correction is applied and the harmonic component used excludes the low frequency components Sa and Ssa (Note: within such a framework it can then be assumed that low frequency errors are dominated by steric effects which may be reduced with later model developments, e.g. using a baroclinic ocean model).
2. Further work should be undertaken to evaluate skew surge sensitivity to type c) and d) errors, and if this is demonstrably reduced relative to non-tidal residual errors, to emphasise its usage within operational forecast systems and as the primary variable that we verify.
3. Further work should be undertaken to evaluate the background errors associated with harmonic predictions (e.g. as an extension of Figure 12). The work would have two purposes: first to better establish target performance data for the surge model (e.g. we recognise that the surge model is performing properly when total water level prediction metrics are close to the error values attributable to the harmonics); second to establish a method to include these information within a probabilistic framework in order to better represent forecast spread (this might not be trivial for residuals since the variations in error will be correlated with tidal phase).
4. A long term aim of surge model development should be to follow the approach of colleagues in the Netherlands, who have developed a model in which the background tide is similarly as skillful as the harmonic prediction [Zijl et al., 2013, 2015]. Due to the complexity of tides around the UK coastline, this is a non-trivial task that may take a number of years to properly develop. However, the process of understanding the models present skill level can be started immediately by regularly evaluating the model total water level forecasts versus observations in addition to the harmonic plus model surge approach.

Crucially, this report and the recommendations above confirm the importance and necessity of applying a level of post-processing within the forecasting frameworks used by UKCFF, both now and in the medium-term future. This should be raised and discussed further within the UKCFF technical group.

Appendices

A Global Tide and Surge Model

In order to perform some quick tests of the typical tide and surge constituents, I have made use of some modelling tools and data already available to me.

GTSM is the forward Global Tide and Surge Model developed at Deltares on a base of Delft-FM. It is a flexible mesh model with resolution from around 50km in the open ocean to around 5km at the coast. I have various runs available, forced with ECMWF ERA-Interim 6-hourly reanalysis. Validation of the major tidal coefficients around the UK has been favourable, and although the surge forecast underpredicts the effect of major hurricanes, due to lack of resolution in the weather forecast, most surge events are captured. For this report a 2013 run is used, with an 11 day spin-up period in December 2012. Tidal coefficients are found by harmonic analysis of the 2013 results.

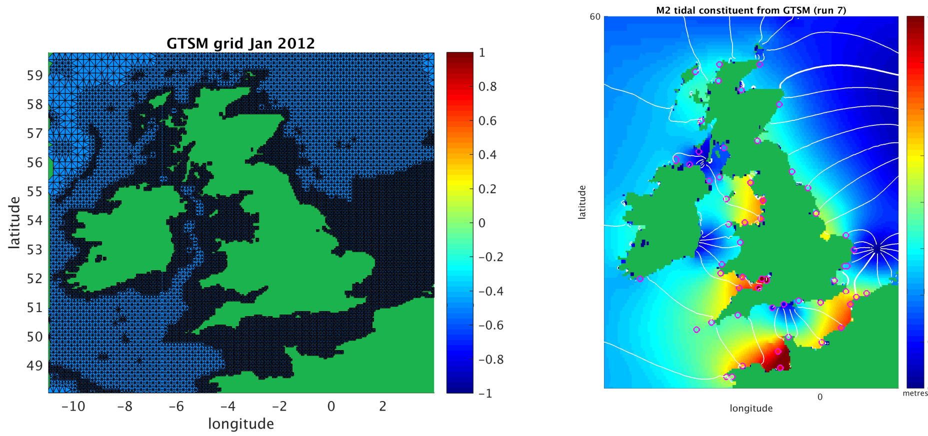

Figure 14: Grid density and M2 amplitude in GTSM model

A.1 Ordering of model sites around the UK

Due to limitations of data storage the model is only output at high frequency at all grid points for one month (Jan 2012) and a selection of coastal points for the year 2013. These coastal points are spaced roughly every 80km, plus wherever a tide gauge is situated. Due to automatic procedures to select output sites, a few may be incorrectly sited at model dry sites - these are clearly seen in plots as lacking sufficient high-frequency variability. The along-coast plots are ordered approximately anti-clockwise around the UK including neighbouring coasts in Europe and Ireland. The order is indicated in figure 15.

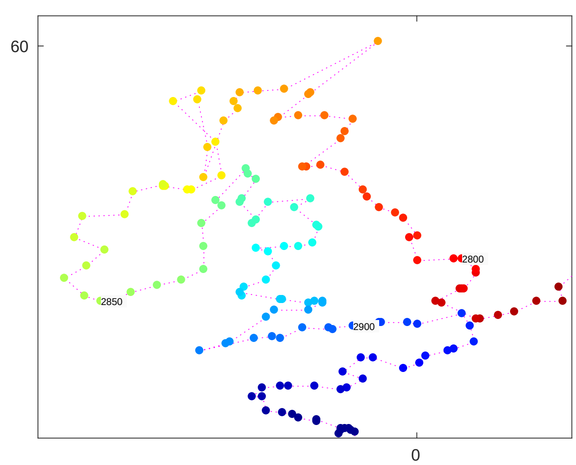

Figure 15: Sites used for analysis and coastal ordering (Red to Blue is top-to-bottom of other plots)

B NEMOSurge model and tide gauge observations

Clare O'Neill has provided 2013–14 NEMO-Surge hindcasts at the 41 ports (15 min snapshots). For comvenient comparison, I have downloaded 15 minute observation and residual data from BODC. These unfortunately are not necessarily based on the same list of harmonics as that in the predictions provided annually by NOC to the Met Office. For example at Portbury, the BODC prediction is based on 102 harmonics, whereas the predictions provided to the Met Office are based on 114 harmonics. Other gauge predictions may used different numbers of harmonics according to the data available.

C Harmonics

The harmonic analysis function is [Pugh and Woodworth, 2014, Chapter 4]

$$H(t) = \sum_{N} H_n f_n \cos\left[\sigma_n t - g_n + (V_n + u_n)\right] \tag{1}$$

where the amplitudes H_n and phases g_n are associated with the tidal constituents with astronomically-determined frequencies σ_n. $f_n(t)$ and $u_n(t)$ are nodal adjustments to amplitude and phase, applied in order to allow for the 18.61 year nodal cycle and 8.85 year longitude of lunar perigee cycle. V_n is the phase of the equilibrium tide, which we take as for Greenwich. We use UTC for all times to enable consistency between local gauges and global maps.

For the harmonic analysis I have applied an in-house matlab code based on the original Fortran code of David Cartwright which underlies the TIRA and TASK software. The underlying mathematics and total list of coefficients is the same. In this document 62 coefficients are fitted for 1 year unless otherwise specified.

The tidal predictions supplied to the Met Office by the NOC Applications team (Colin Bell) are based on a varying number of constituents according to the site and quality and length of data available. In the case of Avonmouth/Portbury, it is 114 constituents. (The 115th, 2MN2, with the same frequency but different nodal corrections as L2, is handled slightly differently).

The constituents used to produce the predictions supplied by BODC alongside the observational data are different again. The most recent Avonmouth output is based on 102 constituents, which omits MVS2, 2MK2, MSV2, SKM2, 2MNS4, MV4, 3MN4, 2MSN4, NA2, NB2, MSO5 and MSK5. This is the prefered list of 4.5 years of data.

References

David Byrne, Gavin Robbins, Neil Counsell, Andrew How, Andy Saulter, Clare O'Neill, and Jennifer Pope. Improving sea level forecasting at Newport. *Internal report*, 2017.

Henryk Dobslaw and Maik Thomas. Atmospheric induced oceanic tides from ecmwf forecasts. *Geophysical Research Letters*, 32(10):n/a–n/a, 2005. ISSN 1944-8007. doi: 10.1029/2005GL022990. URL http://dx.doi.org/10.1029/2005GL022990. L10615.

Roger Flather, Jane Williams, and David Blackman. Causes of seasonal sea level variations and implications for surge prediction. Technical Report 136, Proudman Oceanographic Laboratory, 2000.

Jonathan Flowerdew, Kevin Horsburgh, Chris Wilson, and Ken Mylne. Development and evaluation of an ensemble forecasting system for coastal storm surges. *Quarterly Journal of the Royal Meteorological Society*, 136(651): 1444–1456, 2010. ISSN 1477-870X. doi: 10.1002/qj.648. URL http://dx.doi.org/10.1002/qj.648.

Angela Hibbert, Samantha Jane Royston, Kevin James Horsburgh, Harry Leach, and Alan Hisscott. An empirical approach to improving tidal predictions using recent real-time tide gauge data. *Journal of Operational Oceanography*, 8 (1):40–51, 2015. doi: 10.1080/1755876X.2015.1014641. URL http://dx.doi.org/10.1080/1755876X.2015.1014641.

Clare O'Neill and Andrew Saulter. Nemo-surge: Forecast performance during 2016–2017 winter trial. Technical Report 622, Met Office, 2017.

J Proudman. The propagation of tide and surge in an estuary. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 231(1184):8–24, 1955. ISSN 0080-4630. doi: 10.1098/rspa.1955.0153. URL http://rspa.royalsocietypublishing.org/content/231/1184/8.

D. Pugh and P. Woodworth. *Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes*. Cambridge University Press, 2014. ISBN 9781107028197. URL https://books.google.co.uk/books?id=QiBGAwAAQBAJ.

Firmijn Zijl, Martin Verlaan, and Herman Gerritsen. Improved water-level forecasting for the northwest european shelf and north sea through direct modelling of tide, surge and non-linear interaction. *Ocean Dynamics*, 63(7):823–847, Jul 2013. ISSN 1616-7228. doi: 10.1007/s10236-013-0624-2. URL https://doi.org/10.1007/s10236-013-0624-2.

Firmijn Zijl, Julius Sumihar, and Martin Verlaan. Application of data assimilation for improved operational water level forecasting on the northwest european shelf and north sea. *Ocean Dynamics*, 65(12):1699–1716, Dec 2015. ISSN 1616-7228. doi: 10.1007/s10236-015-0898-7. URL https://doi.org/10.1007/s10236-015-0898-7.