Hydrology and
Earth System
Sciences

Supplement of

Climate change overtakes coastal engineering as the dominant driver of hydrological change in a large shallow lagoon

Peisheng Huang et al.

Correspondence to: Peisheng Huang (peisheng.huang@uwa.edu.au)

Supplementary material: model calibration, validation, and sensitivity test

1. Model calibration

In general, the approaches of model calibration can be categorized into automated calibration approach and manual calibration approach, and each of them has their advantages and disadvantage. Ideally, we could adopt an automated calibration approach, aiming to minimize error in key model predictions using an objective function and a pre-determined acceptable criteria for model acceptability (e.g. Doherty and Johnston, 2003; Arhonditsis et al., 2008; Bahremand and Smedt, 2010). This approach has yet to receive widespread up-take in the hydrodynamic modelling community, particularly where 3-D models are employed to resolve variability in stratification, due to considerable computational burden running these models hundreds or thousands of times. Given the complex nature of the model domain that we wanted to adopt to resolve the river reaches to the tidal limit, and an individual run-time exceeding one-day per year, we therefore could not adopt an automatic optimization approach. Instead, we adopted a structured hierarchical approach to calibration, similar to those described in Muleta and Nicklow (2004) and Hipsey et al. (2020), to manually calibrate the model. This approach first identified the key parameters of importance to the hydrology in the current study related to water mixing and heat exchange at the air-water interface, in order to reproduce the vertical and spatial salinity distribution, and the hypersailinty observed during the dry season. In this stage, the key parameters were identified to be the bottom drag coefficient (which can vary spatially), the bulk aerodynamic coefficients, the bulk transfer coefficient for latent heat flux, and the mixing scheme options associated with the vertical turbulence model (in this case this is parameterized through the GOTM plugin). In the second stage, a matrix of simulations, each with pre-determined parameter vectors and model options (Table S1), was assessed against the observed salinity and temperature data at six stations within the estuary (at both surface and bottom levels). The capability of the model to reproduce the salinity stratification (magnitude of difference between the surface and bottom salinity) created by the interaction of ocean intrusion and freshwater runoff during the wet season were also considered in the model calibration.

The year of 1998 was selected for calibration as this year has a median rainfall and catchment inputs in historical record. The model accuracy in reproducing the key hydrologic features was assessed by using the salinity and temperature data measured at six monitoring stations along the estuary (see Figure 1 of the paper). Monthly salinity and temperature datasets were obtained from the Marine and Freshwater Research Laboratory of the Murdoch University (1977-2001) and the Western Australia Water Information Reporting website

(http://wir.water.wa.gov.au/) (2002-2017). For each variable, we evaluated the model quantitatively against the monitored data using three skill metrics:

- r: regression coefficient, Varies between -1 and 1, with a score of 1 indicating the model varies perfectly with the observations and a negative score indicating the model varies inversely with the observations. A consistent bias may be present even when high score of r is obtained.
- MEF: modelling efficiency, measures the mean magnitude of the difference between model data and observations. This method compares the performance of the model to that only uses the mean of the observed data. A value of 1 would indicate a perfect model, while a value of zero indicates performance similar to simply using the mean of observed data.
- MAE: mean absolute error: Similar to RMSE except absolute value is used. This reduces the bias towards large events. Values near zero indicate good model skill.

The calibration started with a range of bed roughness length scales, a k-ε mixing scheme, a linear method for the aerodynamic coefficient (Wu, 1982), and a value of 0.0013 for the bulk transfer coefficient for latent heat flux. The last two options are default methods used by the TUFLOW-FV model. The performance assessment results (Table S2) indicate the scenario 4 better reproduced the salinity in terms of model performance statistic matrix of R, MEF, and MAE, and the salinity stratification, though the difference of performance between these scenarios were small. The selection of bed roughness length scales is shown to have minor impacts on the water temperatures. After the selection of the bed roughness length scales, we then compared the model performance with various options of mixing scheme, aerodynamic coefficient, and latent heat coefficient. The results suggest the use of the Mellor-Yamada mixing scheme had a similar performance to the k-ε mixing scheme, the selection with constant aerodynamic coefficient produced a relatively lower performance in the salinity and its stratification; whilst a selection of 0.0010 of E_L produced a better performance in the salinity stratification, but failed to match the performance in reproducing the salinity and temperature compared to the selection of 0.0013 of E_L.

Based on the model performance in the calibration progress, numerical schemes and parameters in the scenario 4 were selected and used for validation of other simulated years. For each variable, we evaluated the model quantitatively against the monitored data using three skill metrics: correlation coefficient (r), mean absolute error (MAE), and model skill score (SS) (Table S3). The validation results are presented in section 2 of this supplementary material. In addition, we also carried model sensitivity tests to a few selected environmental factors. The current study focused on the impact of reduced inflow, due to drying climate and the Cut, on the estuary hydrology. However, the perturbations of environmental factors such as air temperature, tide elevation, and benthic vegetation could also affect the local hydrology, and so their influence on the modelling results was explored. To evaluate the effects of these factors, the sensitivity of the τ and salinity was assessed relative to changes in: (1) air temperature (±1 degree, representing 100 year change of local air temperature); (2) tidal elevation (±0.15m, representing 100 year change of local tide record); and (3) bed roughness length (±50%, representing 50% change of bed roughness). The ranges of these environmental factors were carefully selected based on the historical records. Two years, 1990 and 1998, representing a year before the Cut-opening and another year with the Cut, respectively, were selected for these model sensitivity tests. The results of the sensitivity tests are presented in section 3 of this supplementary material.

We acknowledge that this approach is not necessarily providing the most optimum parameter set from a mathematical point of view, however, given other uncertainties in the spatial maps of vegetation (and therefore benthic drag) and potential error or bias in some of the assumed boundary conditions, it is our view that the model performance is close to the optimum and sufficiently accurate for the scale of our assessment.

Table S1. Settings of bed roughness length scale, mixing scheme, aerodynamic coefficient (E_A), and bulk transfer coefficient for latent heat flux (E_L) in calibration scenarios.

Scenario ID	bed roughness length scale (m)				mixing scheme	E_A	E_L
	channel	river	lagoon	ocean			
1	0.002	0.002	zone 1: 0.006 zone 2: 0.004 zone 3: 0.002	0.002	k-ε	linear	0.0013
2	0.005	0.002	zone 1: 0.015 zone 2: 0.010 zone 3: 0.005	0.002	k-ε	linear	0.0013
3	0.008	0.0033	zone 1: 0.024 zone 2: 0.016 zone 3: 0.008	0.002	k-ε	linear	0.0013
4	0.01	0.0033	zone 1: 0.030 zone 2: 0.020 zone 3: 0.010	0.002	k-ε	linear	0.0013
5	0.01	0.0033	zone 1: 0.030 zone 2: 0.020 zone 3: 0.010	0.002	Mellor-Yamada	linear	0.0013
6	0.01	0.0033	zone 1: 0.030 zone 2: 0.020 zone 3: 0.010	0.002	k-e	constant: 0.0013	0.0013
7	0.01	0.0033	zone 1: 0.030 zone 2: 0.020 zone 3: 0.010	0.002	k-e	constant: 0.0016	0.0013
8	0.01	0.0033	zone 1: 0.030 zone 2: 0.020 zone 3: 0.010	0.002	k-e	linear	0.0010
9	0.01	0.0033	zone 1: 0.030 zone 2: 0.020 zone 3: 0.010	0.002	k-e	linear	0.0016

Table S2. Model performance against the observed salinity and temperature, and the mean surface and bottom salinity difference during the wet season July-September (ΔSAL) of calibration scenarios. The values are the average from 6 sites.

Scenario ID	Salinity			Temperature			ΔSAL[*] (PSU)
	R	MEF	MAE	R	MEF	MAE	
1	0.9378	0.7256	2.1452	0.9459	0.7224	1.6722	3.5722
2	0.9381	0.7282	2.1291	0.9455	0.7223	1.6729	3.5782
3	0.9389	0.7332	2.0953	0.9459	0.7240	1.6710	3.6128
4	0.9398	0.7418	2.0433	0.9463	0.7293	1.6641	3.6175
5	0.9379	0.7255	2.1444	0.9457	0.7213	1.6761	3.5796

6	0.9360	0.7238	2.1647	0.9463	0.7280	1.6543	3.3086
7	0.9354	0.7221	2.1853	0.9428	0.7291	1.6553	3.2103
8	0.9361	0.7308	2.0772	0.949	0.8167	1.2934	3.7060
9	0.9316	0.7125	2.1874	0.9449	0.7211	1.7054	3.1605

*: observed mean salinity difference: 4.9653 PSU

2. Model validation

After the model calibration, we then used the selected numerical schemes and parameters to run all the selected years in history, and conducted comprehensive evaluation for all the simulation years, except for 1970 when the long term monitoring had not started yet. The evaluation focused on the salinity and water temperature of the surface and bottom water at the 6 monitoring stations within the estuary. Surface elevation records obtained from the gauged stations in the centre of Peel Inlet, provided by the Department of Transport of Western Australia, were also used to validate the modelled surface elevation in year 1990 (a modelled year before the Cut opening) and 1998 (a modelled year after the Cut opening).

In general, the model reproduced the temporal variations of salinity and temperature in both the surface and bottom well (Table S3). The mean regression coefficient r for the salinity from six monitoring sites is above 0.81, and for the water temperature is above 0.85 except in the year 1970, when a mean r of 0.72 was obtained, which may have been due to poor boundary forcing for this year. The model skill scores are generally higher than 0.61 for both salinity and temperature in all historical years, suggesting the model has captured the major features of the hydrodynamic response to the external forcing of tide and freshwater inputs.

Table S3. Model performance statistics at 6 monitoring stations (indicated as dark crosses in Figure 1) in the selected historical years.

Simulated year	Model performance of salinity (mean±std)			Model performance of temperature (mean±std)		
	r	*MAE*	*SS*	*r*	*MAE*	*SS*
1970	N/A	N/A	N/A	N/A	N/A	N/A
1978	0.95±0.02	3.24±0.54	0.85±0.02	0.72±0.16	1.89±1.10	0.58±0.23
1985	0.97±0.01	2.18±0.14	0.92±0.01	0.97±0.01	0.94±0.09	0.91±0.03
1990	0.95±0.02	2.60±0.18	0.87±0.05	0.94±0.02	1.31±0.10	0.81±0.04
1998	0.94±0.03	2.04±0.43	0.74±0.29	0.95±0.02	1.66±0.28	0.73±0.14
2004	0.81±0.20	3.78±1.47	0.61±0.34	0.85±0.05	1.60±0.28	0.62±0.15
2011	0.88±0.14	3.69±0.84	0.75±0.19	0.94±0.03	1.43±0.19	0.65±0.15
2016	0.87±0.11	3.54±0.81	0.74±0.23	0.96±0.01	1.13±0.08	0.88±0.02

3. Sensitivity to air temperature, benthic properties, and sea level variation

The sensitivities of modelled salinity and τ to air temperature, tide elevation, and bed roughness are shown in Figure S1. The changes in the air temperature of ±1 °C have minor effects on both the salinity and τ in both years of 1990 and 1998. The influence of air temperature on the hydrology was mostly through evaporation, and resulted in changes in salinity of less than 0.9 PSU, and 0.5 days changes in τ. Secondly, the changes in the mean tide elevation of ±0.15 m led to changes in salinity of up to 2.2 PSU and 8.4 days in τ. Thirdly, the bed friction also had a noteworthy impact on the salinity and τ by modifying the water movement and therefore benthic layer mixing at near-bed level. The presence of benthic vegetation was shown to affect salinity by up to 2.8 PSU in the Harvey Estuary, while a maximum change in the τ of 8.6 days was observed in the same location.

In summary, the modelled salinity and τ were affected by the changes in the sea level variation and bottom vegetation presence, but the effects of these environmental factors were still small when compared to that caused by the reduced flow over the past decade and the Cut-opening. For example, the maximum change in τ observed in the sensitivity test runs was 8.6 days, caused by the enhanced bottom roughness in the 1990 scenario, compared to the magnitude of 20-100 days caused by the reduced flow from 1970 to 2016. The maximum changes in the salinity observed in the sensitivity test runs was 2.8 PSU, caused by the reduction of tide level in the Harvey Estuary, compared to the magnitude of 10-30 PSU changes in the salinity caused by the reduced flows from 1970 to 2016. These results suggested the changes in the climate and the ocean connectivity are the major drivers of the hydrology of the Peel-Harvey Estuary.
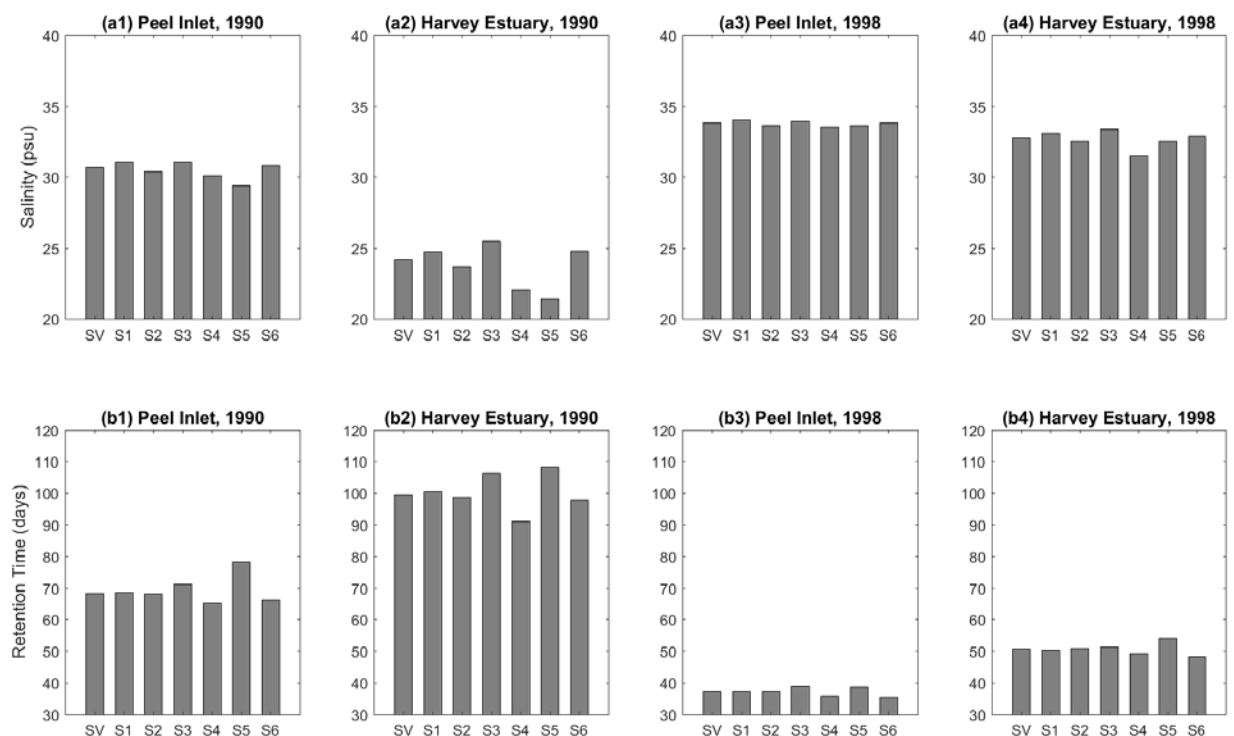

Figure S1. Sensitivity of the modelled annual-mean salinity and retention time in the Peel Inlet and Harvey Estuary. SV: standard scenario; S1: +1 degree in air temperature scenario; S2: -1 degree in air temperature

scenario; S3: +0.15 m in tide elevations scenario; S4: -0.15 m in tide elevations scenario; S5: +50% in bed roughness scenario; S6: -50% in bed roughness scenario.

Reference:

Arhonditsis, G. B., Perhar, G., Zhang, W., Massos, E., Shi, M., Das, A., 2008, Addressing equifinality and uncertainty in eutrophication models, Water Resour. Res., 44, W01420, doi:10.1029/2007WR005862.

Bahremand, A., Smedt, F.D. 2010, Predictive Analysis and Simulation Uncertainty of a Distributed Hydrological Model, Water Resources Management volume 24, 2869–2880.

Doherty, J., Johnston, J.M., 2003. Methodologies for Calibration and Predictive Analysis of a Watershed Model, Journal of the American Water Resources Association (JAWRA), 39(2):251-265.

Hipsey, M.R., Gal, G., Arhonditsis, G.B., Carey, C.C., Elliott, J.A., Frassl, M.A., Janse, J.H., de Mora, L., Robson, B.J. A system of metrics for the assessment and improvement of aquatic ecosystem models. Environ. Model. Soft 2020, 128, 104697.

Muleta, M.K., Nicklow, J.W., 2004, Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model, J. Hydrol., 306, 127–145.

Wu, J. (1982). Wind-stress coefficients over sea surface from breeze to hurricane. J. Geophys. Res., 87 (C12) 9704–9706.