Paper Review of Learning functionals via LSTM neural networks for predicting vessel dynamics in extreme sea states

José del Águila Ferrandis, Michael Triantafyllou, Chryssostomos Chryssostomidis,
George Karniadakis

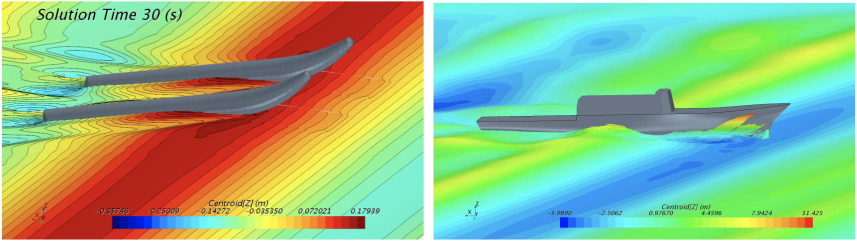
Department of Mechanical Engineering, MIT
DAM, Brown

January 24, 2020

Problem Setup

- Goal: Predict motions of vessels in extreme sea states
- Involves computing complex nonlinear wave-body interactions, hence taxing heavily computational resources
- Use traditional method to generate dataset
- Use supervised learning methods to save computation time when making predictions.
- Input: Sea surface elevation at specific locations
- Output: Vessel Motion in 3 DOF(heave, pitch and roll)

Problem Setup



(a) (b)

Figure 1: *Snapshots of the URANS simulations*; the color scale (meters) represents sea surface elevation. (a) Catamaran sailing in regular 5th-order Stokes Waves. This constitutes a relatively easy condition to simulate using potential flow methods. (b) Notional DTMB battleship sailing in World Meteorological Organization (WMO) sea state 8 at Froude number 0.4. Meshes for this case involve more than 20 million finite volume cells and require several days to compute on a parallel computer with 300 cores. (See video for DTMB sailing through sea state 8 in supplementary materials).

Sea Surface Elevation

Ocean waves used to induce motions in the vessels are reconstructed from experimental sea spectra (see fig. 2) that characterize the stochastic process of sea surface elevations. At a particular spatial location, let $\zeta(t)$ be the sea surface elevation as a function of time. Then, this time signal can be defined as a sum of sinusoidal waves with random phases ε_i between $-\pi$ and π sampled from a uniform distribution, and incommensurate frequencies ω_i spanning the frequency range of the spectrum, i.e.,

$$\zeta(t) = \sum_{i=1}^{n} a_i \cos\left(\omega_i t + \varepsilon_i\right), \tag{4}$$

which is a Gaussian probability density function as n becomes large in accordance with the central limit theorem. The wave amplitude for a given frequency is obtained from the following relation,

$$\frac{1}{2} a_i^2 \cong S\left(\omega_i\right) \Delta\omega, \tag{5}$$

where $S\left(\omega\right)$ is the modified Pierson-Moskowitz spectrum [28] with T_1 as the mean wave period:

$$\frac{S(\omega)}{H_s^2 T_1} = \frac{0.11}{2\pi} \left(\frac{\omega T_1}{2\pi}\right)^{-5} \exp\left[-0.44\left(\frac{\omega T_1}{2\pi}\right)^{-4}\right], \tag{6}$$

Vessel Motion

The data set is obtained from a viscous Volume-of-Fluid (VOF) URANS solver (STAR-CCM+). The equations solved are the averaged continuity and momentum equations for incompressible fluids:

$$\frac{\partial(\rho \overline{u}_i)}{\partial x_i} = 0, \tag{1}$$

$$\frac{\partial(\rho \overline{u}_i)}{\partial t} + \frac{\partial}{\partial x_j}(\rho \overline{u}_i \overline{u}_j + \rho \overline{u_i' u_j'}) = \frac{\partial \overline{p}}{\partial x_i} + \frac{\partial \overline{\tau}_{ij}}{\partial x_j}, \tag{2}$$

$$\overline{\tau}_{ij} = \mu \left(\frac{\partial \overline{u}_i}{\partial x_j} + \frac{\partial \overline{u}_j}{\partial x_i} \right), \tag{3}$$

where $\overline{\tau}_{ij}$, in eq. (2), are the components of the averaged viscous force tensor, $\overline{p}$ is the averaged pressure and $\overline{u}$ are the Cartesian components of the averaged velocity. In eq. (2), $\overline{u_i' u_j'}$ are the Reynolds stresses, ρ the fluid density, and μ the dynamic viscosity.

Functional Approximation

The results presented in [1] can be summarized as follows. Given very mild conditions, a functional defined on a compact set in $C[a,b]$ or $L^p[a,b]$ (spaces of infinite dimensions) can be approximated arbitrarily well by a neural network with just one hidden layer. Particularly, given U a compact set in $C[a,b]$, σ (a bounded sigmoidal function) and $\mathcal{F}$ a continuous functional defined on U, then $\forall u \in U, \mathcal{F}(u)$ can be approximated by

$$\mathcal{F}(u) = \sum_{i=1}^{N} c_i \sigma \left(\sum_{j=0}^{m} \xi_{i,j} u \left(x_j \right) + \theta_i \right). \tag{9}$$

In the above expression, c_i, ξ_{ij}, θ_i are real numbers and $u(x_j)$ is the value that u takes at x_j.

Operator Approximation

1. If $u \in U$, then $u|_{\Gamma_{\alpha,a}} \in U, \forall \alpha \in \mathbf{R}^n, a > 0$.
2. $\forall\ \alpha\ \in\ \mathbf{R}^n, a\ >\ 0, U_{\alpha,a}$ is a compact set in $C_V\left(\prod_{k=1}^n [\alpha_k - a_k, \alpha_k + a_k]\right)$ or a compact set in $L_V^p\left(\prod_{k=1}^n [\alpha_k - a_k, \alpha_k + a_k]\right)$, where V stands for $\mathbf{R}^{q_1}$.
3. Then, if we let $(Gu)(\alpha) = \left((Gu)_1(\alpha), \ldots, (Gu)_{q_2}(\alpha)\right)$, consequently each $(Gu)_j(\alpha)$ will be a continuous functional defined over $U_{\alpha,a}$, with the corresponding topology in $C_V\left(\prod_{k=1}^n [\alpha_k - a_k, \alpha_k + a_k]\right.$ or $L_V^p\left(\prod_{k=1}^n [\alpha_k - a_k, \alpha_k + a_k]\right)$.

We consider that a map G from X_1 to X_2 has *approximately finite memory* if $\forall \epsilon > 0$ there is $a > 0$ such that:

$$\left|(Gu)_j(\alpha) - \left(GW_{\alpha,a}u\right)_j(\alpha)\right| < \epsilon, \quad j = 1, \ldots, q_2 \quad \forall \alpha \in \mathbf{R}^n, u \in U.$$

Theorem: If U and G satisfy all the assumptions (1-3) made previously, and G is of *approximately finite memory*, then $\forall \epsilon > 0$, $\exists a > 0$, m a positive integer, $(m+1)^n$ points in $\prod_{k=1}^n [\alpha_k - a_k, \alpha_k + a_k]$, N a positive integer, constants $c_i(G, \alpha, a)$ that only depend on G, α, a, and $q_2 \times (m+1)^n$ - vectors ξ_i, $i = 1, \ldots, N$, such that:

$$\left|(Gu)_j(\alpha) - \sum_{i=1}^{N} c_i(G, \alpha, a)\sigma\left(\bar{\xi}_i \cdot \bar{u}_{q_1,n,m} + \theta_i\right)\right| < \epsilon, \quad j = 1, 2, \ldots, q_2.$$

To conclude, we would like to place emphasis on the assumption of *approximately finite memory*. This assumption provides the blocks to build the functional approximation as a sum of functionals defined in the subsets given by the window operator, previously defined. During the empirical analysis, we benchmark against a case for which we hope that *approximately finite memory* will allow representing the functional arbitrarily well from the subsets given by the window operator.

Operator Approximation

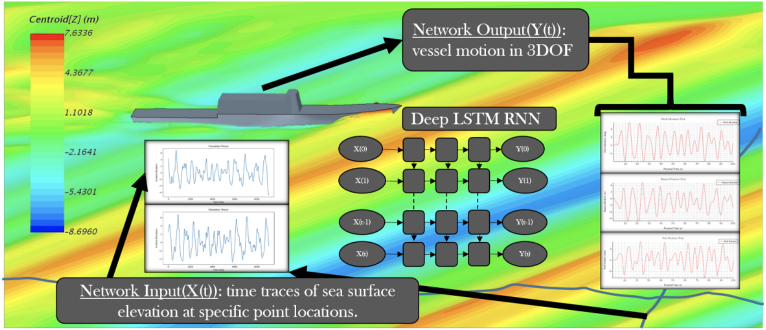
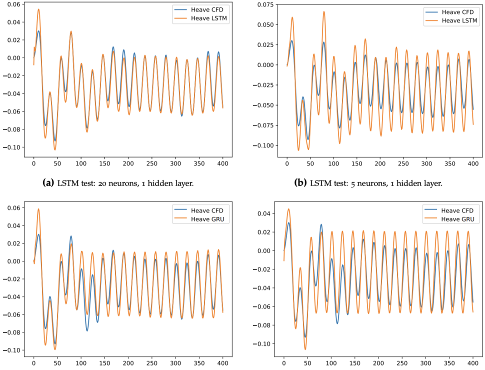


Figure 3: *Schematic of the physical problem simulated and inputs and outputs of the deep LSTM RNN.* The inputs for training are sea surface elevations in the form of time series, while the corresponding outputs are the vessel motions. Sea surface elevations are recorded at specific point locations that can be chosen from lines over the free surface. Vessel motions in the training cases are obtained from an URANS solver. Shown here as inputs ($X(t)$) are two unseen surface elevations, which serve as test cases in our simulation example for the DTMB vessel.

Catamaran Vessel



(a) LSTM test: 20 neurons, 1 hidden layer.

(b) LSTM test: 5 neurons, 1 hidden layer.

(c) GRU test: 20 neurons, 1 hidden layer.

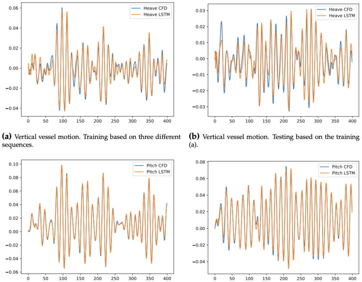
(d) GRU test: 5 neurons, 1 hidden layer.

Figure 4: *Comparison of LSTM and GRU for a 5th-order Stokes regular wave (wave amplitude is* $0.15m$*) for the catamaran vessel.* The data-set is composed of 5 waves of varying amplitude. The first 4 waves are used as training cases (for 20000 steps) while the last wave (shown here) is used for testing. Each time step corresponds to $\Delta t = 0.0625s$.

Loss: $L = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$

Catamaran Vessel

- The pitch motion is modelled better than the heave motion.
- Shallow network is adequate in approximating the two DOF catamaran vessel subject to irregular waves.



(a) Vertical vessel motion. Training based on three different sequences.

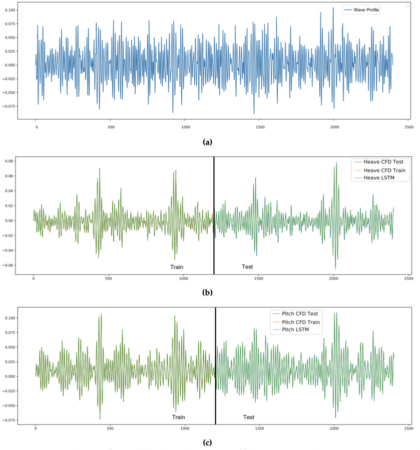
(b) Vertical vessel motion. Testing based on the training of (a).

(c) Angular vessel motion. Training based on three different sequences.

(d) Angular vessel motion. Testing based on the training of (c).

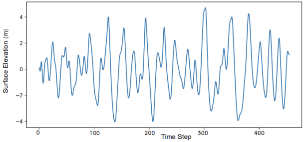
Figure 6: *LSTM network (20 neurons, 1 layer) for the catamaran vessel subject to irregular waves.* The left column (a, c) shows the vertical and angular vessel motions after training with three different sea state realizations; one such realization is shown in fig. 6a. The right column (b, d) shows the vertical and angular vessel motions for testing given the inputs indicated in fig. 6b. Each time step corresponds to $\Delta t = 0.0625s$.

- The model is capable of making long time prediction.

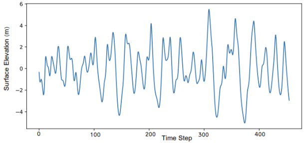


DTBM Vessel

- Add one dimension in input: Use profile of the wave in the longitudinal direction and transversal direction.
- Add one dimension in output: Model heave, pitch and roll.



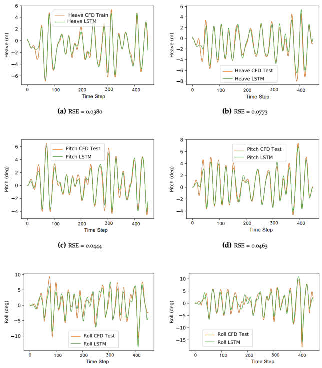
(a) Profile of the wave in the longitudinal direction corresponding to test case one.



(b) Profile of the wave in the transversal direction corresponding to test case one.

DTBM Vessel

- Loss $L = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - mean(\hat{Y}_i))^2}$
- Result is worse than the previous one.



Conclusion

- Good prediction results of the motion of vessel can be obtained in a much shorter time than traditional methods, without sacrifice of too much accuracy.
- Maybe(?) the link between theory and experiment need more careful explanation. Can try DeepONet in the future.

The End