Review for “A dynamic informed deep learning method for future estimation of laboratory stick-slip” by Yue et al.

Line 11: As far as I know, there is no actual successful application to real-world slow slip events.

Line 22: Natural fault activity is not quasi-periodic.

Line 30: Remove “firstly”.

Lines 32-33: Lubbers et al. (2018) worked on laboratory earthquake: this needs to be specified because it is otherwise misleading.

Line 41: “discovered” → “showed”.

Line 66: Remove “shear”.

Lines 67-68: “and the equipment is used to record the changes in system properties such as pressure recorded during the shear process” is too vague. What observables are you using?

Line 69: “modified”: How did you modify the model?

Line 72: To claim that the system is quasi-periodic you should run a frequency analysis. The system is likely not quasi-periodic in a strict sense, and it’s more appropriate to say that it is cyclic.

Line 74: “robust” → “clear”

Figure 1: caption: remove “fault slip”.

Section 2.2.1 needs references. You are largely taking the description from existing literature, and it needs to be cited.

Line 82: “fits” seems the inappropriate verb here. Maybe “described by”?

Line 83: Eq. 1 is wrong. The shear stress at time $t+1$ is not a function of solely the shear stress at time t.

Line 86: The mapping g is not the mapping of the Koopman operator. The wording is unclear. From eq. 2, the Koopman operator maps one step ahead the system described by the function g applied to the variables that describe the system.

Lines 91-92: Technically, you need infinite modes to properly describe with a linear operator a generic non-linear dynamic.

Line 94: “are the eigenvector” and “are the eigenvalues”.

Line 96: Is j the time step?

Line 97: “of the k-th dynamic mode”.

Line 98: “we’re aware of” → “we know”.

Line 98: “the mapping function”.

Line 99: “the linear operator”.

Line 100: You are now talking about control systems, but you do not have a controlled system in eqs. 1-3. It seems like you copy-pasted from existing literature, but it is not an accurate description of what you are actually doing.

Line 106: What are “majority observations”?

Line 118: What “As shown in Figure 3”? The sentence is incomplete. Figure 3 does not show g and K. It is unclear what is the link with the previous sentence.

Line 121: “Lorentz” → “Lorenz”. Furthermore, it needs a citation.

Eqs. 8-9: Is the running index i going to d instead of n? What is n? Furthermore, I think there is an error about the j running index. Shouldn’t it go to L instead of h?

Line 159: “to address” → “that addresses”.

Line 161: “earth” → “Earth”.

Eqs. 10-12: Either you describe what σ, ω and b are, or you do not show the equations and just refer to the literature.

Lines 192-194: You should show the results (for example, with a figure) to confirm that an embedding dimension of 100 is suggested by the method that you mention (Cao, 1997). The optimal time delay embedding is likely not 1, but you are using a value of 1 in order to build the Hankel matrix. The fact that the dimension of the Koopman operator is set to 10 further suggests that the dimensionality of the system is likely not of the order of 100, so it is difficult to understand how Cao’s method could provide such a high minimal embedding dimension. The reason is likely because there is a lot of redundant information with the small time delay embedding of 1.

Line 197: How are they adapted? Just by trial and error? Or with some consistent procedure?

Lines 197-199: What is a multi-step trick?

Tab. 1: adopted instead of adapted? These parameters are fixed, not adapting.

Eq. 14: Is the mean the mean over the whole time series or in the window used to make the forecast? Using the whole time series mean seems wrong to me: why should the forecast tend to the mean of the whole time series when using a specific window that has a subset of the whole time series?

Eq. 17: The notation is quite odd. What you name “slip” or “stress” it’s actually just the index, so it is a time. When writing “slip_center_index” what is typically understood is that that is the value taken by the slip at the time index given by the index. Furthermore, “the central moment of a slip behaviour” is not clear. What is a “slip behaviour”?

Eq. 18: The fraction and parenthesis are likely wrong: check the consistency of your equation.

Lines 237-239: SVD is not applied to validate the effectiveness of the dynamic modelling.

Lines 243-244: Is it really quasi-periodic?

Lines 249-253: The wording is very cryptic. What do you mean by “a single observation dimension”? Furthermore, a dynamical system will follow a trajectory that defines the attractor. How is it possible to have an unstable stick-slip system where the data does not exactly follow the attractor trajectory during the experiment?

Line 257: Experiment p5198 does not have a stable mode at 9 s.

Lines 259-260: Why a zero-frequency mode with low amplitude is representative of a strong time-varying component? What does it mean "strong time-varying component"?

Line 278: Why only two experiment settings?

Line 284: You claim that the HKAE results are clearly better when using the DeltaT evaluation metric. OK, it is better, but the R2 is most of the time negative, and, in the best situation, slightly positive. This seems to be a very poor result for all the architectures. Even if for the HKAE it is better, it is still a poor result.

Line 285-287: The sentence is unclear. It is surprising that the RMSE and R2 metrics are worse for the HKAE in the first few forecasted steps. This would suggest that for short term predictions LSTM and TCN are better. But the dynamics that you are extracting should degrade with longer lead prediction times. Something that instead is not really observed.

Lines 289-294: Which metric? You tested more than one. The setting 20-10 s is shown in Fig. 8b, not 8a. Finally, how can you claim that the lead prediction can cover 1-2 complete cycles? When we look at Fig. 8b, the DeltaT metric shows that many points are not on the diagonal, meaning that the prediction is wrong in many cases regarding the timing of the events. This class misrepresentation (DeltaT small or large are the two classes) indicates that your lead prediction cannot cover multiple cycles.

Line 296: If the slip events were periodic, it would be very easy to predict them, and you should have zero error. The events are not periodic.

Lines 318-322: There is confusion on what you mean by underestimation. For example, if you say that the minimum value is underestimated, it means that the predicted shear stress is smaller than the actual minimum shear stress. But the figure shows that the predicted minimum shear stress is higher. What you mean is that the stress drop is underestimated. Please be consistent and describe properly the results.

Line 326: "moment of " → "instant for".

Line 343-346: I disagree with the way you are reading the figure. In Exp. 4679 LSTM clearly has a better RMSE for most of the lead time between 3 and 10 s (non-grey area). It is not even clear how you are drawing those lines. If you look at just the points (and not the interpolated line, which is not an actual result), LSTM is in fact comparable to HKAE until 4.5 s, and then always better or with same results. For Exp. 4581 the RMSE is basically the same for all algorithms. Finally, for Exp. 5198 the RMSE of LSTM is the best until about 7 s, where it becomes like the one of HKAE, and only slightly worse after 9 s. But what does it mean to be able to have a slightly better forecast at 9-10 s lead time? Based on your results, if I had to pick an algorithm to make earthquake forecasts, I would use LSTM: it clearly outperforms HKAE in the forecast of the immediate next steps, which is what matters the most. For long-term forecasts (like 9-10 s, which is a full cycle ahead), one can very likely use purely statistical methods and obtain similar results.

Lines 369-370: The two "stable" dynamics clearly show an amplification over time, which makes them not as stable as you depict them. From Fig. 11b it is not clear what they represent. What do you mean with "input"? These are modes obtained after processing the input data in some way.

Lines 387-388: This is not true. For short-term predictions HKAE seems to perform poorly compared to other algorithms. If we look at either Fig. 8 or Fig. 10, we see that for the first few time-steps of forecast the HKAE is consistently not the best, with LSTM being often the best, followed by TCN.

Lines 392-394: This is not true. On many occasions HKAE was not showing superior performance in terms of prediction accuracy.

Figure 2: The caption is too synthetic. You need to explain what the panels represent. You need to add letters to the various panels. All the panels need labels for the axes. You are using the Lorenz63 system time series, but never mentioned it in the text. If you are using the figure taken from another publication, you need to state it and you need to have the permission to reproduce it here.

Figure 3: I find the usage of the Lorenz system as an example not appropriate. The Lorenz system is not the focus of this work, and you could (and, in my opinion, should) use the dynamical system that mimics the laboratory earthquakes instead.

Figure 4: My understanding is that the output of the Koopman operator applied to $g(x_t)$ is $g(x_{t+1})$, so this should be the reconstruction in the embedding space. From your figure this reconstruction comes after the application of $g^{-1}(\cdot)$, but this should give you a signal back in the original input space (and indeed you have an arrow coming back down as "One-step Evolution"). I think that this figure is inaccurate and needs to be corrected.

Figure 8: In the main text you start from experiment 5198, and this should be the first to appear in the figure (i.e., on the left). Furthermore, you have also shown ODE simulations earlier on. You should show the results on the simulations as well. They are periodic, so you should be able to recover them perfectly. If not, this would cast some doubts on the procedure.