The manuscript entitled 'Coupled estimation of incoherent inertia gravity wave field and turbulent balanced motions via modal decomposition' investigates the identification problem of inertia gravity waves interacting with turbulent jets using POD and extended POD methods. The authors test their method on a rotating shallow water model. Given the two-point statistics of the turbulent field and the incoming wave, and assuming time scale and magnitude separations and strong correlation between the jet and the wave, they introduce a methodology that can distinguish the wave and the jet components from a single observation of the sea surface height. While I believe these assumptions significantly limit the method's applicability to more realistic scenarios, I find the methodology itself quite innovative. I appreciate that the authors managed to construct a model problem satisfying these assumptions.

My primary concern with the manuscript is its readability; the dense mathematical details required multiple readings to fully grasp. The text is mathematically quite involved, and I think, it can be simplified. For instance, the authors initially permit the complex amplitude of the inertia wave to vary over time, but later disregard this time dependency. I wonder whether it would be possible to reach the final resolvent equation by imposing classical Reynolds decomposition and linearization about the mean. Furthermore, the inconsistent use of terminology—particularly terms that carry different meanings across disciplines—complicates the understanding of the methodology. For instance, using an expectation operator, they decompose the jet into mean and fluctuating parts. But, the same expectation operator separates the wave into its coherent and incoherent parts.

I have listed specific areas that require further clarification; however, beyond addressing these, I recommend that the authors re-evaluate the entire text to enhance overall clarity. I believe the study merits publication, though its impact could be significantly enhanced with additional efforts to improve readability.

Details:

lines 107 and 143: I don't see why the authors define \tilde{q}_\omega as a function of time, which is assumed to be constant later on anyway.

lines 148-149: The authors do not make any assumption about the frequency band they choose. How do they conclude about the fact that the resolvent operator is approximately constant? I would rephrase this sentence as '... can also be interpreted as assuming that the resolvent operator is approximately constant'

line 151: Why do the authors define this decomposition as coherent and incoherent while it is a Reynolds decomposition?

Eq. (11): If I understand it correctly, the first term in this equation is not fluctuating since the expectation operator (which amounts time or ensemble averaging, I guess) applies to the bilinear operator. So \tilde{q}'_omega is actually fluctuating about this term. If that is the case, I find it confusing that a prime term has a nonzero mean. Regarding the first use of the prime on line 140, I would relate it to quantities with zero mean.

lines 167-170: This spectral broadening effect applies to any term in eq. (11). But I think the authors particularly think of the term B(q'_jet,\tilde{q}'_omega). If that's the case, I think it would help if they explicitly stated that such as 'Taking B(q'_jet,\tilde{q}'_omega) for instance, ...'. It took me a while to figure out which incoherent component they were mentioning on line 170.

Eq. (13): How do the authors come up with this norm? Is it common in oceanography?

Algorithm 1 – Training stage, last equation: This assumes a strong correlation with h_omega and q_jet, as stated later in the text. Is not it a very limiting assumption? How realistic is it?

Eq. (22): How is this minimization achieved? The coefficients at the training stage are computed using the full vector q_jet, while in the minimization problem, only h is used. Does not this potentially cause a uniqueness issue? How do we know that the result of the minimization is unique?

Section 4.1 - A schematic showing the domain, discretization, the forcing and how the wave is introduced would be very helpful to visualize the test case.

lines 359-360: How are the EPOD modes calculated, applying POD to [q_jet, q_\omega] or first applying to POD to q_jet and then using the coefficients a_n(t) to reorganize q_\omega? How much error would be introduced by the former?

line 366: 'jet POD mode wave BBPOD mode' → 'jet POD mode and the wave BBPOD mode'

lines 371-375: Isn't it actually possible to calculate the single scattering term in eq. (19) and compare it against the EPOD modes?

Figure 10. The orange line is not mentioned in the caption.

lines 443-444: Why does the method fail for the W2 case?

lines 450-451: Is there a way to predict where the cut off the modes a priori?