Thank you very much for your time and your comments and suggestions. We appreciate your help to improve our manuscript. The following is our point-to-point replies:

1.p. 3, ln. 84/85: Is the pressure-gradient discretization in FV3 fourth-order accurate including the metric terms related to the coordinate transformation?
Yes, the pressure-gradient discretization in FV3 with fourth-order accuracy includes the metric terms related to the coordinate transformation. The discretization in FV3 uses a finite-volume integration method to compute the pressure gradient force in general vertical coordinates (Lin 1999). It is based on fundamental physical principles in the discrete physical space, rather than on the common approach of transforming analytically the pressure gradient terms in differential form from the vertical physical (i.e., height or pressure) coordinate to one following the bottom topography. This is analogous to the finite-volume of the advective process in which the accuracy depends on the assumed subgrid distribution of the advected constituent.
2. p. 4, ln. 121: "but that did not fix the model stability issue for all eight crashed cases": That means, it did help for some of the cases, or for none of them? Please clarify.
It helped for some of the cases. The line 121 is revised to emphasize that some cases can finish 16-day.

3. p. 5, ln. 124/125: As the instability was related to a collapse of the surface layer thickness, did the changes of the Rayleigh damping / sponge layer beneath the model top have any impact at all? If yes, how can this be explained?
It is a good question. You are right that the changes of Rayleigh damping/sponge layer near the model top cannot solve the issue related to the collapse of the surface layer thickness. We believed that the occurrence of surface layer thickness is a rare and extreme event in the model. Applying stronger damping near the model top likely change the forecasts with model integration and we are just "lucky" to avoid the crash.

4. p. 5, caption of Fig. 1: From which model level are the wind fields taken? And the unit for the reference arrow is missing (hopefully m/s).
The wind fields are at the model lowest layer with unit (m/s). It is added in Fig. 1 caption.

5. p. 6, Eq. (1): What is m? I suppose mass, but this should be indicated here.
Yes, m is mass, it is added in the manuscript
6. p. 7, ln. 218: "close to 0 200s ..." please write "zero".
Revised.
7. p. 8, Fig. 4 and corresponding caption: please indicate the unit of virtual potential temperature and explain the meaning of its values. Is it a deviation from a reference state? In addition, it appears quite odd that the height thicknesses in panel (e) are generally negative. This may be a model-internal sign convention, but a thickness is usually a positive-definite quantity, and it would be better to adjust the figure accordingly.
The definition of the virtual potential temperature in FV3 is added and its unit is indicated in figure 4. height thinkness is revised to positive value in the figure 4(e).
8. p. 8, ln. 234–237: The sentence starting with "Note that the update" is hard to understand and should be formulated more clearly.
This sentence was rewritten
9. p. 12, ln. 329: $\frac{dq}{dx}$ 0 – analogous in y-direction? Moreover, I would expect that zero slope means $\frac{dh}{dx} = 0$. Please clarify.

$\frac{dq}{dx}$ should be $\frac{dq}{dz}$. It is corrected.

10. p.15, ln. 391: Is replacing upper and lower boundary conditions at the same time needed for consistency, or would it be sufficient to use zero-gradient for the lower bc's?

These two kinds of BCs make have little different impact on the model upper BCs. But we replaced both lower and upper boundary for consistency.

11. p.15, ln. 392: "forecasting"!"forecasts"

Revised to forecasts

12. p. 15, ln. 409: "zerio"!"zero"

Corrected

13. p. 15, ln. 411: "Planetary Boundary Lateral"!"... Layer"

Corrected

14. p. 17, ln. 469: "value of 2 to 6 can ..." insert "meters"

"meters" are added.