Supplementary Information

Supplementary Tables

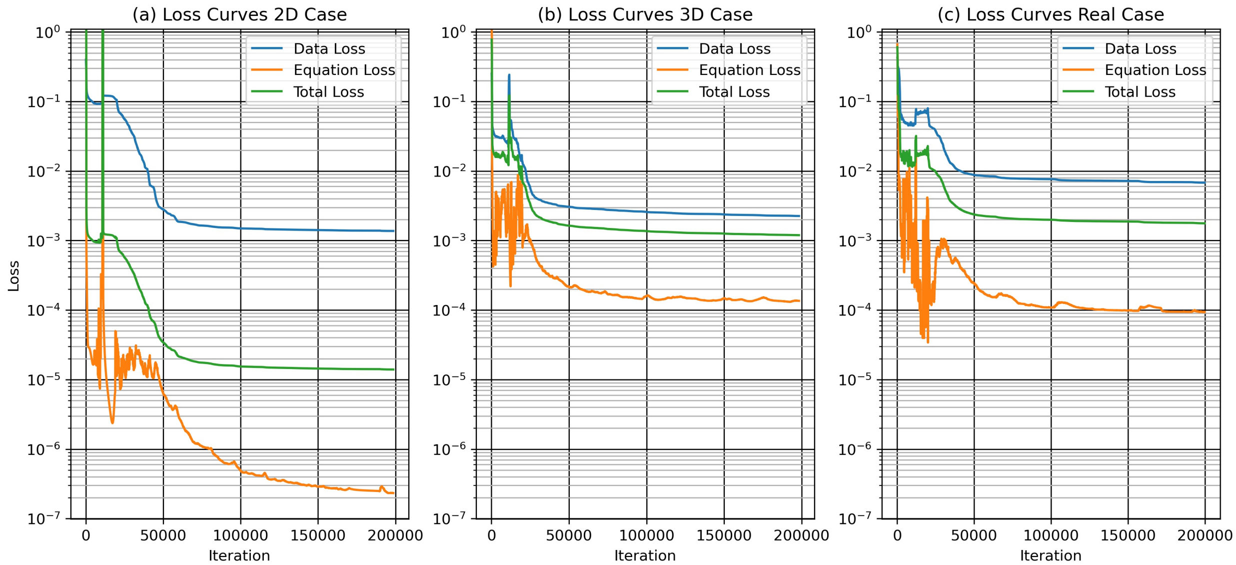
Data Point Density Ratio	Collocation Density Ratio	Total RMSE	Inner Core RMSE	Maximum Wind Value
5.0	0.2	2.20	5.90	58.54
5.0	1.0	2.20	5.96	55.84
5.0	5.0	2.24	5.91	53.78
1.0	0.2	2.20	6.13	55.31
1.0	1.0	2.22	6.66	52.97
1.0	5.0	2.26	6.72	52.05
0.2	0.2	2.27	8.76	53.12
0.2	1.0	2.29	7.19	50.42
0.2	5.0	2.35	7.70	48.26

Supplementary Table 1: PINN results from the 3D Case when trained with different densities of data and collocation points within the inner core and outside the inner core. The inner core is defined as the inner 100km cylinder around the storm center extending in the full vertical direction. In all cases, 4,158 data points and 10,000 collocation points are used. In the first two columns, a density ratio of 5.0 means that the density of points in the inner core is 5 times higher than the density of points outside the inner core. The total RMSE is the RMSE of the PINN output against the SHiELD output throughout the entire domain. The inner core RMSE is just the RMSE in the inner 100km of the storm. The maximum winds of the storm were 69 m/s, so in this table the higher max wind values are generally better. All results are averaged over an ensemble of 5 independently trained samples. These results highlight that the best results were obtained with higher data point density ratios (5.0) and lower collocation point density ratios (0.2).

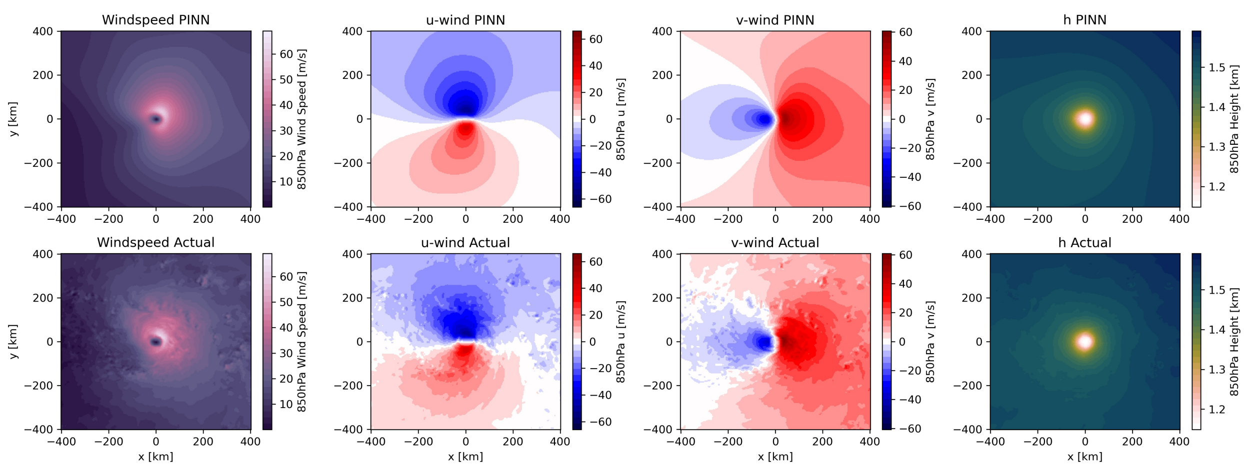
Data Points	Collocation Points	Hidden Layers	Layer Sizes	Training Time (minutes)
4158	10000	4	50	48
4158	1000	4	50	35
4158	5000	4	50	39
4158	20000	4	50	48
4158	100000	4	50	118
41581	10000	4	50	50
415807	10000	4	50	76
4158066	10000	4	50	351
4158	10000	4	100	50
4158	10000	8	50	62
4158	10000	8	100	77

Supplementary Table 2: Training times from PINN training for different numbers of data points and collocation points and network structures. The first line in the table is the baseline model, then the next block shows how training time changes with collocation points, the next block for changing data points, and the final block for changing network structure. We see the training times are very similar and generally scale well, except for $O(10^5)$ collocation points and $O(10^6)$ data points. Note all PINNs used a single NVIDIA A100 GPU core for training.

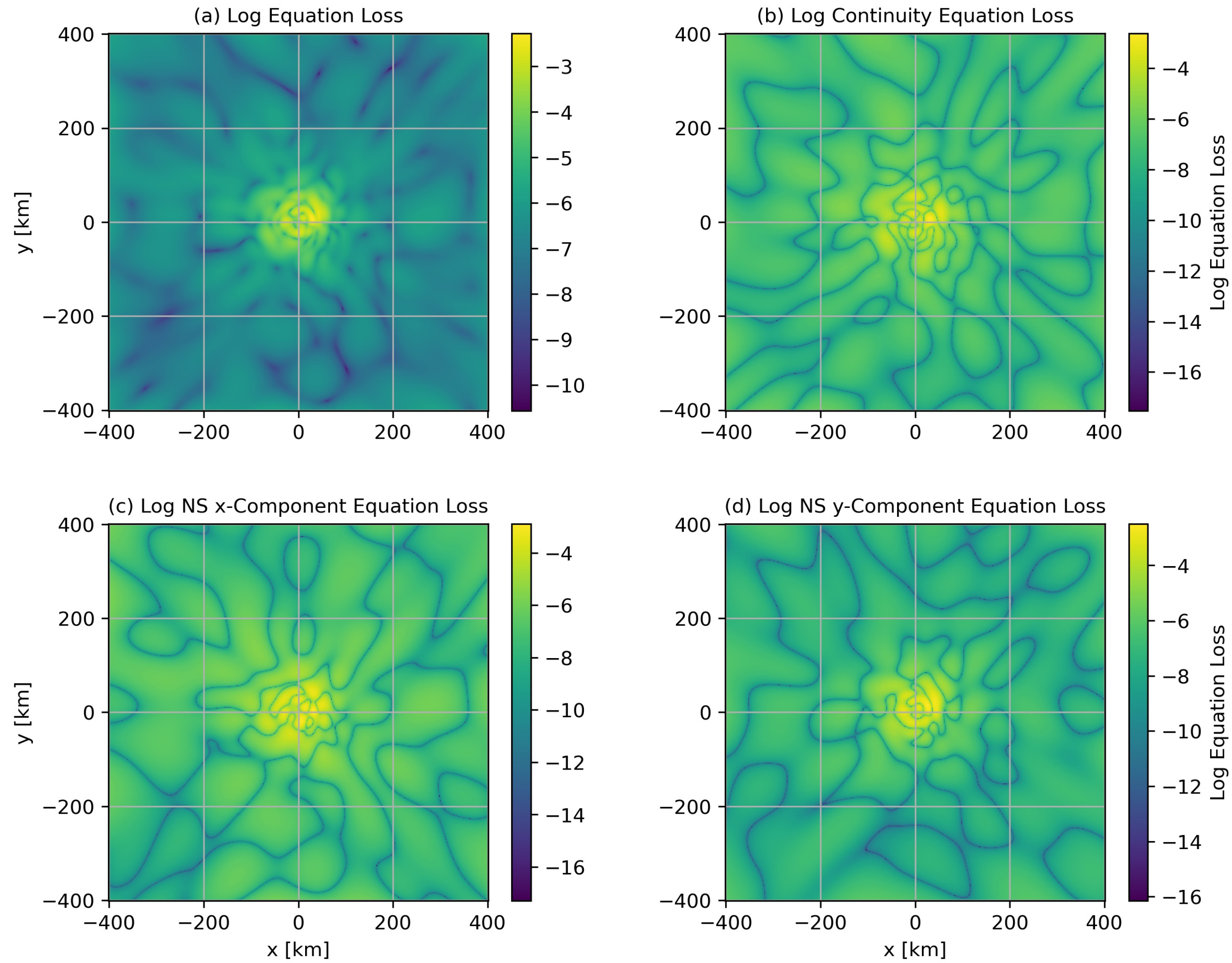
Supplementary Figures



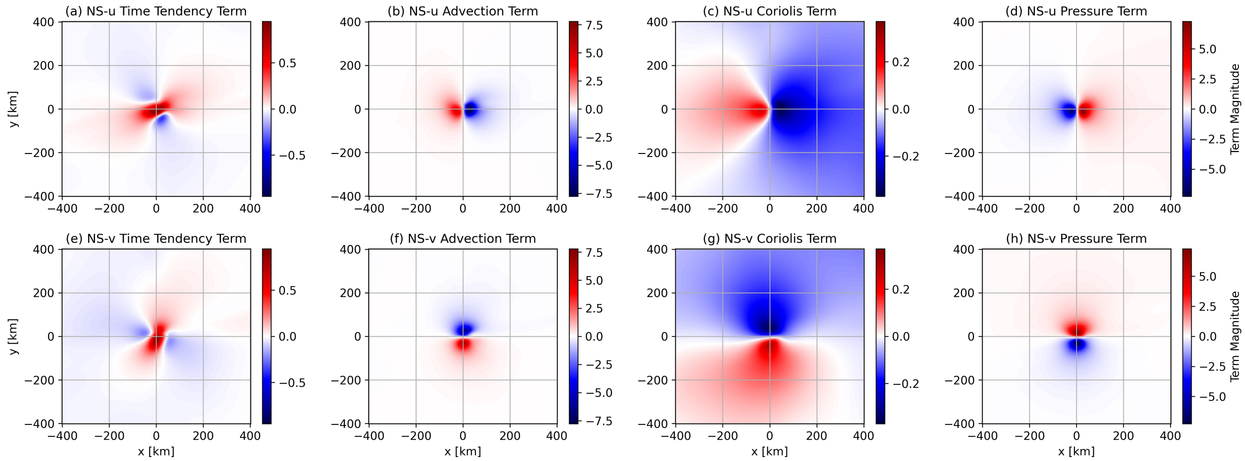
Supplementary Figure 1: **PINN loss curves.** Data loss, equation loss, and total loss curves from training for the PINN used in the (a) 2D case, (b) 3D case, and (c) real case. Referring to the methods section, the data loss follows equation 7; the equation loss follows equation 8; the total loss follows equation 9.



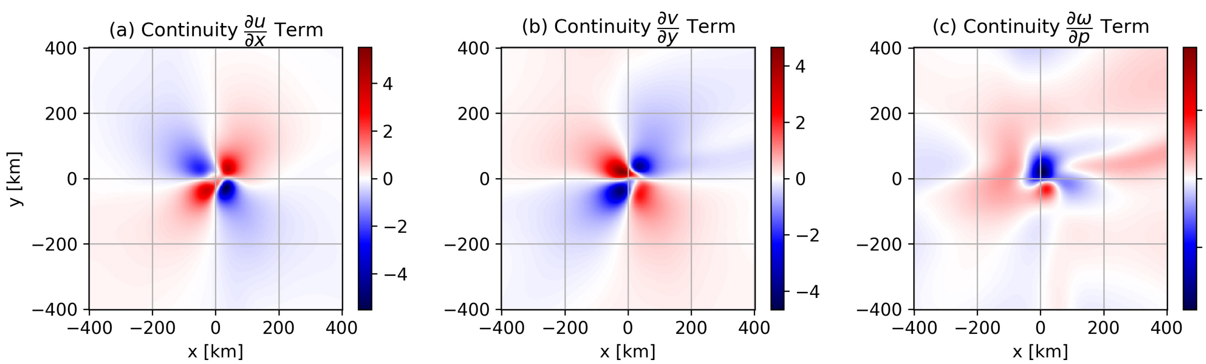
Supplementary Figure 2: **PINN output compared to target SHiELD.** From left to right we see the wind speed, the u-component of the wind, the v-component of the wind, and the geopotential height. The top row is the PINN output after training and the bottom row is the target output from SHiELD. This is hour 60 of the SHiELD forecast of Hurricane Ida initialized at Aug 27, 2021 00z.



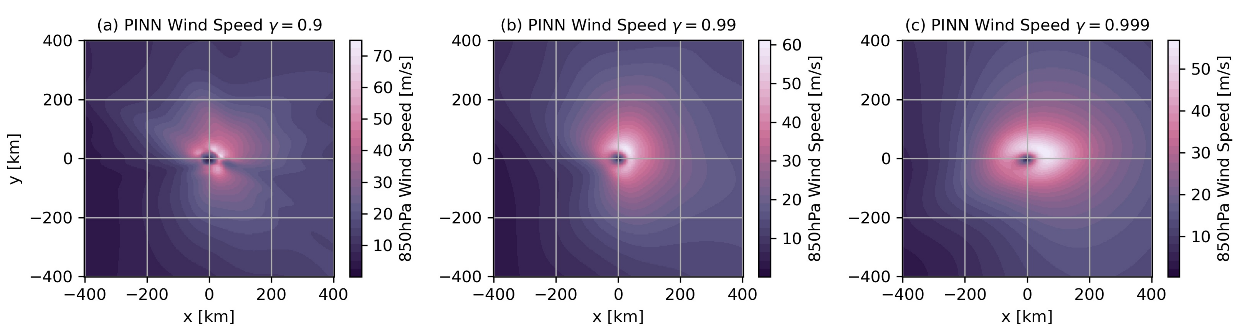
Supplementary Figure 3: **The log of the equation loss grids in the 2D case and contributions from each equation.** (a) Log equation loss, defined as the log of equation 8. (b) The log continuity equation loss (i.e. the squared equation residual from equation 5). (c) x-component of the Navier-Stokes log equation loss (i.e. squared equation residual from equation 3). (d) y-component of the Navier-Stokes log equation loss (i.e. squared equation residual from equation 4).



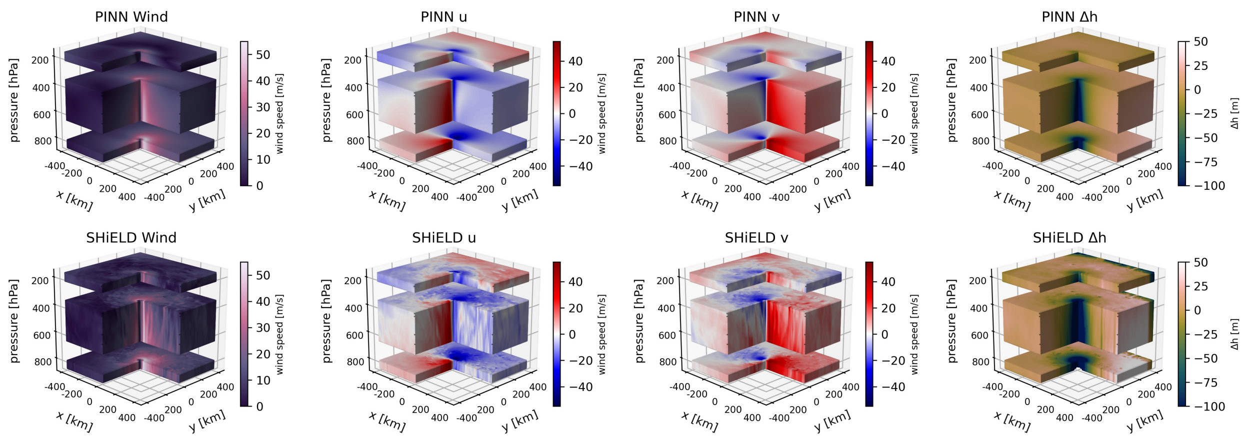
Supplementary Figure 4: **Navier-Stokes equation non-dimensionalized term magnitudes for the PINN trained on the storm in the 3D case.** (a-d) are for the x-component of the Navier-Stokes equations and (e-h) are for the y-component. From left to right, we have the time tendency term, the advection term, the Coriolis term, and the pressure term.



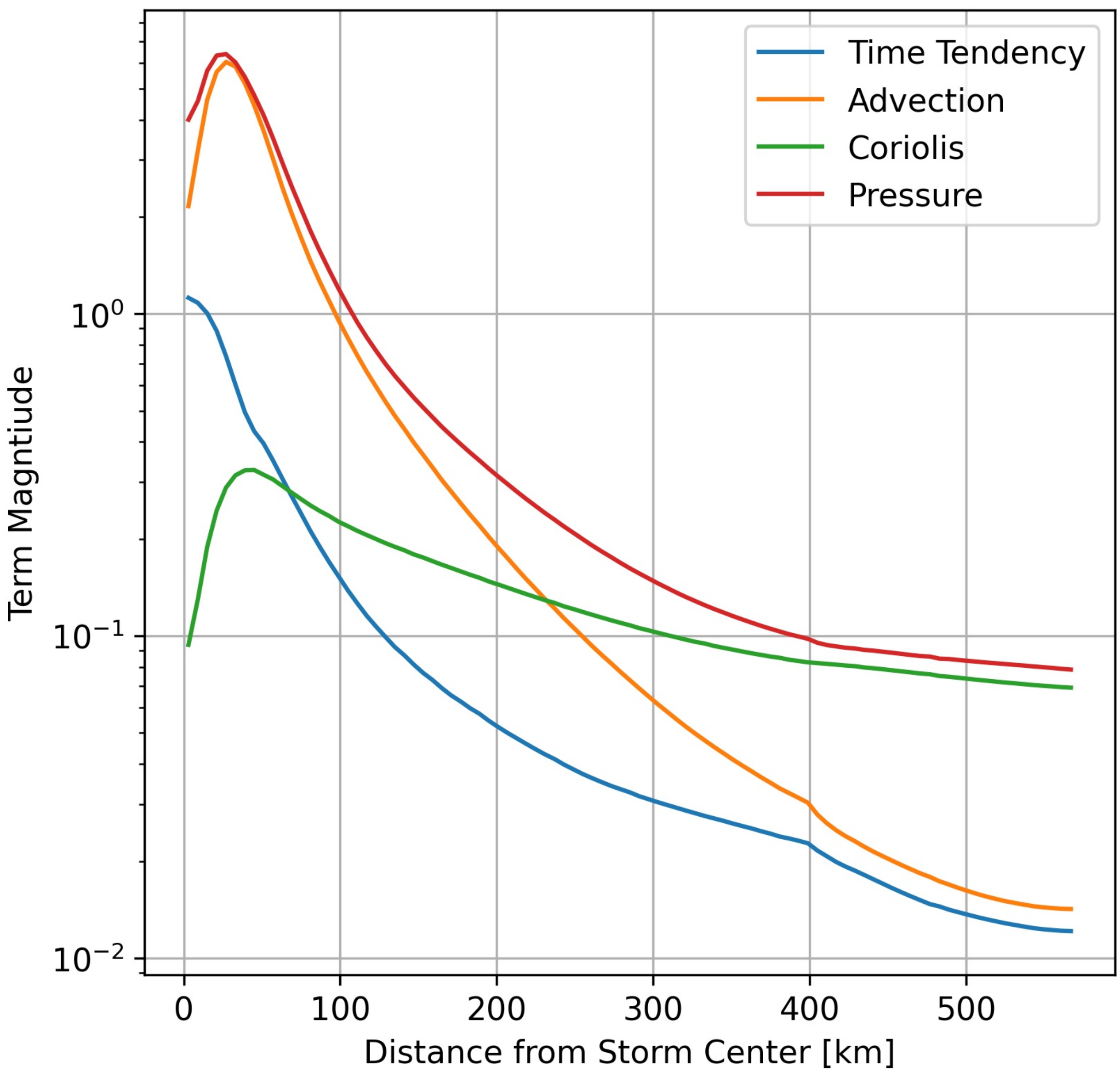
Supplementary Figure 5: **Continuity equation non-dimensionalized term magnitudes for the PINN trained on the storm in the 3D case.** Equation terms are labeled on the axis titles.



Supplementary Figure 6: **The effect of the γ parameter on the PINN results in the 2D case**. PINN results after full training using γ parameters of (a) 0.9, (b) 0.99, and (c) 0.999.



Supplementary Figure 7: Same as Supplementary Fig. 2, but for the 3D case. For the geopotential height panel on the right, we display Δh - the relative pressure for each pressure surface (each point is its geopotential height subtracted by the mean pressure surface height across the grid).



Supplementary Figure 8: **Radial Navier-Stokes Term Magnitudes** Magnitudes of the various Navier-Stokes Equation Terms in the 3D case by radius from the storm center, illustrating the transition from cyclostrophic balance near the core to geostrophic balance far from the core.