

Name / Source	Description / Variables	Usefulness / Notes	Link / Access
Measurements of the solar wind propagation delay (L1 → Earth)	Contains ~380 interplanetary shock events with ACE data and measured propagation delays. (zenodo.org)	This is good for supervised “delay prediction” tasks. You could incorporate physics constraints (e.g., from MHD models) and use delays as labels or part of loss terms.	Zenodo dataset page (zenodo.org)
Wind Solar Wind Weimer Propagation Details at 1-min resolution	Propagated solar wind data (magnetic field, plasma, delays etc.) at 1-minute resolution. (catalog.data.gov)	Useful for time series modeling of solar wind, possibly as “input at upstream / boundary” to propagate to Earth or magnetosphere.	data.gov page (catalog.data.gov)
NOAA Real-Time Solar Wind / Space Weather Data	Observations (e.g. from DSCOVR, ACE) of magnetic field, solar wind plasma, density, velocity, etc. (swpc.noaa.gov)	Good “raw” data for building PINNs, e.g. as boundary or input conditions. You may need to preprocess (interpolate, align).	NOAA / SWPC real-time solar wind data (swpc.noaa.gov)
NASA / NOAA Solar Wind Dataset (Kaggle)	Collections combining spacecraft measurements (solar wind, magnetic field) over time. (kaggle.com)	Good for prototyping, experiments, benchmarking.	Kaggle dataset pages (kaggle.com)
SuryaBench (Benchmark for Heliophysics / Space Weather ML)	A benchmark-level dataset derived from NASA’s SDO (solar imagery + auxiliary variables) spanning many years (2010–2024). (arxiv.org)	Very promising: includes inputs relevant for solar wind speed prediction, which can feed into a propagation PINN. You could use parts of it as “driving boundary data.”	Paper / dataset reference (arxiv.org)

Name / Source	Description / Variables	Usefulness / Notes	Link / Access
Physical & Numerical Experimental Datasets for Forchheimer Flow	Laboratory & simulated data of porous media flow (Darcy and non-Darcy regimes). Includes flow rates, pressure gradients, permeability, and media parameters.	Excellent for validating or training PINNs on controlled Darcy flow experiments. Subset (low Re) fits standard Darcy's law.	figshare.com
Darcy Groundwater v2 (GIS Raster Dataset)	Spatial data of groundwater head, flux, and conductivity derived from Darcy's law analysis.	Good for spatially distributed PINNs — can extract $(K(x,y))$, $(h(x,y))$, etc.	hub.arcgis.com
LamaH (Large-Sample Data for Hydrology, Central Europe)	Basin-scale hydrologic attributes, including baseflow and groundwater components.	Useful for broader hydrology PINNs linking surface and subsurface flow.	LamaH Dataset Info
Germasogeia Aquifer Data (Cyprus)	Field data from coastal aquifer used for modeling and optimization studies. Includes head, pumping, and salinity info.	Real-world aquifer test case for transient or steady-state PINNs.	arxiv.org/abs/1901.10810
Aquifer Modeling Example (Spain)	Numerical solution data and measurement comparisons for a confined aquifer.	Small-scale case suitable for reproducing with a PINN (steady-state flow).	arxiv.org/abs/1802.00890