

HASDM-SOL

ML THERMOSPHERE

← 175 - 825 KM

a Machine-learned Model of Thermospheric Density Trained on Two Solar Cycles

Alexander R. Lozinski, Shaylah M. Mutschler, Marcin D. Pilinski, W. Kent Tobiska, Andong Hu

OVERVIEW

HASDM-Sol is designed to be a surrogate model for the High Accuracy Satellite Drag Model (HASDM). HASDM-Sol uses space weather indices that are available operationally in near real-time to derive input features, and was trained using a 25.5 year dataset of **HASDM nowcast density** ρ_H as the target (Tobiska et al., 2021).

Inputs are derived from time series of: F10.7 solar radio flux, S10, M10, Y10, ap index, Disturbance storm time index and Auroral Electrojet index.

Output is global mass density $\rho(h, LST, \lambda)$ over the range $h=[175, 825\text{km}]$ at present time, where LST stands for local solar time.

SH COMPRESSION OF DENSITY

HASDM-Sol was not trained to predict ρ_H directly. Instead, a target variable was created by pre-processing the HASDM database $\rho_H(t, h, LST, \lambda)$ into sets of **spherical harmonic (SH) coefficients** f_{lm} :

- $\rho_H(LST, \lambda)$ was extracted for each time and altitude on the HASDM density grid, then used to calculate $g(LST, \lambda) = \log(\rho_H(LST, \lambda) \times 10^{12})$, following Acciarini et al. (2024).

- $g(LST, \lambda)$ was then considered to be a sampling of the function

$$g(LST, \lambda) = \sum_{l=0}^L \sum_{m=-l}^l f_{lm} Y_{lm}(\lambda, LST)$$

where f_{lm} are SH coefficients, Y_{lm} are real, 4π -normalized SH functions, l and m are the SH degree and order respectively, and L is the degree of the expansion.

- We calculated f_{lm} then truncated the expansion at $L=2$. This effectively reduced the 24×19 HASDM $LST \times \lambda$ grid from 456 to nine parameters for each time and altitude, a compression of $\sim 98\%$.

Figure 1 shows the reconstruction error as a result of evaluating the SH expansion up to limited order for $L=2, 4$ and 6 . **The error is small ($<0.3\%$), enabling an efficient reparameterization $\rho_H(t, h, LST, \lambda) \rightarrow f(t, h, q)$.**

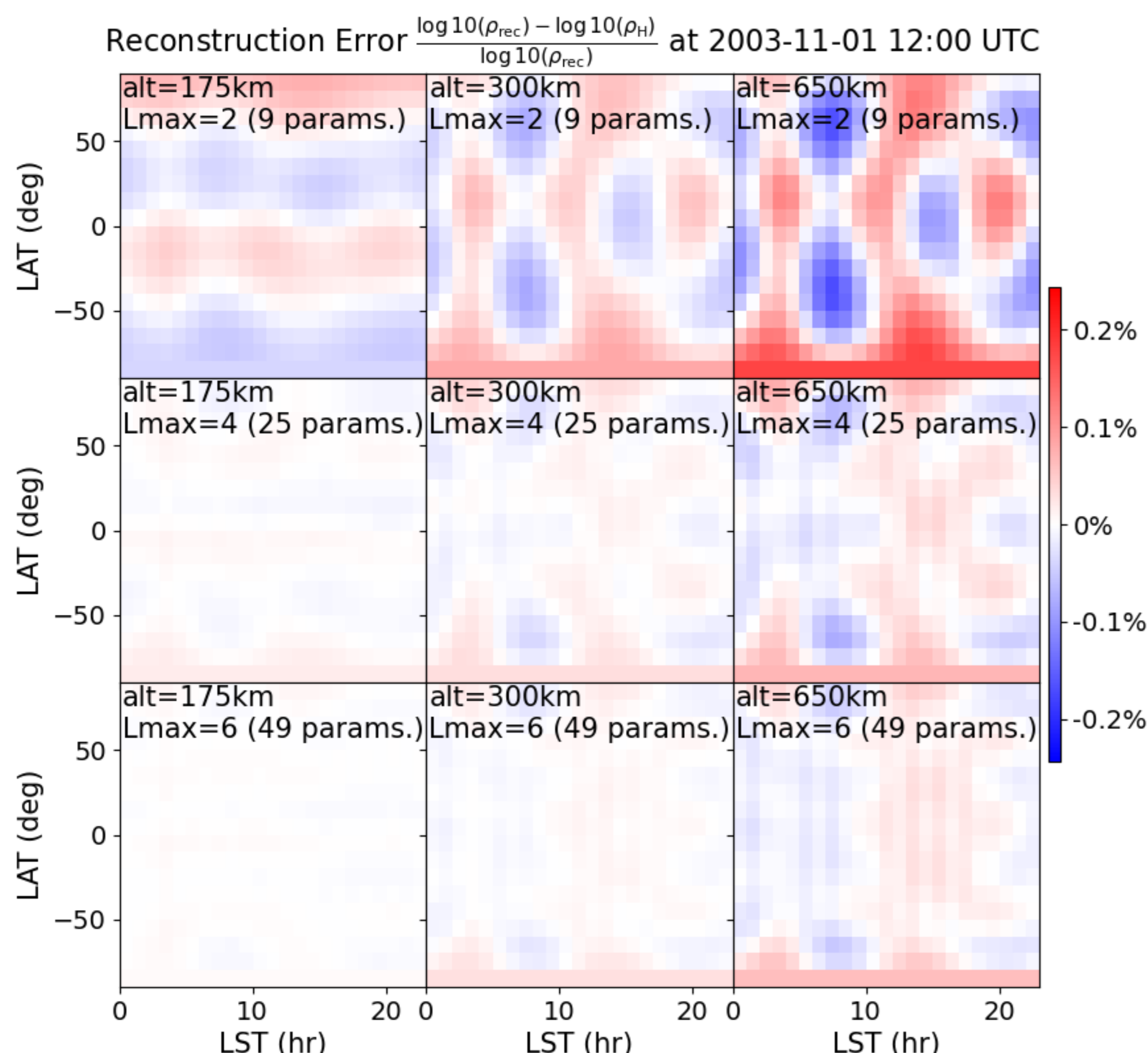


Figure 1: The error between HASDM density data ρ_H , and reconstructed HASDM density ρ_r calculated by evaluating the SH expansion up to degrees $L=2, 4$ and 6 (rows), shown at various altitudes (columns).

MODEL PIPELINE

A second parameter reduction method was applied within the model training loop: the altitude and SH coefficient dimensions were flattened then compressed into a set of **principal components**, i.e. $f(t, h, q) \rightarrow f(t, h \times q) \rightarrow z(t, k)$ with $nk < nh \times nq$.

However, this transformation must be derived from $f(t, h \times q)$ at a subset of time indices corresponding to training data, therefore it is part of the model pipeline, as shown below:

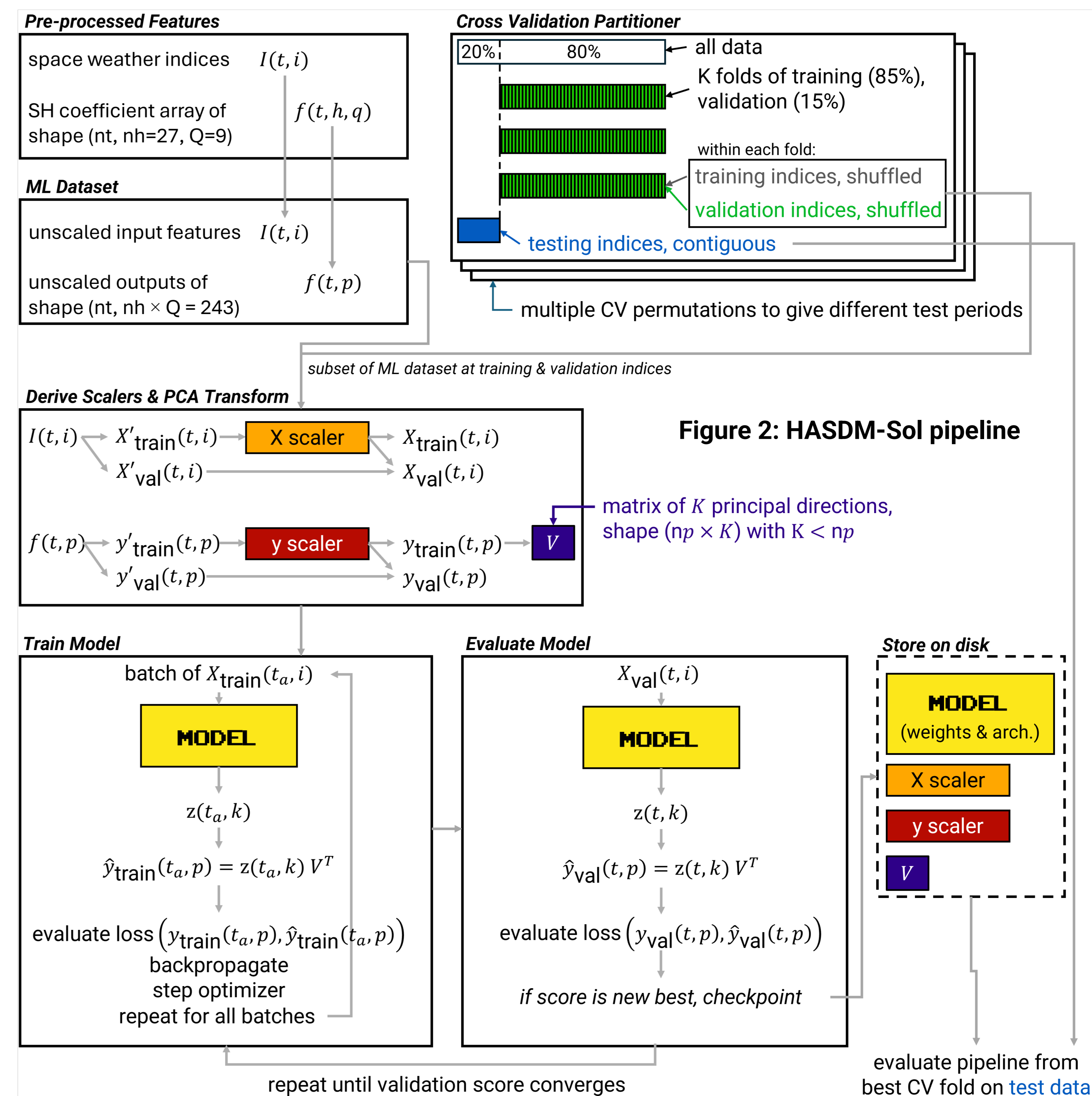


Figure 2: HASDM-Sol pipeline

RESULTS

A key objective for HASDM-Sol was to outperform the JB2008 empirical model (Bowman et al., 2008) in its agreement with HASDM. Table 1 shows the mean absolute percentage error (MAPE) in $\log_{10}(\rho)$ vs. HASDM at three altitudes (leftmost col.), binned by geomagnetic activity (second col.) and solar cycle (top row). The errors show **HASDM-Sol statistically outperforms JB2008**.

h (km)	activity bin	Solar Max 23		Solar Max 24		Solar Min (24 & 25)	
		HASDM.Sol	JB2008	HASDM.Sol	JB2008	HASDM.Sol	JB2008
175	G0	4.5%	5.7%	2.9%	3.9%	6.7%	9.2%
	G1	4.5%	5.4%	3.5%	4.2%	5.9%	10.0%
	G2	4.5%	5.9%	4.9%	4.7%	7.1%	5.8%
425	G0	8.5%	10.3%	8.8%	11.1%	13.1%	27.6%
	G1	12.3%	14.8%	11.7%	14.4%	18.8%	23.0%
	G2	12.0%	15.4%	10.3%	15.0%	18.6%	23.7%
525	G0	10.9%	13.6%	11.1%	14.4%	13.0%	27.1%
	G1	15.6%	19.6%	15.1%	18.6%	22.6%	27.5%
	G2	15.4%	20.3%	13.6%	20.2%	21.3%	28.6%

Table 1: Mean Absolute Percentage Error model vs. HASDM

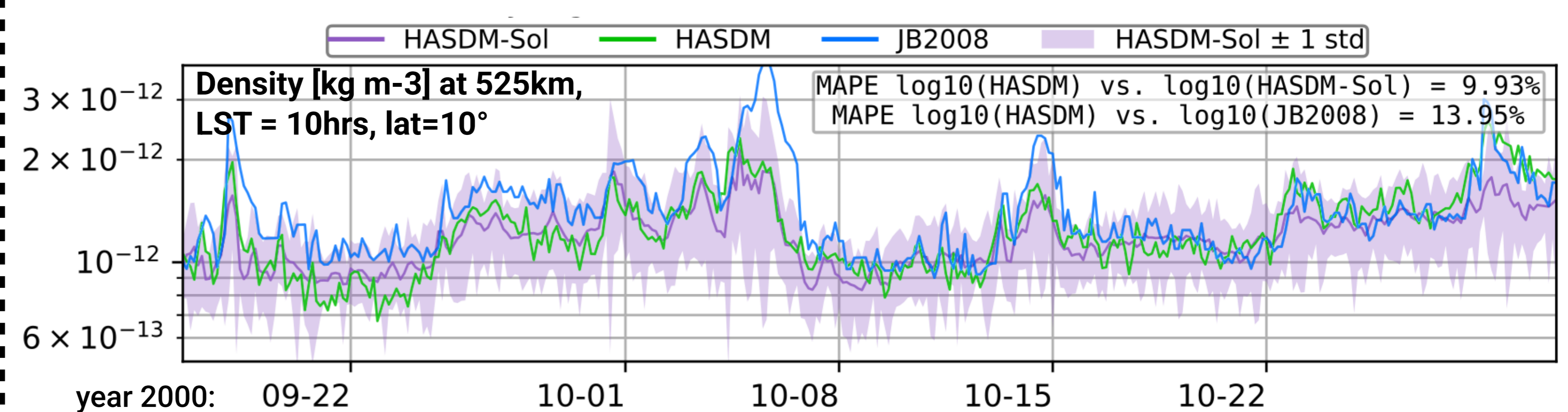


Figure 3: HASDM-Sol density predictions (purple) versus JB2008 (blue) and HASDM density data (green).