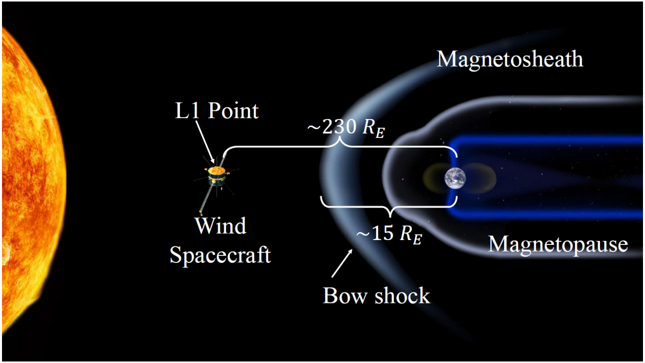
Non-Deterministic Model of the Earth's Magnetosheath

PRIME-SH

Connor O'Brien[1], Brian Walsh[1], Ying Zou[2], Ramiz Qudsi[1], Samira Tasnim[3], Huaming Zhang[4], David Sibeck[5]

October 6, 2023

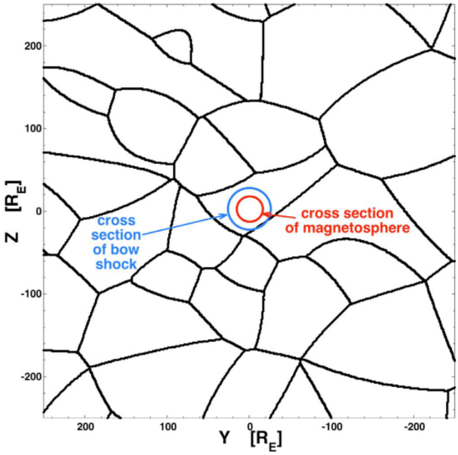
1: Center for Space Physics, Boston University, Boston, MA, USA
2: Applied Physics Lab, Johns Hopkins University, Laurel, MD, USA
3: German Aerospace Center (DLR), Institute for Solar-Terrestrial Physics, Neustrelitz, Germany
4: Computer Science Department, University of Alabama in Huntsville, Huntsville, AL, USA
5: Heliophysics Science Division, NASA/GSFC, Greenbelt, MD, USA



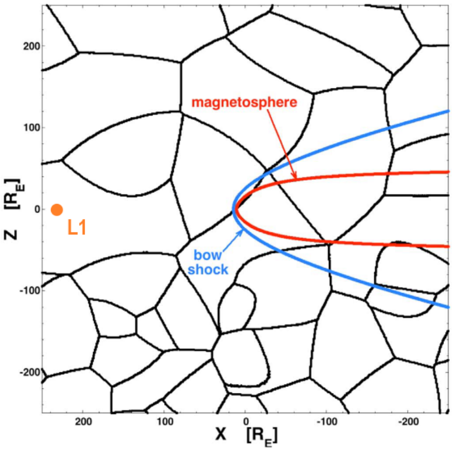
Location of L1 relative to important magnetospheric features (Bow shock, magnetopause, and magnetosheath). Not to scale, courtesy Ramiz Qudsi.

- The only continuous in-situ solar wind monitors are located at L1 (1,500,000km away from the Earth)
- In order to study solar wind driving at the Earth, it is necessary to propagate data from L1 monitors to the Earth

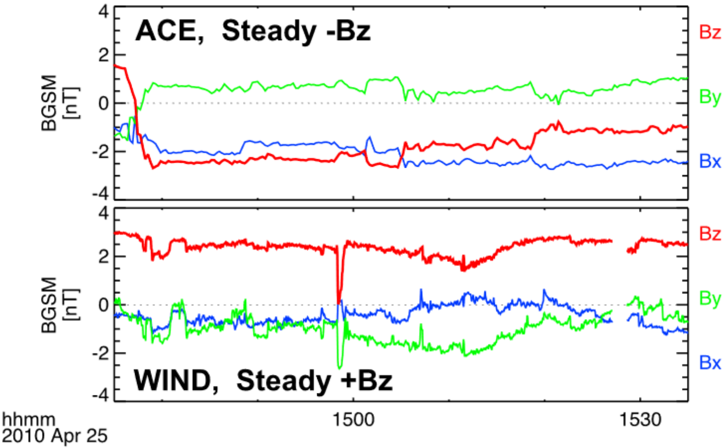
Solar Wind Flux Tube Structure



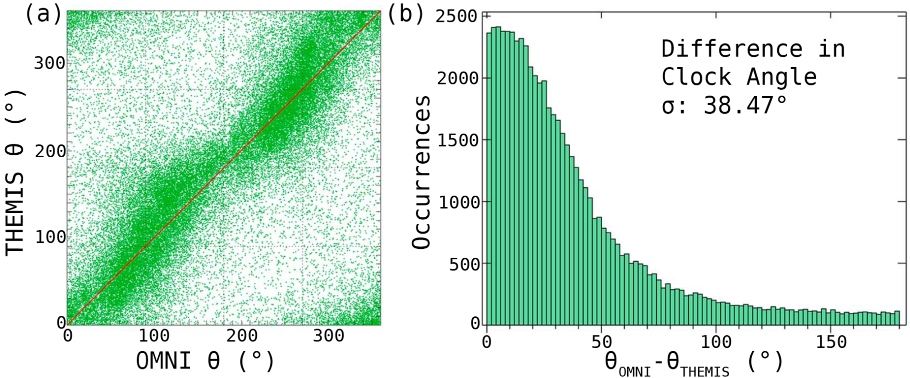
Sun's view of the magnetosphere with flux tube current sheet boundaries plotted (Borovsky 2018).



Side view of the magnetosphere with flux tube current sheet boundaries plotted (Borovsky 2018).



Discrepancy in the interplanetary magnetic field between two L1 monitors, each in a different flux tube (Walsh et al. 2019). Both monitors measure Bz to be steady but of opposite sign.



Difference in IMF clock angle between solar wind (OMNI) and magnetosheath (THEMIS) measurements. (Walsh et al. 2019)

Current approaches to determining inputs to the magnetosphere from L1 data have the following limitations:

1. Unphysical assumptions that ignore processing from L1 to the bow shock and the boundaries between flow parcels.
2. Failing to include physically meaningful uncertainties.
3. Stopping propagation at the nose of the bow shock, rather than the subsolar magnetopause where energy transfer actually takes place.

1 and 2 have been recently addressed by the Probabilistic Regressor for Input to the Magnetosphere Estimation (PRIME).

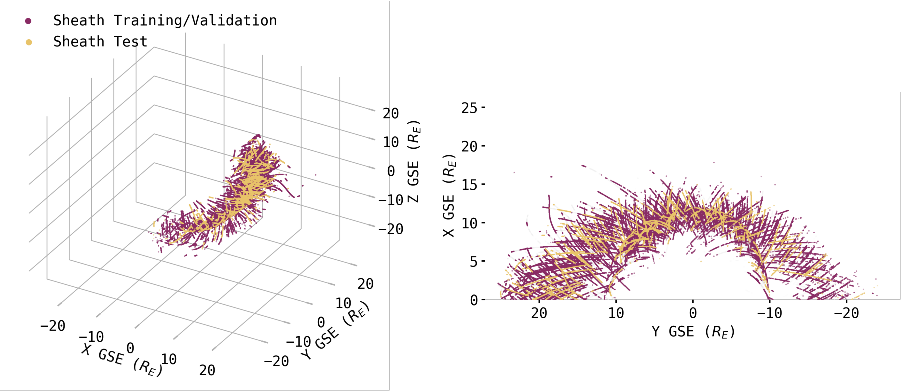
Can we develop an algorithm that predicts conditions in the magnetosheath in the same manner?

Scan for PRIME solar wind propagation manuscript:

The training of a new version of PRIME that predicts magnetosheath conditions (PRIME-SH) requires three ingredients:

1. MMS-1 magnetosheath data
2. Wind L1 solar wind data
3. Sufficiently representative network architecture

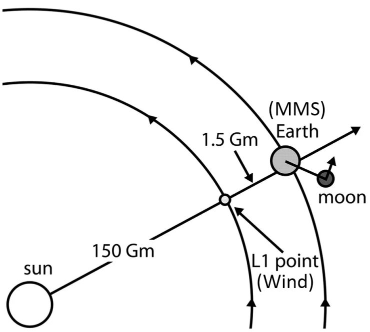
MMS Solar Wind Data Spatial Distribution

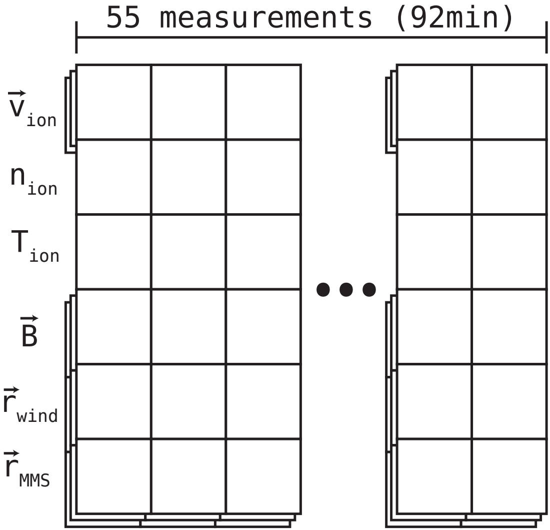


Full MMS-1 dataset of 117,427 samples (100s each) split into training/validation (purple, 80%) and test (yellow, 20%) datasets.

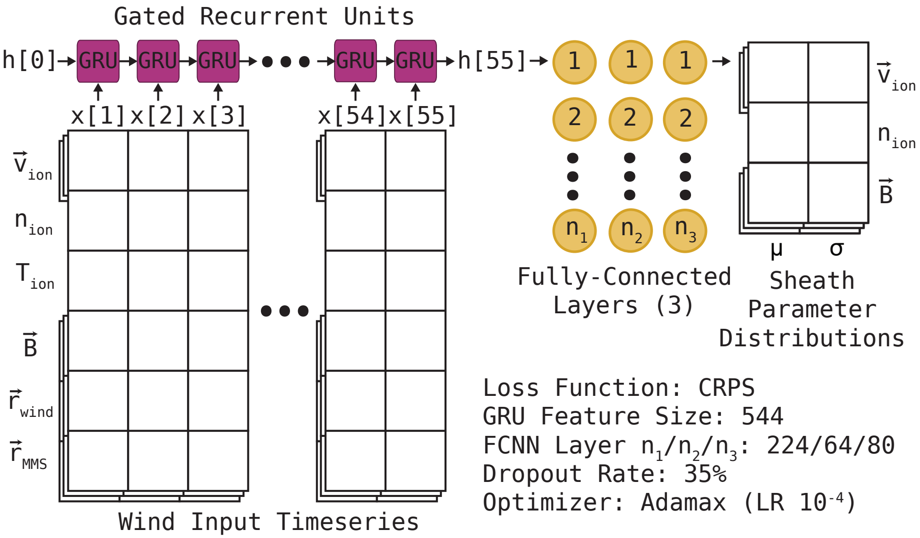
Parameters include $\vec{B}_{IMF}$, $\vec{v}_{sw}$, and n_i (seven components).

Wind Input Dataset





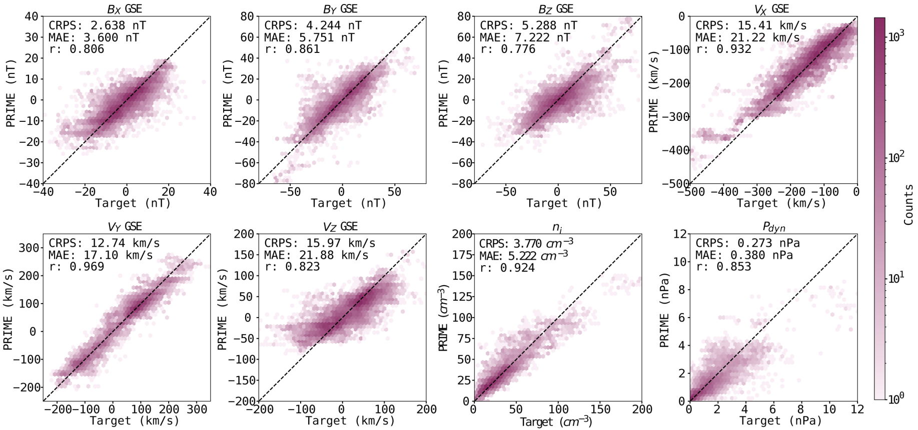
For each MMS target, 92 minutes of Wind data is assembled. Lead time is 27 minutes.



Loss Function: CRPS
GRU Feature Size: 544
FCNN Layer $n_1/n_2/n_3$: 224/64/80
Dropout Rate: 35%
Optimizer: Adamax (LR 10^{-4})

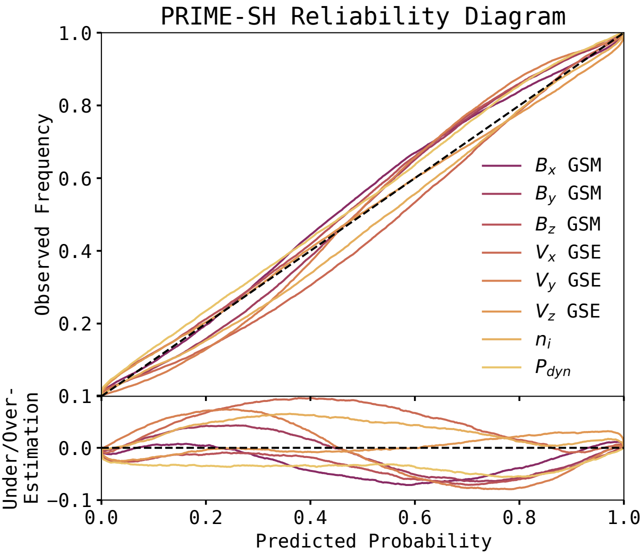
PRIME architecture is designed to output a probability distribution (μ and σ) for each parameter, rather than a single value.

PRIME-SH Overall Performance

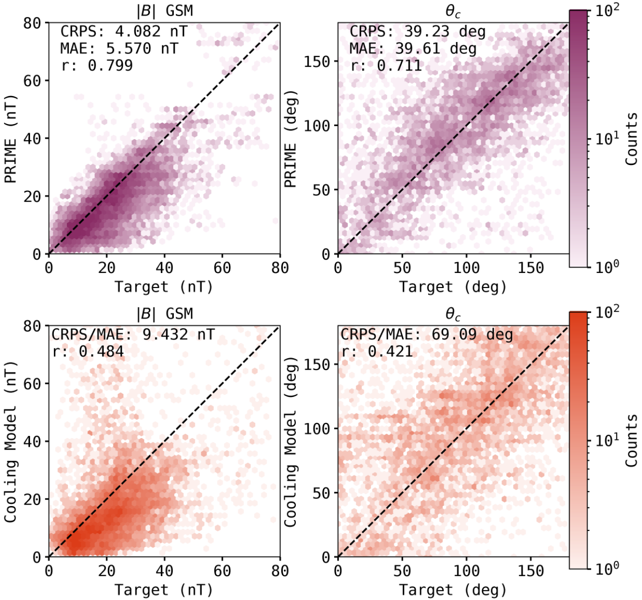


Joint distributions of MMS-1 data (x-axis) with predicted parameters from PRIME-SH. CRPS, the mean absolute error (MAE), and Pearson's r correlation coefficient for each parameter shown in the top left of each distribution. The MAE is calculated between the peaks of its predicted distributions and each MMS observation (thereby throwing away uncertainty information).

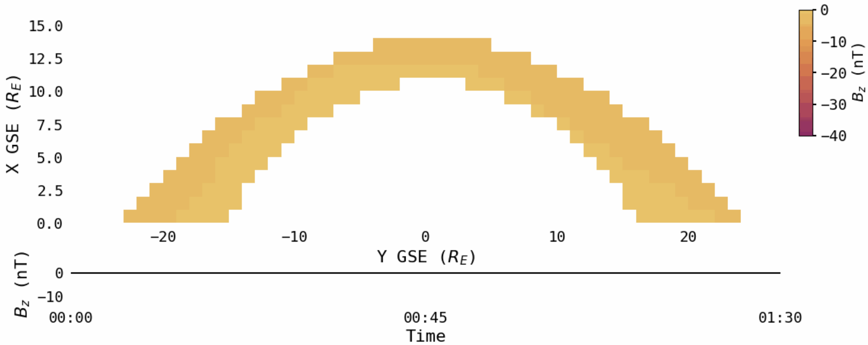
A forecast is reliable/calibrated if the predicted probabilities are matched by observed frequencies.
For bottom, a given parameter being over (under) the line corresponds to PRIME over (under) predicting the occurrence frequency.



Joint distributions of MMS-1 data (x-axis) with predicted parameters from PRIME-SH (top, purple) and the Cooling et al. 2001 model applied to OMNI IMF data (bottom, orange). CRPS, the mean absolute error (MAE), and Pearson's r correlation coefficient for each parameter shown in the top left of each distribution.



Preliminary Application: Dynamical Modeling



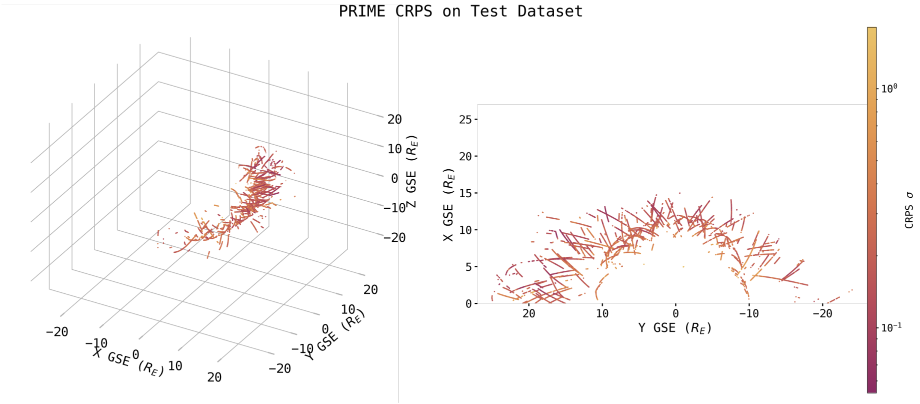
PRIME-SH is lightweight enough to generate predictions on a grid for real or synthetic data. Here we have average SW with synthetic southward turning of $\vec{B}_{IMF}$.

Conclusions-

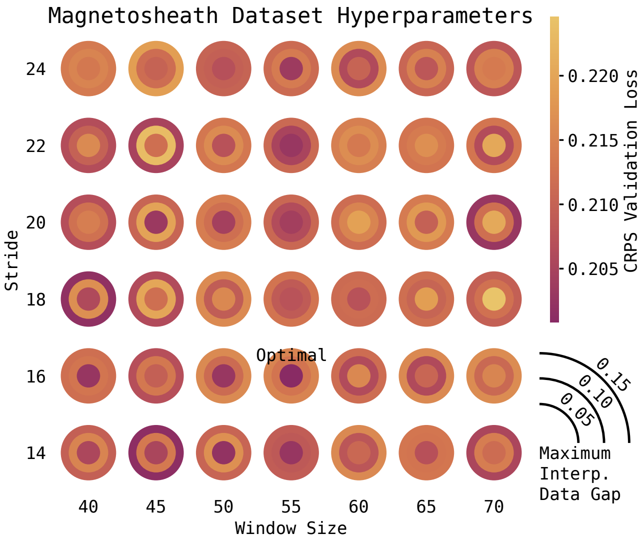
- PRIME's recurrent Bayesian neural network architecture proved capable of outperforming traditional solar wind propagation algorithms, motivating the development a magnetosheath prediction algorithm based on PRIME.
- PRIME-SH is a reliable, accurate magnetosheath prediction algorithm. Its performance is better than the current approach to statistical magnetosheath modeling (i.e. analytical draping codes).
- PRIME-SH can be used to investigate spatial and temporal evolution of the magnetosheath.

Follow PRIME-SH on GitHub:

Quick Summary:
Following success at modeling solar wind propagation, PRIME's recurrent Bayesian neural network architecture has been applied to the problem of magnetosheath modeling/prediction. Initial results show it to be reliable, accurate, and an improvement over analytical models.



PRIME-SH performance on the test dataset arranged in GSE coordinates.

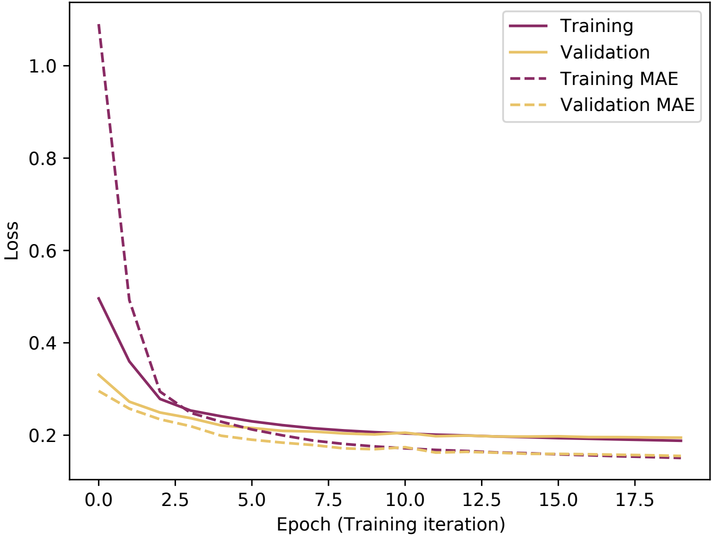


Search space for dataset hyperparameters, and the model performance for each set.

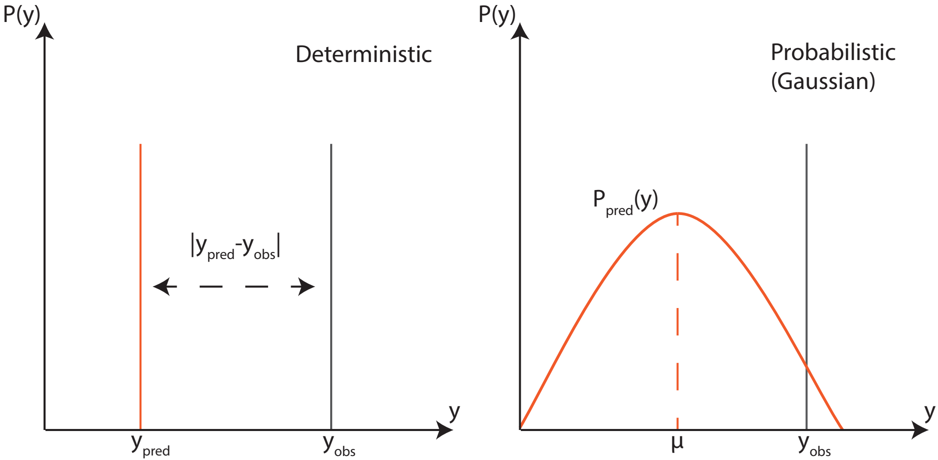
Window: input timeseries length.

Stride: lead time from end of window to prediction time.

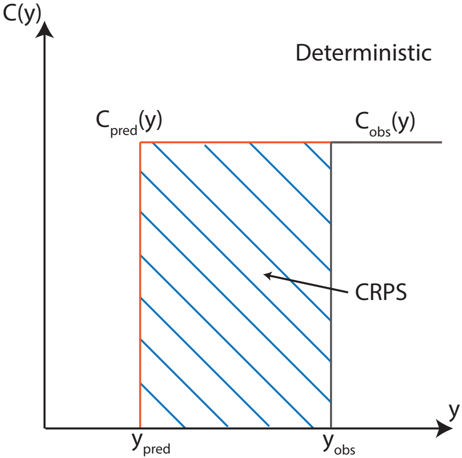
Interpolation fraction: percentage of input data in a row that is allowed to be interpolated.



CRPS and MAE loss on training and validation datasets during training. As can be seen, no overfitting has occurred.



Scoring metrics often take the form $Score = |Prediction - Observation|$. For deterministic predictions and observations, this formulation becomes the absolute error $AE = |y - y_{obs}|$.
For predictions that are probability distributions (i.e. have some uncertainty), the formulation is less obvious.



Reformulating in terms of the cumulative distribution function ($C(y) = P(y_{obs} \leq y)$), allows the generalization of the absolute error to probabilistic predictions (Continuous Rank Probability Score).

Continuous Rank Probability Score (CRPS) loss function:

$$CRPS = \int_{\infty}^{\infty} [C(y) - H(y - y_{obs})]^2 dy$$

$C(y)$: Cumulative distribution $H(y - y_{obs})$: Heaviside function
For a Gaussian distribution with mean μ and variance σ^2 this becomes

$$CRPS(\mu, \sigma) = \sigma[\frac{y_{obs} - \mu}{\sigma} erf(\frac{(y_{obs} - \mu)}{\sqrt{2}\sigma}) + \sqrt{\frac{2}{\pi}} exp(-\frac{(y_{obs} - \mu)^2}{2\sigma^2}) - sqrt\frac{1}{\pi}]$$

This function is often better than the negative log likelihood for training models because it:

- Symmetrically punishes over- and underconfident predictions
- Collapses to the mean absolute error for $\sigma \to 0$
- Has the same unit as the variable of interest