Machine Learning Approach to Determine Surface Radiative Fluxes based on CERES Footprint Observations at Low Latency

Jay Garg[1],

PC Sawaengphokai[1], Paul Stackhouse[3],

Hunter Winecoff[2], Walter Miller[1]

[1]ADNET Systems, Inc.

[2]Analytical Services and Materials, Inc.

[3]NASA Langley Research Center

Surface Flux Introduction

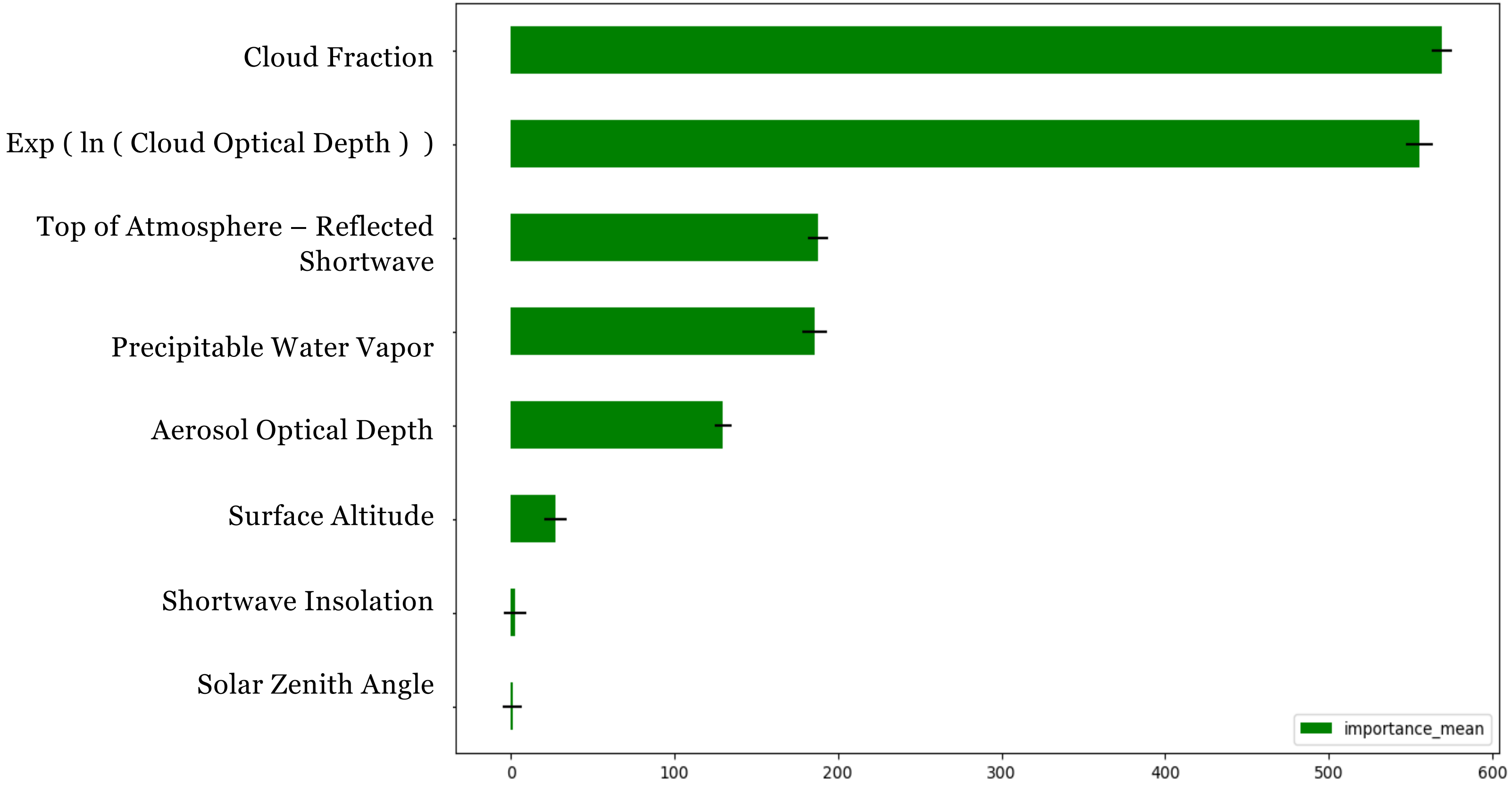
- ***FLASHFlux Objective:***
  - Provide satellite-based observations of TOA radiative fluxes, Clouds properties, and Surface radiative fluxes within 2-3 days for Level 2 SSF and up to a week for Level 3 TISA.
- ***Context***
  - The Langley Surface Parameterizations (LPSA/LPLA) are parameterized radiative transfer methods using various approximations for broadband scattering/absorption/transmission of the atmosphere and limited inputs for surface fluxes.
    - The models are older and efforts to update them tend to cause unintended effects
  - The Fu-Liou surface algorithm employed in CRS data products requires inputs that are not available and/or would require too much processing time for low latency processing.
- ***Surface Flux Objectives***
  - To replace the Langley Surface Parameterization (LPSA/LPLA) models with:
    - a Neural Network trained model that better represents the physics of radiative transfer in the atmosphere according to the CERES Fu-Liou algorithm
    - provide efficient estimates of surface fluxes for the FLASHFlux SSF data product, more consistent with the CRS (which is to become the primary surface flux product for CERES footprints).
  - Allows for more reliable estimates of surface flux models for intercomparison with other products (i.e., NOAA-ABI, CM-SAF SARAH) and users using overpass data for solar power and related applications

LOFO Variable Importance – SW

- Leave One Feature Out (LOFO)
  - involves repeatedly retraining the model and omitting a variable for each retrain, calculating difference in loss from the base model.
  - An increase in loss (green) indicates high importance and a decrease in loss (red) indicates negative impact to model performance.
  - Selecting the correct feature helps improve the performance and runtime of the model.
  - Running LOFO gives us insight into how different features cross-correlate and its impact on the model.
  - Shortwave Insolation and Solar Zenith Angle have high correlation, so LOFO doesn't give full story here

Terra Daytime

Train dataset: CRS_2019



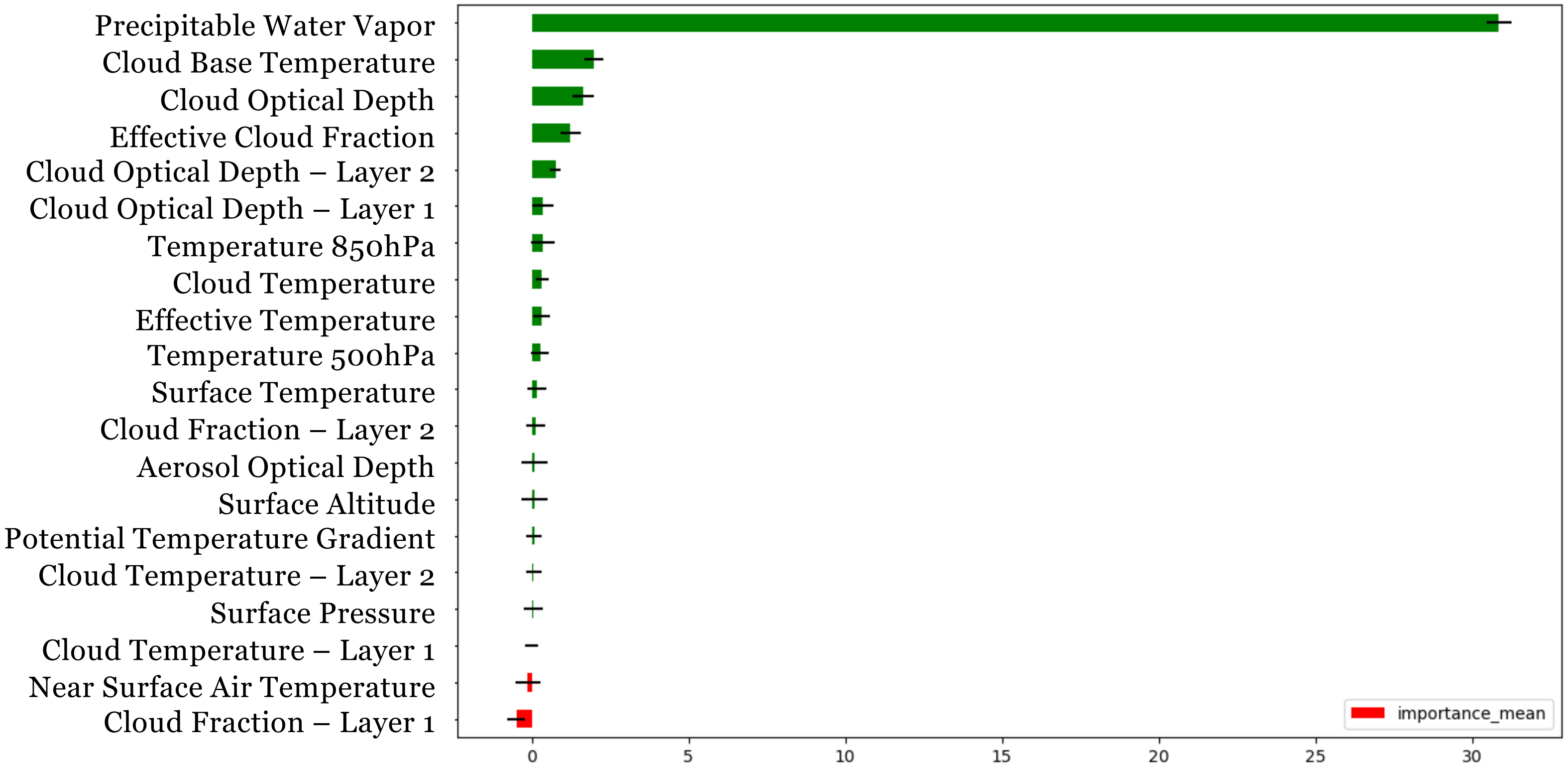
Shortwave Training Features

LOFO Variable Importance – LW

- **Leave One Feature Out (LOFO)**
  - **An increase in loss (green) indicates high importance and a decrease in loss (red) indicates negative impact to model performance.**
  - **LOFO plot doesn't give the entire story**
  - **Currently investigating effects of layer 1 variables of Cloud Temperature and Cloud Fraction and NSAT on model performance**
- **Why LOFO?**
  - **Many methods to select important features like Permutation Feature Importance, Partial Dependence Plots, Recursive Feature Elimination, and SHAP values.**
  - **Advantage of LOFO is that it is model agnostic and yield negative feature that hurt performance upon inclusion**

Terra Daytime

Train dataset: CRS_2019



Longwave Training Features

Model Comparisons – SW (All sky) January

Surface SW Down

Terra Daytime

ANN-CRS Total January Stats:

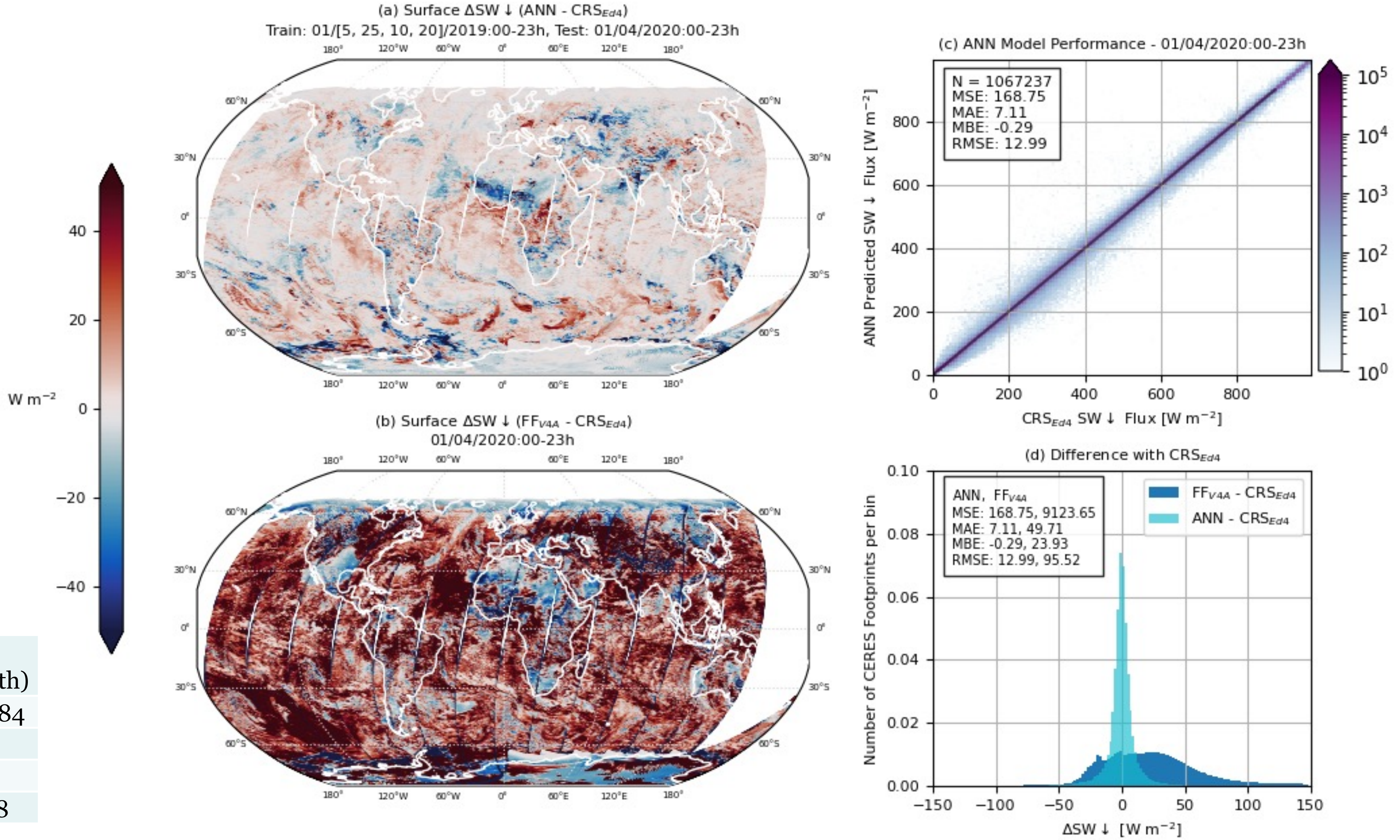
For Surface SW Down Daytime, we can see a tighter range of differences and a smaller bias relative to FF. Our differences are consistently small across all flux values. We are investigating the cause behind the systematic differences for the area along the sub-Sahara and Central Asia.

Train dataset: CRS_2019

Test dataset: CRS_2020

****FF uses GMAO GEOS5124 while CRS and ANN uses GMAO CERES541***

Difference with CRS	FF (day)	ANN (day)	ANN (month)
Mean Squared Error	9123.65	168.75	163.284
Mean Absolute Error	49.71	7.11	7.036
Mean Bias Error	23.93	-0.29	0.152
RMS Error	95.52	12.99	12.778



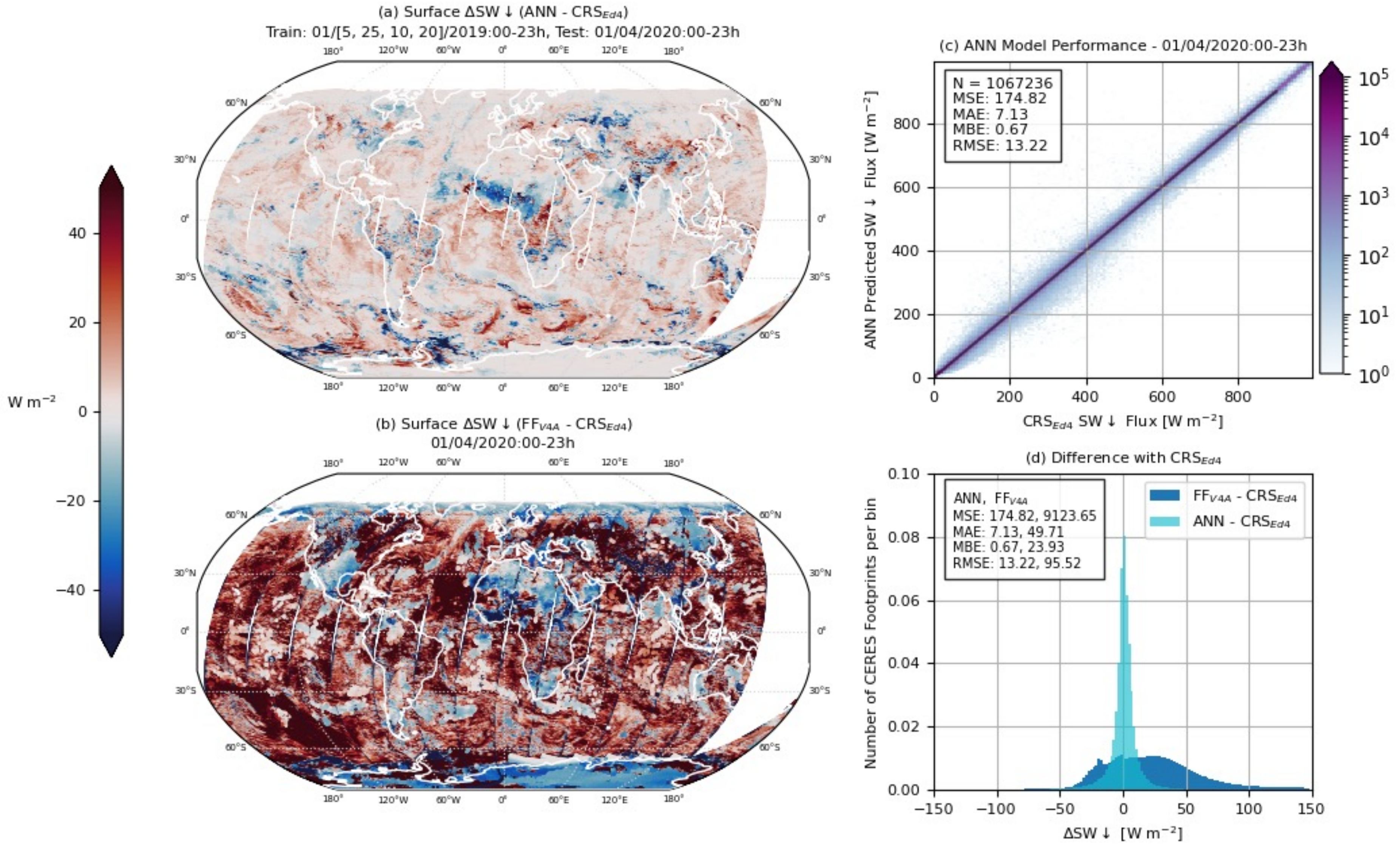
Model Comparisons – SW Combined Models

Surface SW Down

Terra Daytime
ANN-CRS Total January Stats:

Using separate clear sky – cloudy sky models did not give us better results

Does not reduce differences in sub-Sahara Africa and Central Asia or provide any other advantage



Difference with CRS	FF (day)	ANN (day)
Mean Squared Error	9123.65	174.82
Mean Absolute Error	49.71	7.13
Mean Bias Error	23.93	0.67
RMS Error	95.52	13.22

Model Comparisons – LW Day January

Surface LW Down

Terra Daytime

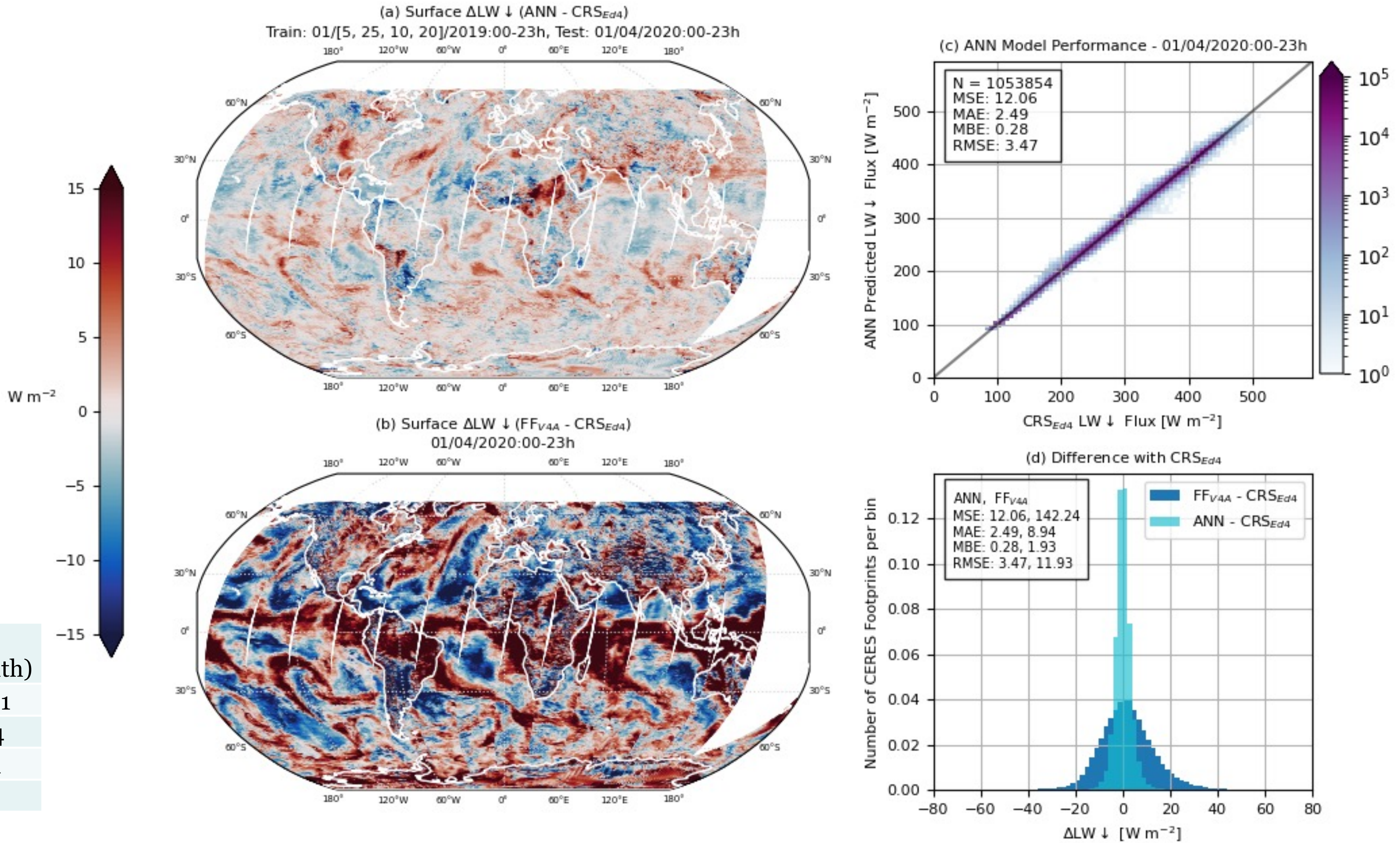
ANN-CRS Total January Stats:

Surface LW Down Daytime,

ANN differences with CRS are consistently small across all flux values compared to FF.

The bottom plot highlights the differences between FLASHFlux and CRS. These differences are at least partly attributed to differing meteorological inputs.

Difference with CRS	FF (day)	ANN (day)	ANN (month)
Mean Squared Error	142.24	12.06	12.331
Mean Absolute Error	8.94	2.49	2.504
Mean Bias Error	1.93	0.28	0.251
RMS Error	11.93	3.47	3.512



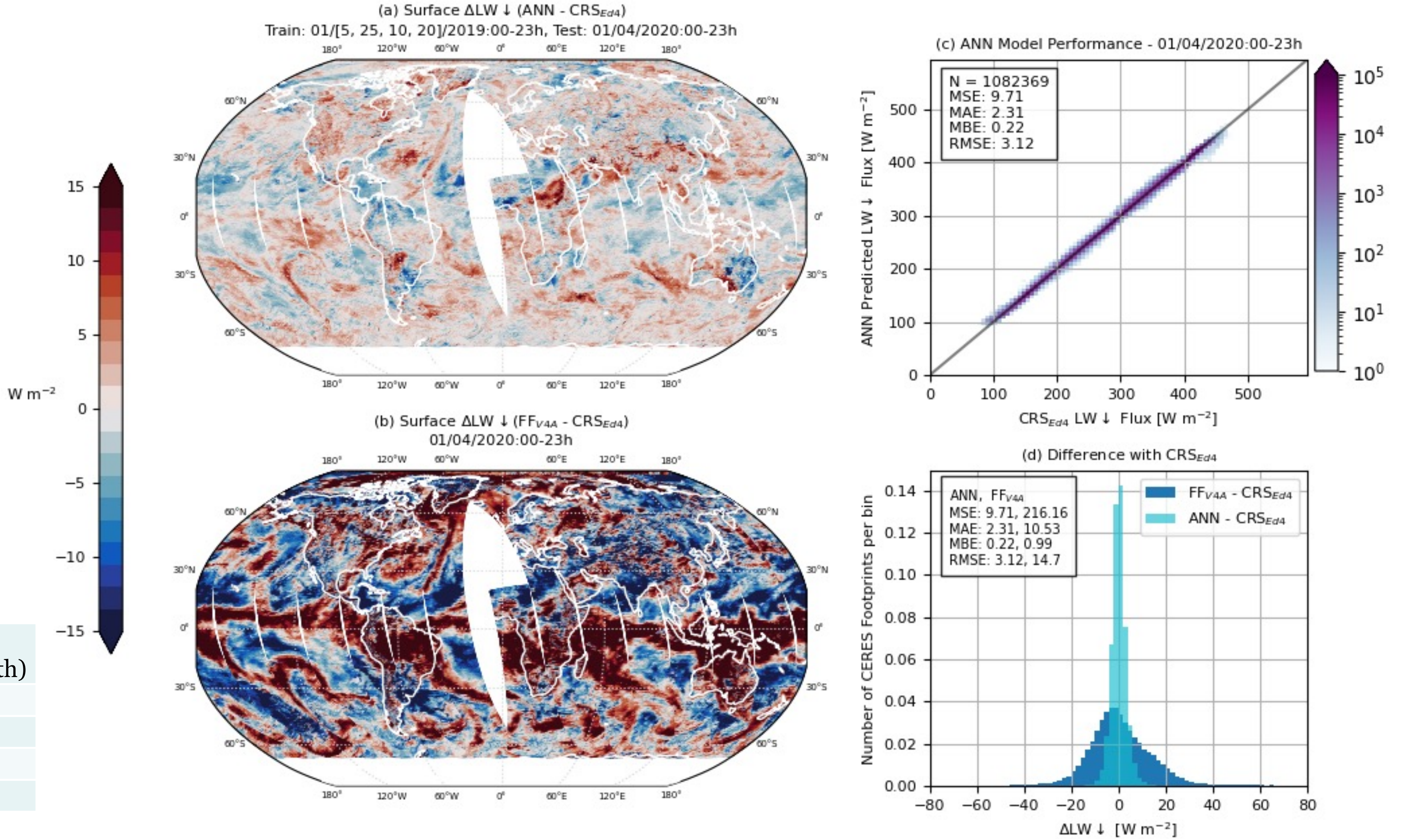
Model Comparisons – LW Night January

Surface LW Down

Terra Nighttime
ANN-CRS Total January Stats:

ANN differences are consistently small across all flux values. Some areas over continents and areas involving cloud patterns can be improved.

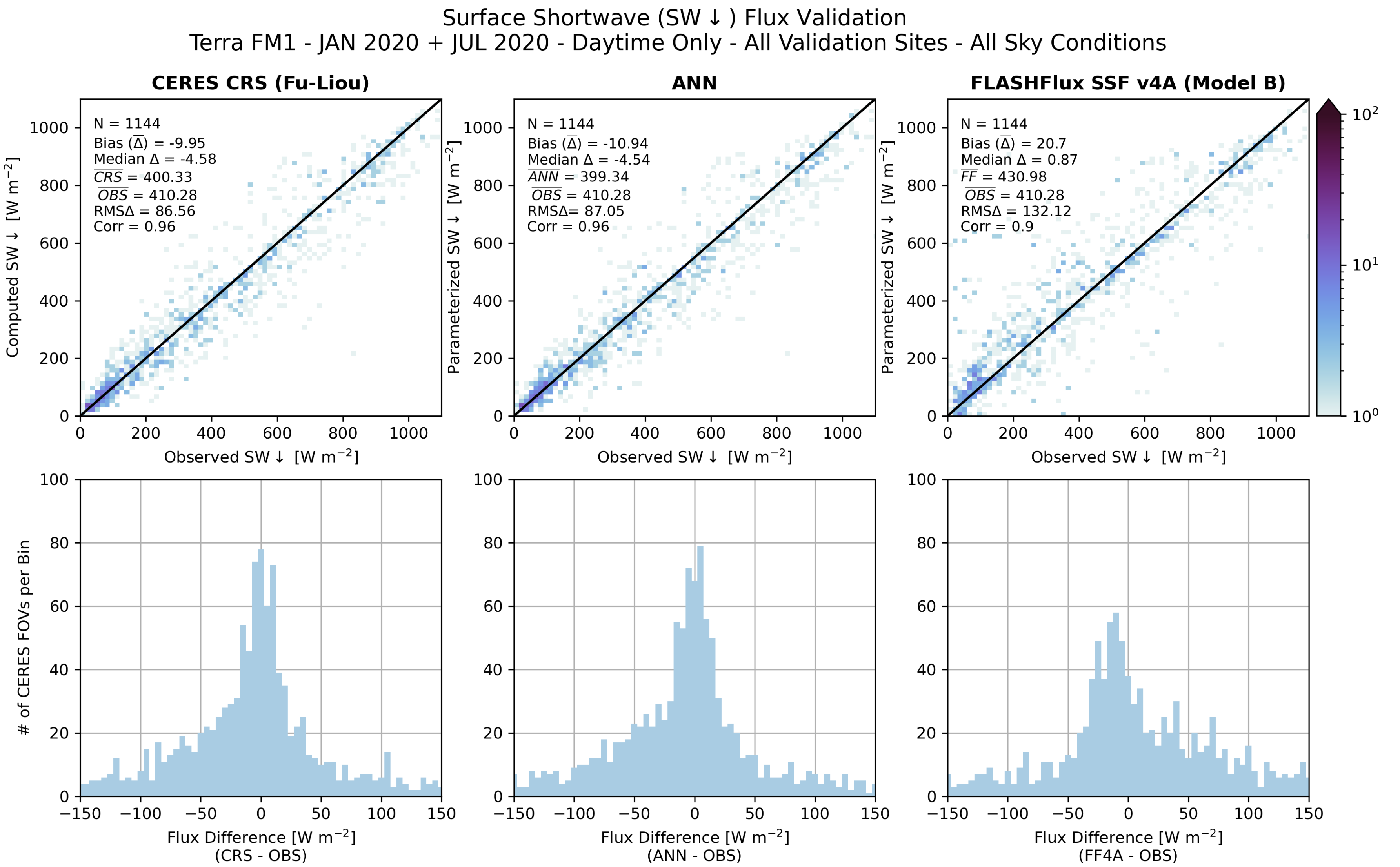
Difference with CRS	FF (day)	ANN (day)	ANN (month)
Mean Squared Error	216.16	9.71	9.497
Mean Absolute Error	10.53	2.31	2.298
Mean Bias Error	0.99	0.22	0.244
RMS Error	14.7	3.12	3.082



Surface Validations – SW Jan & July

SW Surface Flux Validation

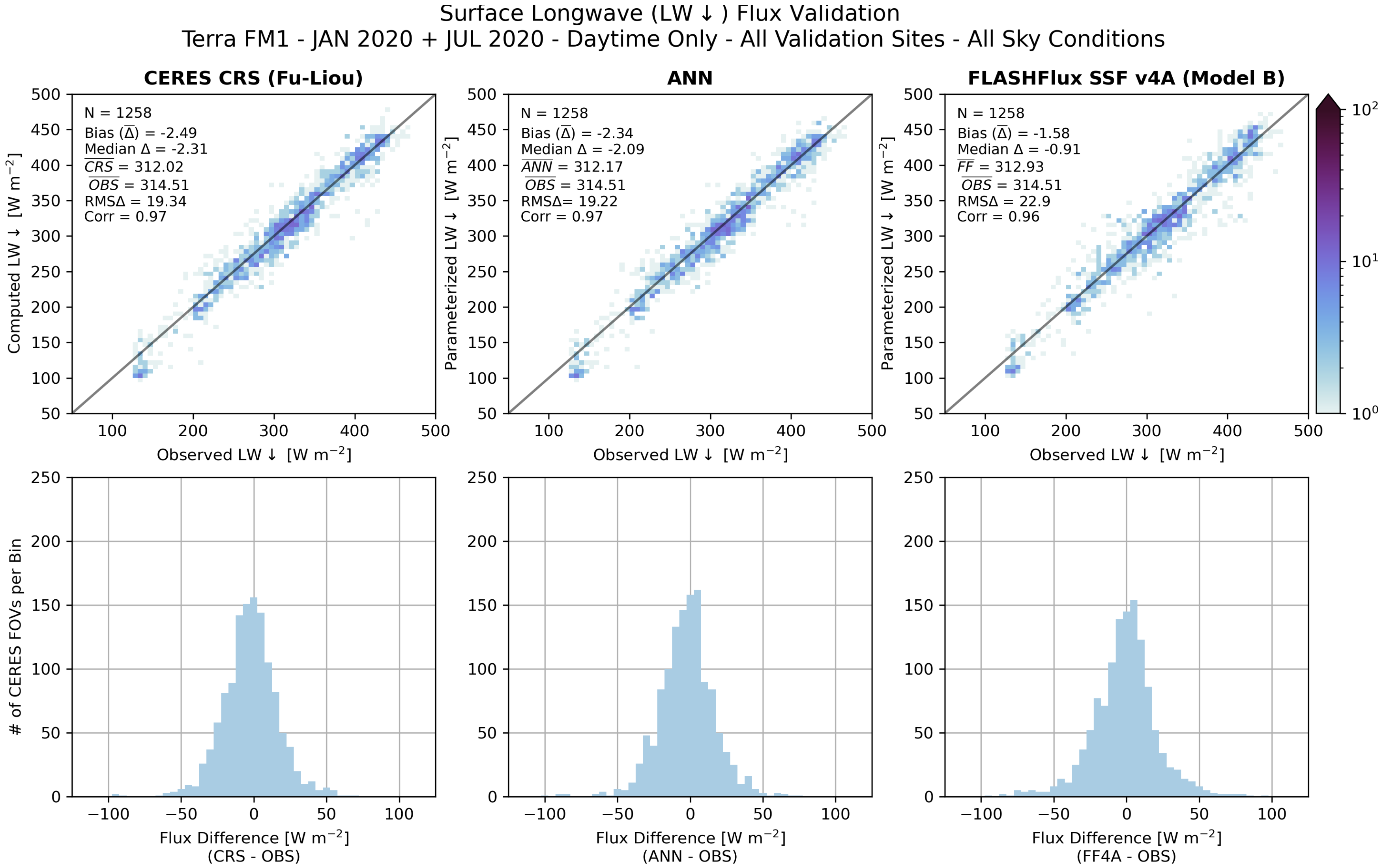
- Using Jan and July 2019 training sets, computed flux estimates for Jan and July 2020
- Compared CRS, ANN and archive FLASHFlux (v4A) with BSRN surface measurements
- CRS and ANN have nearly identical statistics and similar difference distributions => showing success of the ANN
- Both CRS and ANN, for these months reduce the bias (from +21 to ~ -10 W m^{-2}) but show a nearly 50 W m^{-2} reduction in RMS and correlation
- ANN SW is significantly improving fluxes over FF



Surface Validations – LW Day Jan & July

Daytime LW Flux Validation

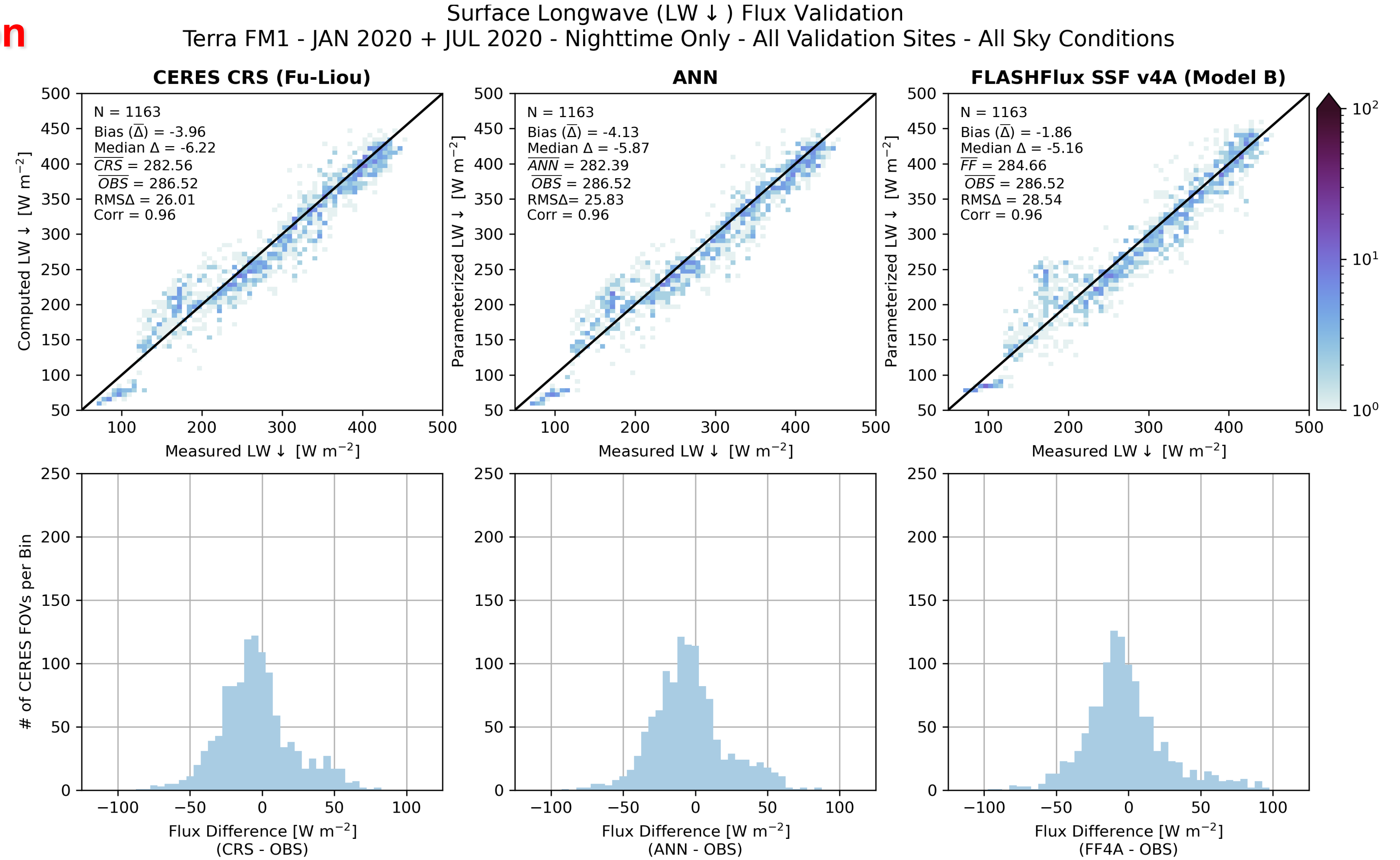
- Using Jan and July 2019 training sets, computed flux estimates for Jan and July 2020
- Compared CRS, ANN and archive FLASHFlux (v4A) with BSRN surface measurements
- CRS and ANN have nearly identical statistics and similar difference distributions => showing success of the ANN
- For daytime LW: CRS, ANN and LPLA are compatible, but CRS and ANN show slightly better RMS and correlation



Surface Validations – LW Night Jan & July

Nighttime LW Flux Validation

- Using Jan and July 2019 training sets, computed flux estimates for Jan and July 2020
- Compared CRS, ANN and archive FLASHFlux (v4A) with BSRN surface measurements
- CRS and ANN have nearly identical statistics and similar difference distributions => showing success of the ANN
- For daytime LW, CRS, ANN and LPLA are compatible:
  - LPLA shows slightly better bias
  - CRS and ANN show slightly better RMS and correlation



Conclusion

- An ANN was developed by training with the CRS data for all-sky surface SW and LW fluxes
  - Utilized the Leave-One-Feature-Out to evaluate the importance and then select various key inputs
- The resulting surface fluxes from the ANN shows near identical statistics and distribution to ground observations when compares to CRS.
- ANN SW model shows significant improvements over the LPSA model use in FLASHFlux.
- ANN LW model shows a larger bias compares to the LPLA model use in FLASHFlux.

• Future Work

- Assess key areas of uncertainty:
  - Investigate systematic differences in the ANN SW model in comparison to the CRS over the sub-Sahara and Central Asia.
  - Assess impact of adding a near surface temperature inversion correction (Gupta 2010) as implemented in the LPLA to help dry-arid region and polar regions
- Use FLASHFlux SSF with GMAO GEOS-IT meteorology as input to test the sensitivity of the ANN fits.
- Determine if January trained fits can be use seasonally or if monthly training for each fits are required.
- Testing on multiple years' data using same training model
- Training new ANN using Aqua CRS data and test it with Aqua SSF and NOAA-20 SSF.
- Training SW & LW Net model
- Refining ANN architecture for preparation for operations

Gupta et. al., J. Appl. Meteor. Climatol, **49,** 1579-1589

• *Extra Slides*

Intentional Left Blank

CERES FLASHFlux Overview

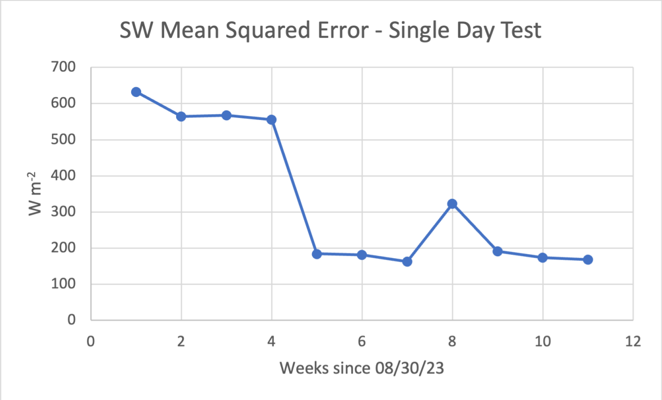
- ***FLASHFlux Overview***
  - Provides satellite-based observations of radiative fluxes and cloud systems within a week
  - CERES typically takes several months before release
  - FLASHFlux uses CERES based production system through inversion
  - FLASHFlux uses a simplified calibration, an operational meteorological product from GMAO anand it's own surface parameterization.
  - Terra and NOAA20 Satellites
- ***FLASHFlux Latency Objectives***
  - SSF products within 3-4 days
  - Global 1x1 daily averages from FF TISA; goal: 5-7 days latency
- ***FLASHFlux Uses***
  - Primarily used for applied science and education (i.e., POWER and Globe Clouds)
  - Supports also QC for selected missions (e.g., NOAA NESDIS)
  - TOA gridded fluxes; normalized to TOA EBAF for annual "State of the Climate" assessments

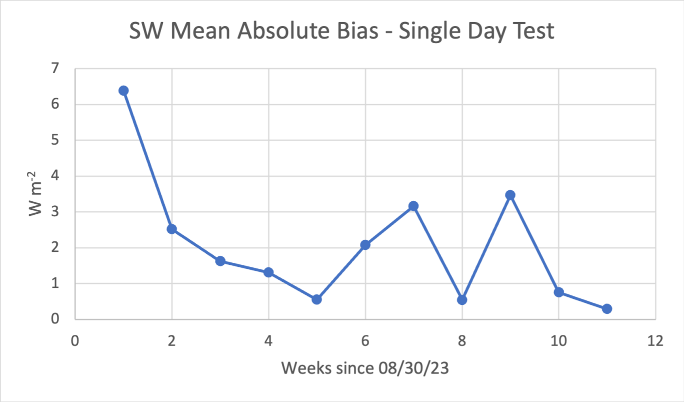
Model Improvement

**4 training days across all models*

Mean Squared Error | *Bias*

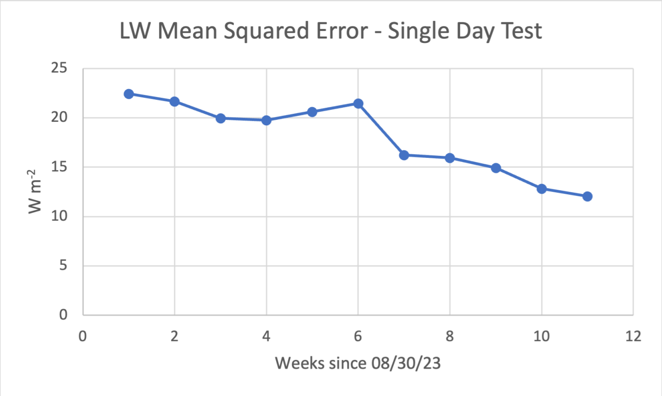
Shortwave

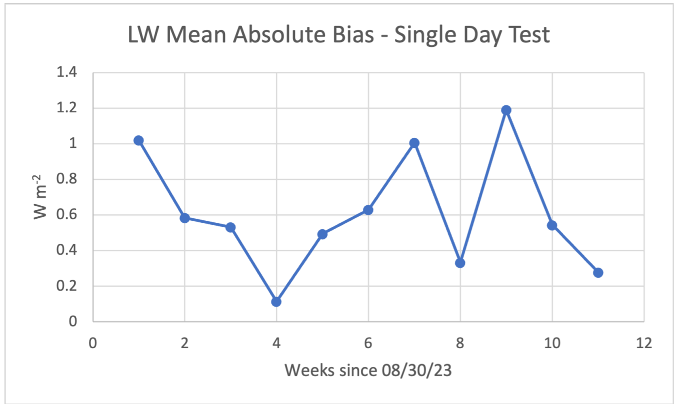




- *Our initial run for SW surface flux down included the training variables TOA insolation, solar zenith angle, cloud fraction, cloud optical depth, precipitable water vapor, aerosol optical depth, and altitude.*
- *Our final run included all the above variables except for cloud optical depth, we instead ran the exponential function of the weighted average of the log value of COD for each of the 2 cloud layers, we also included TOA reflected shortwave in our features.*

Longwave





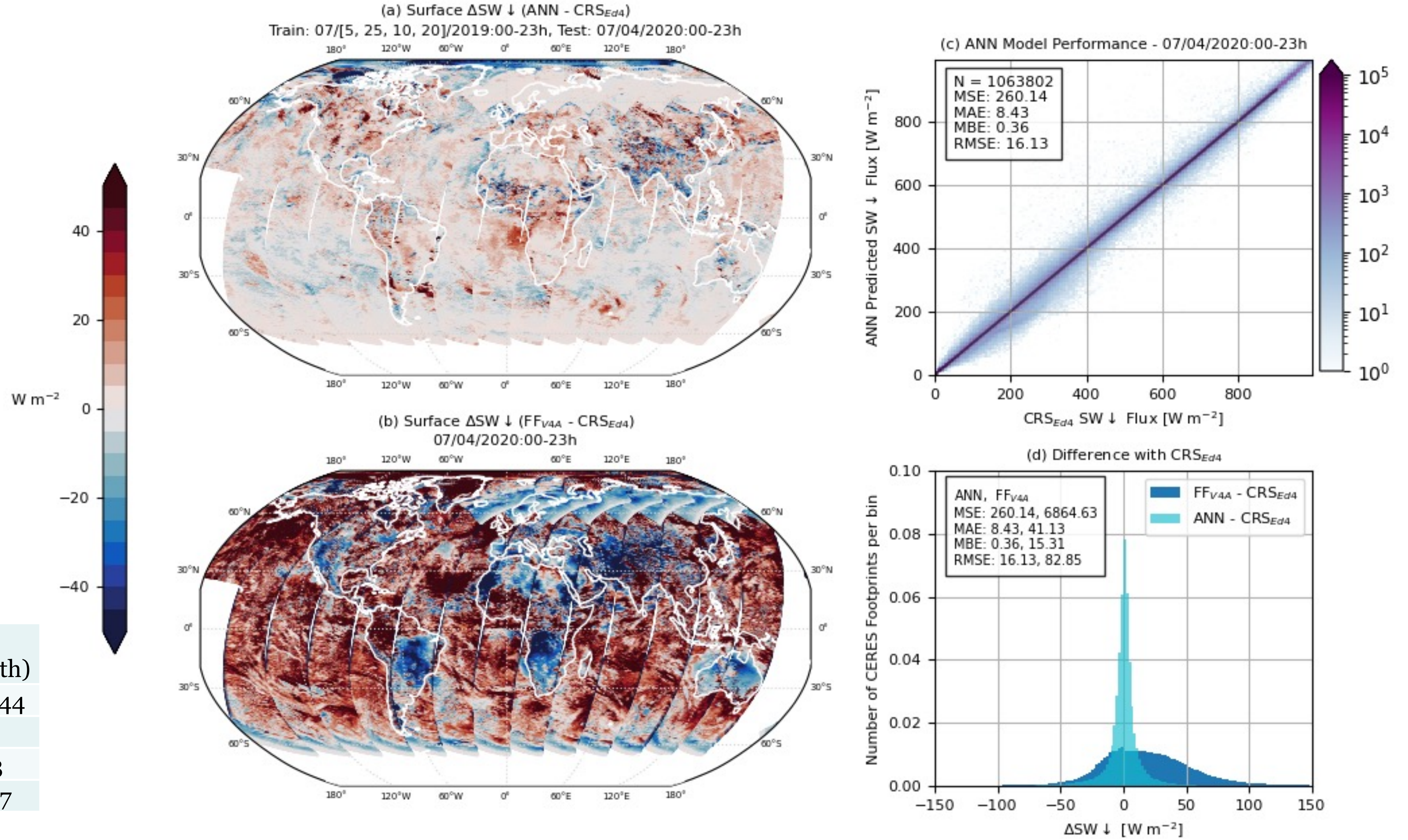
- *Our initial run for LW surface flux down included the training variables effective temperature, cloud fraction, cloud optical depth, precipitable water vapor, potential temperature gradient, cloud temperature, and altitude.*
- *Our final run (daytime and nighttime) included all the above variables and near surface air temperature, aerosol optical depth, temperature at surface and 500 hPA and 850 hPa, surface pressure, cloud fractions and temperatures and optical depths for each layer, and cloud base temperature.*

Model Comparisons – SW July

Surface SW Dowı

Terra Daytime
ANN-CRS Total July Stats:

- Significantly worse if using January trained model



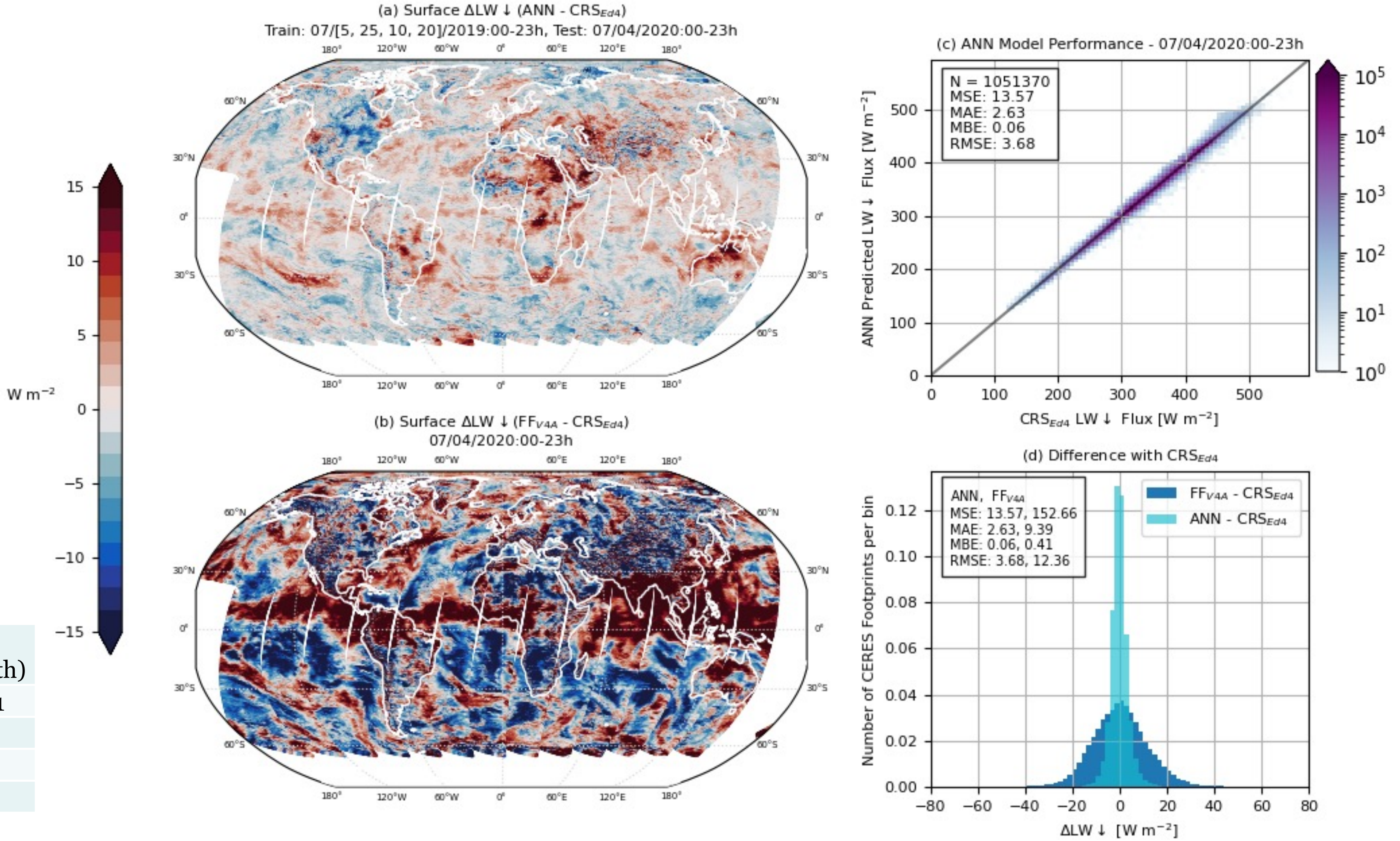
Difference with CRS	FF (day)	ANN (day)	ANN (month)
Mean Squared Error	6864.6	260.14	195.644
Mean Absolute Error	41.13	8.43	7.412
Mean Bias Error	15.31	0.36	0.493
RMS Error	82.85	16.13	13.987

Model Comparisons – LW Day July

Surface LW Down

Terra Daytime
ANN-CRS Total July Stats:

- Significantly worse if using January trained model



Difference with CRS	FF (day)	ANN (day)	ANN (month)
Mean Squared Error	152.66	13.57	12.991
Mean Absolute Error	9.39	2.63	2.589
Mean Bias Error	0.41	0.06	0.263
RMS Error	12.36	3.68	3.604

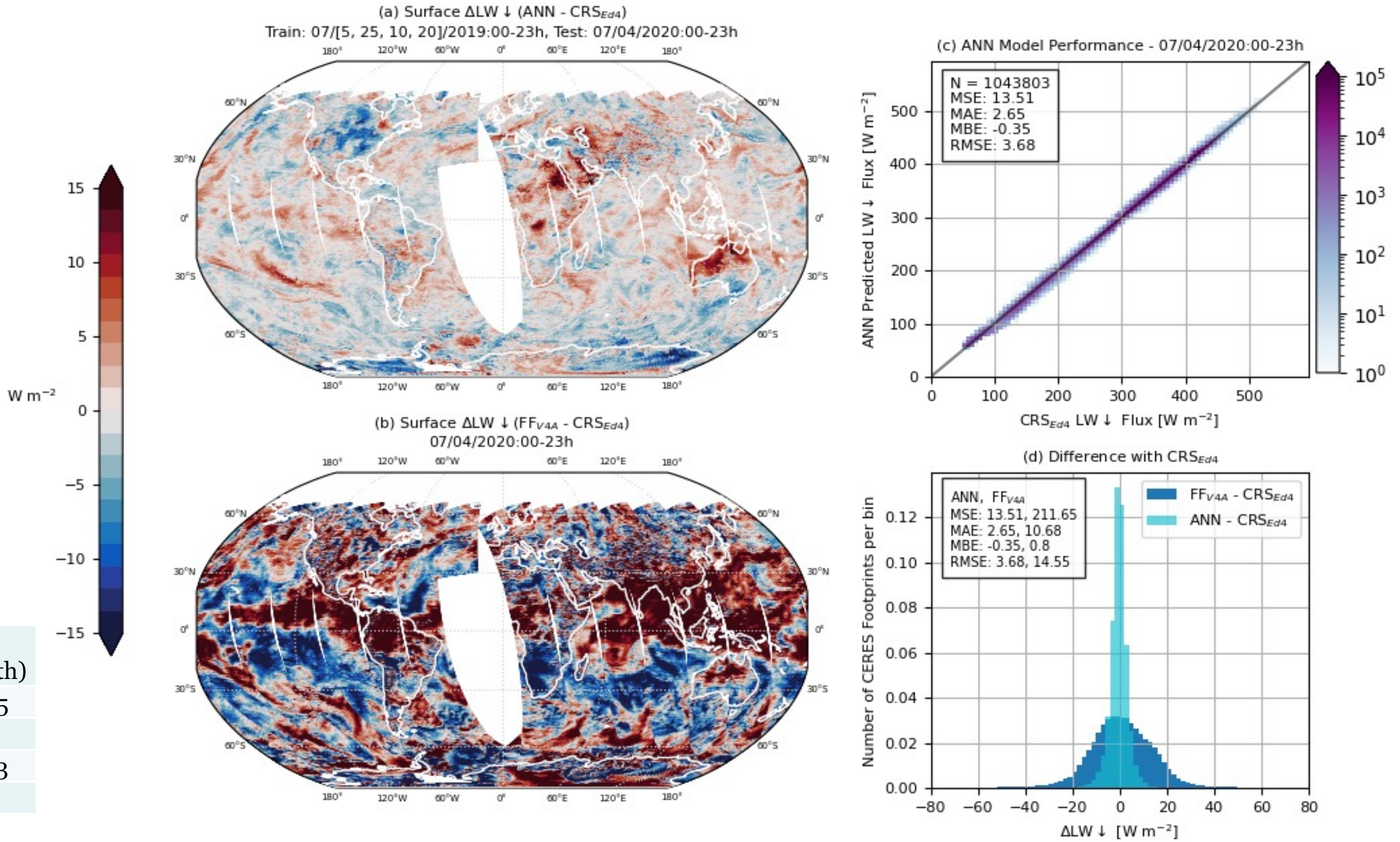
Model Comparisons – LW Night July

Surface LW Down

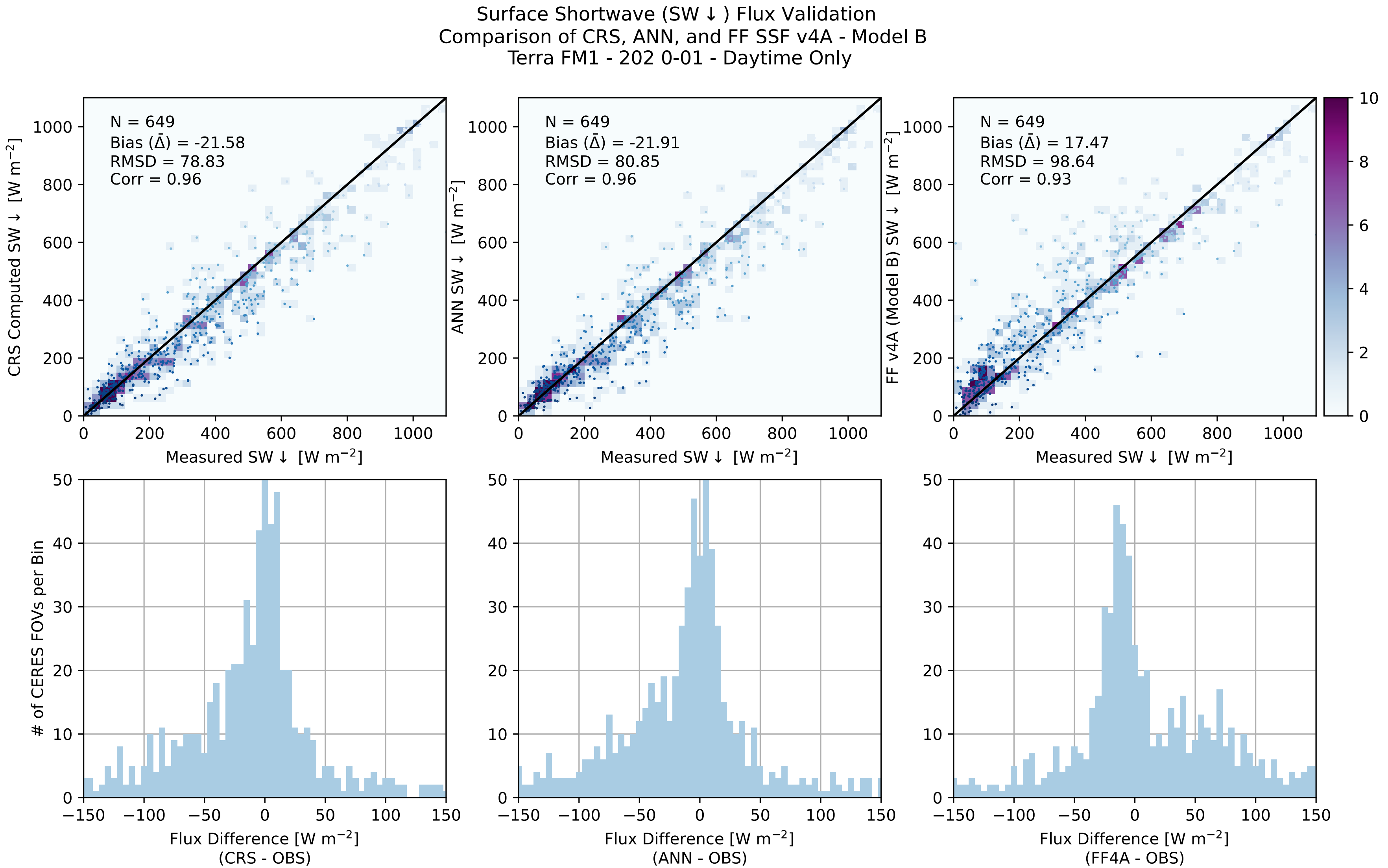
Terra Nighttime
ANN-CRS Total July Stats:

- Significantly worse if using January trained model

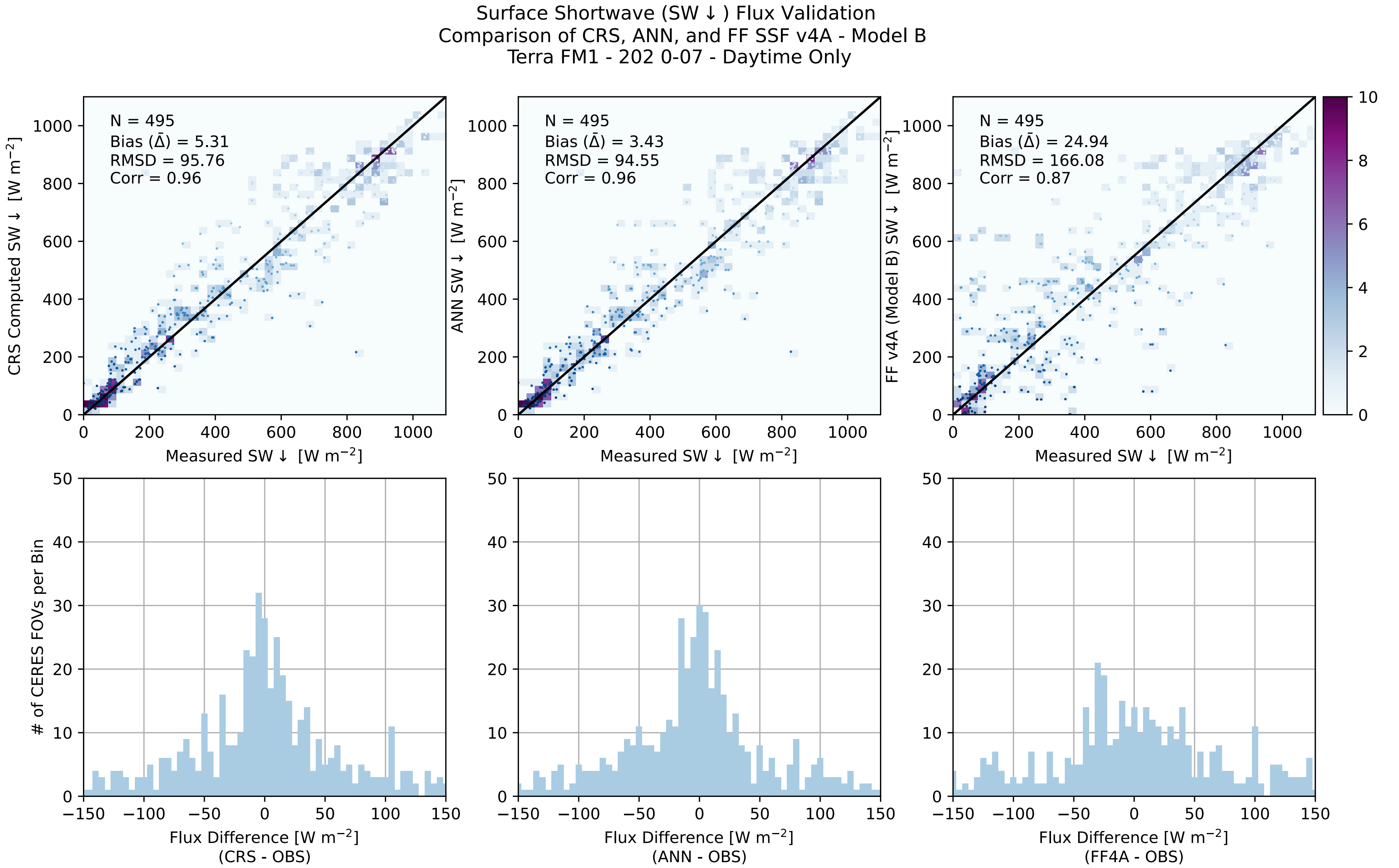
Difference with CRS	FF (day)	ANN (day)	ANN (month)
Mean Squared Error	211.65	13.51	12.985
Mean Absolute Error	10.68	2.65	2.622
Mean Bias Error	0.8	-0.35	-0.263
RMS Error	14.55	3.68	3.603



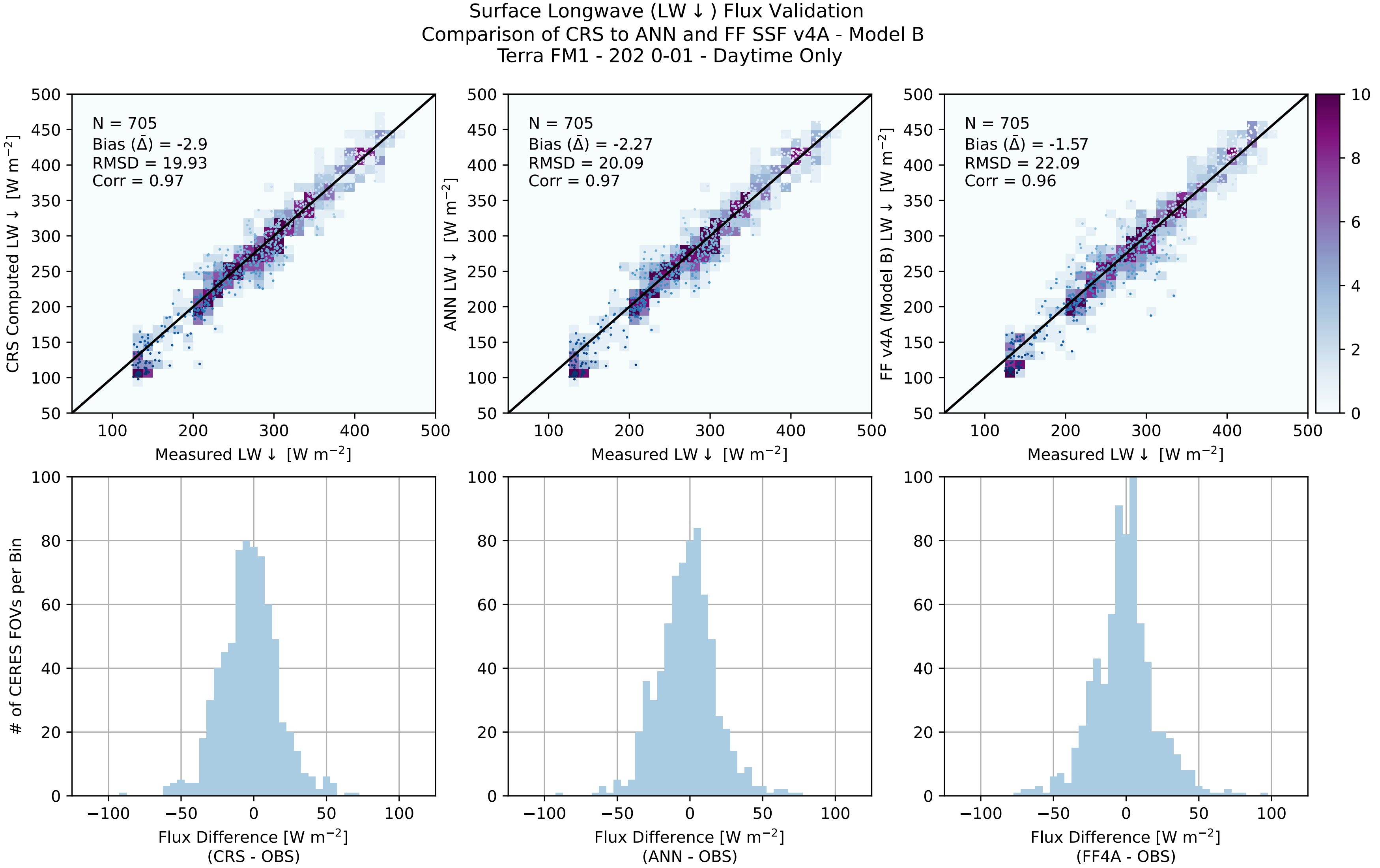
Surface Validations – SW Jan



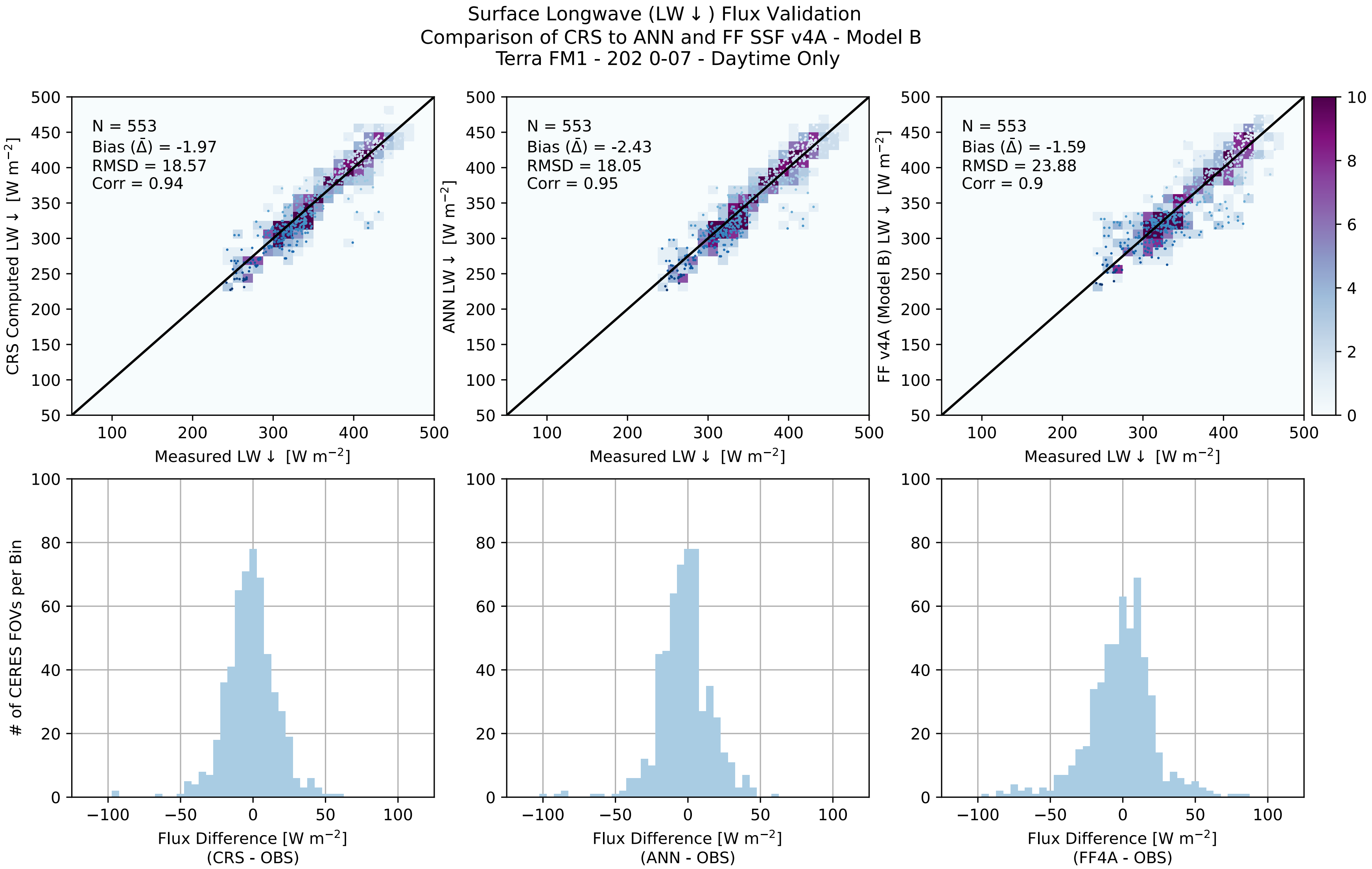
Surface Validations – SW July



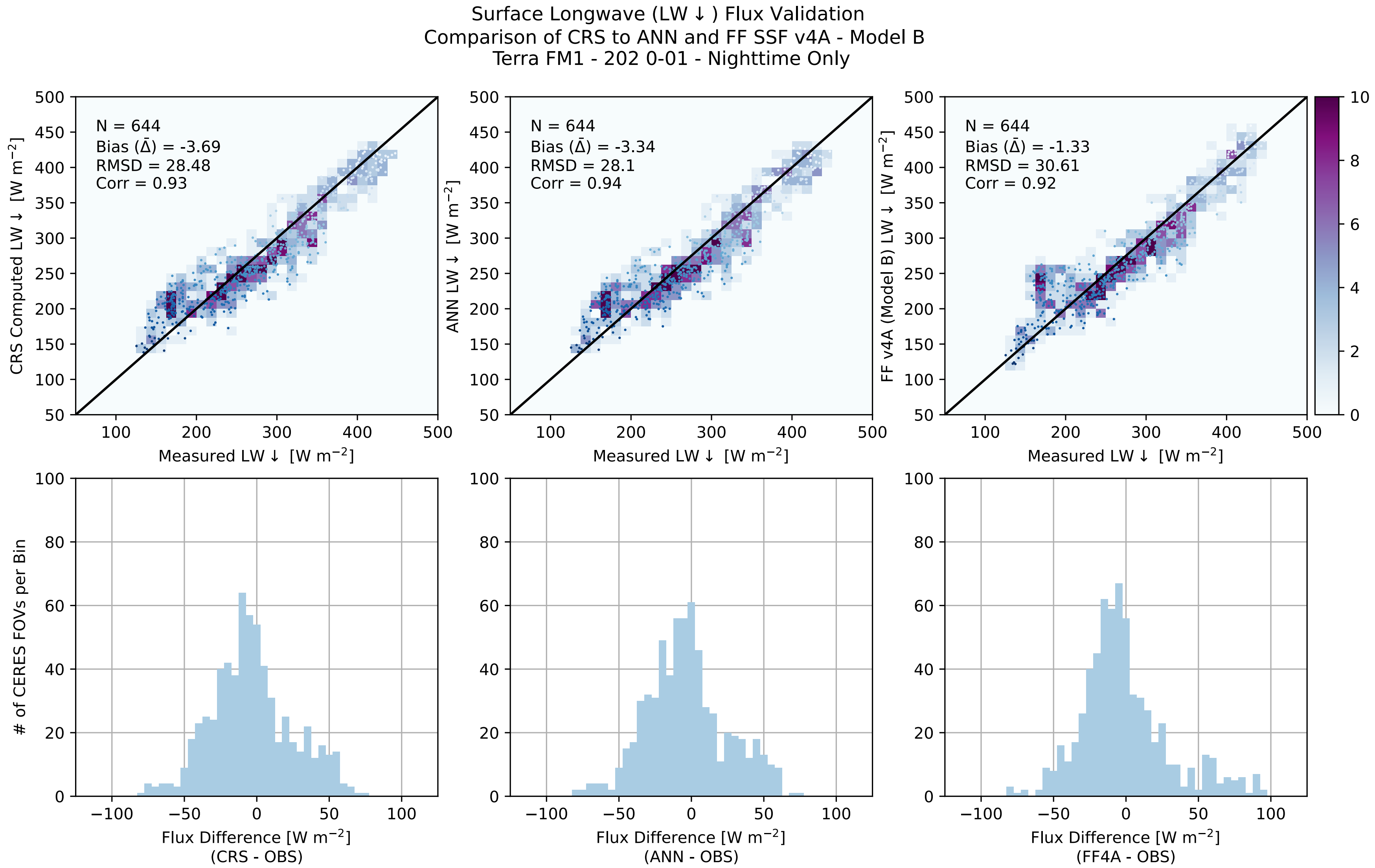
Surface Validations – LW Day Jan



Surface Validations – LW Day July



Surface Validations – LW Night Jan



Surface Validations – LW Night July