APPLYING DESIGN OF EXPERIMENTS TO NUMERICAL WEATHER PREDICTION

Jeffrey A. Smith, Judah L. Cleveland, Richard S. Penc, John W. Raby

Objective: Use Design of Experiments methods to identify which WRF-ARW parameter schemes drive the bias.

Consider a forecast as a map from an observed atmospheric state to some future state. Mathematically, we say:

$$f: x \rightarrow y$$

Where:

- f here is the NWP code WRF (Skamarock et al. 2008). We treat f as a set of functions where each $f_i \in f, i = 1, 2, \ldots, N$ is a specific configuration of the NWP code. N, although countable, is so large as to preclude a search of f using brute force methods.
- x represents the observed atmospheric state and/or initialization conditions
- y is a forecast.

Goal: Find $f_{(i)} \in f$ that minimizes the difference between the forecast and observed values at a specific point in time, a value we call "*bias*".

Data:

- The matched forecast – observation pair data elements produced by MET Point-Stat (National Center for Atmospheric Research 2016)
- The model employed uses WRF (version 3.8.1) in a 9/3/1 nest.
- Initialization 0.5° GFS (NOAA 2018d) forecast data providing gridded background fields with raw observations analyzed onto the background fields; 1/12° (~9 km) RTG high resolution SST (NOAA2018c); and 1 km NOHRSC SNODAS (NOAA 2018b) snow data when available and GFS snow data elsewhere.
- 6-h pre-forecast with observation nudging (12-18 UTC).
- Observation nudging during data assimilation uses TAMDAR (AirDat 2018) aircraft data and various MADIS (NOAA 2018a) datasets [standard surface observations, mesonet surface observations, maritime surface observations, profiler data, rawinsondes, and ACARS (aircraft) data (Mamrosh 1997)] (Dumais and Reen 2013; Dumais et al. 2015).
- The model top was set at 10mb for all runs.

Initial Design:

- Forty WRF runs
- Six different parameters (but the cumulous scheme was dropped from subsequent analysis because it was only applied to the outer domain)
- Two domains, San Diego and San Francisco
- Five case days
- Almost balanced and nearly orthogonal in order to be able to calculate the factor effects
- Some runs crashed and were unrecoverable so only 25 runs were actually completed.

Runs

Run	Domain	Case	Boundary Layer (PBL)[1]	Microphysics (MP)	Shortwave (SW)	Longwave (LW)	Land Surface (LSM)
1	SAN	02/07	MYNN2	Eta	RRTMG	New Goddard	CLMv4
2	SAN	02/07	SH	Goddard	Goddard	RRTM	NOAH
3	SAN	02/09	MYJ	Thompson	Dudhia	RRTM	5ITD
4	SAN	02/09	MYNN2	Lin	Dudhia	New Goddard	RUC
5	SAN	02/09	MYNN2	Lin	Goddard	New Goddard	RUC
6	SAN	02/09	SH	Lin	GFDL	RRTMG	NOAH
7	SAN	02/16	YSU	Goddard	GFDL	GFDL	CLMv4
8	SAN	02/16	MYJ	WSM5	Goddard	GFDL	NOAH
9	SAN	02/16	ACM2	Thompson	GFDL	New Goddard	5ITD
10	SAN	03/01	MYJ	WSM5	Dudhia	GFDL	NOAH
11	SAN	03/01	MYJ	Thompson	Goddard	New Goddard	RUC
12	SAN	03/05	YSU	Lin	GFDL	RRTM	CLMv4
13	SAN	03/05	SH	Eta	Goddard	RRTM	CLMv4
14	SFO	02/07	YSU	WSM5	Goddard	RRTMG	RUC
15	SFO	02/07	MYJ	Thompson	Dudhia	RRTMG	RUC
16	SFO	02/07	MYNN2	Thompson	RRTMG	New Goddard	RUC
17	SFO	02/09	MYNN2	Lin	Dudhia	New Goddard	CLMv4
18	SFO	02/16	YSU	Lin	GFDL	RRTM	RUC
19	SFO	02/16	MYJ	Lin	RRTMG	GFDL	5ITD
20	SFO	02/16	SH	Goddard	GFDL	RRTMG	RUC
21	SFO	03/01	MYNN2	Thompson	Goddard	GFDL	CLMv4
22	SFO	03/01	ACM2	WSM5	GFDL	RRTMG	5ITD
23	SFO	03/01	ACM2	Goddard	RRTMG	GFDL	RUC
24	SFO	03/05	YSU	Goddard	Dudhia	New Goddard	NOAH
25	SFO	03/05	YSU	Thompson	Dudhia	RRTM	5ITD

1: Includes Surface Layer Scheme, but only Boundary Layer Scheme noted.

Design

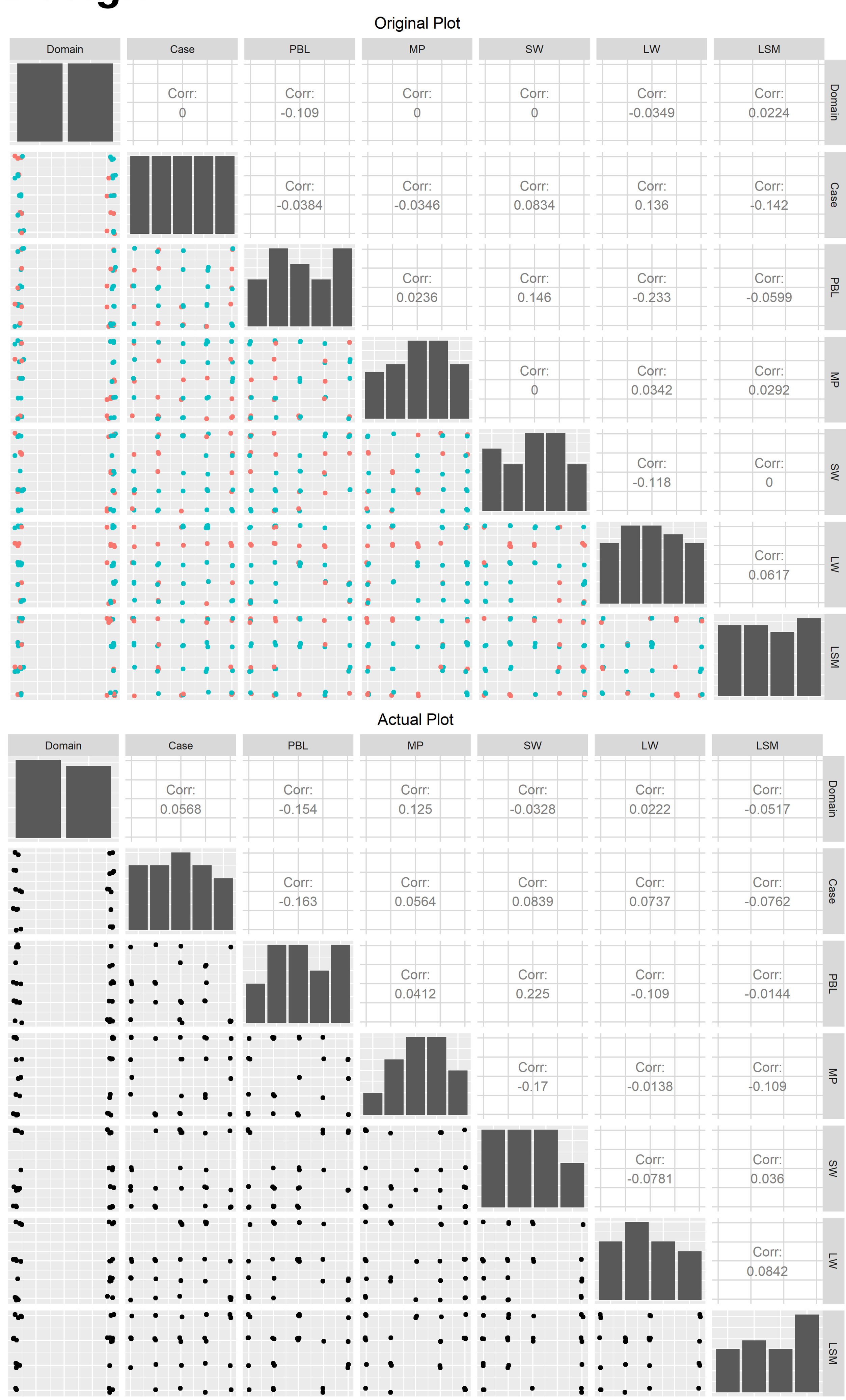


Figure 2: The correlation and density of the original design matrix (top) and the as-ran matrix (bottom). The orange points are the failed runs.

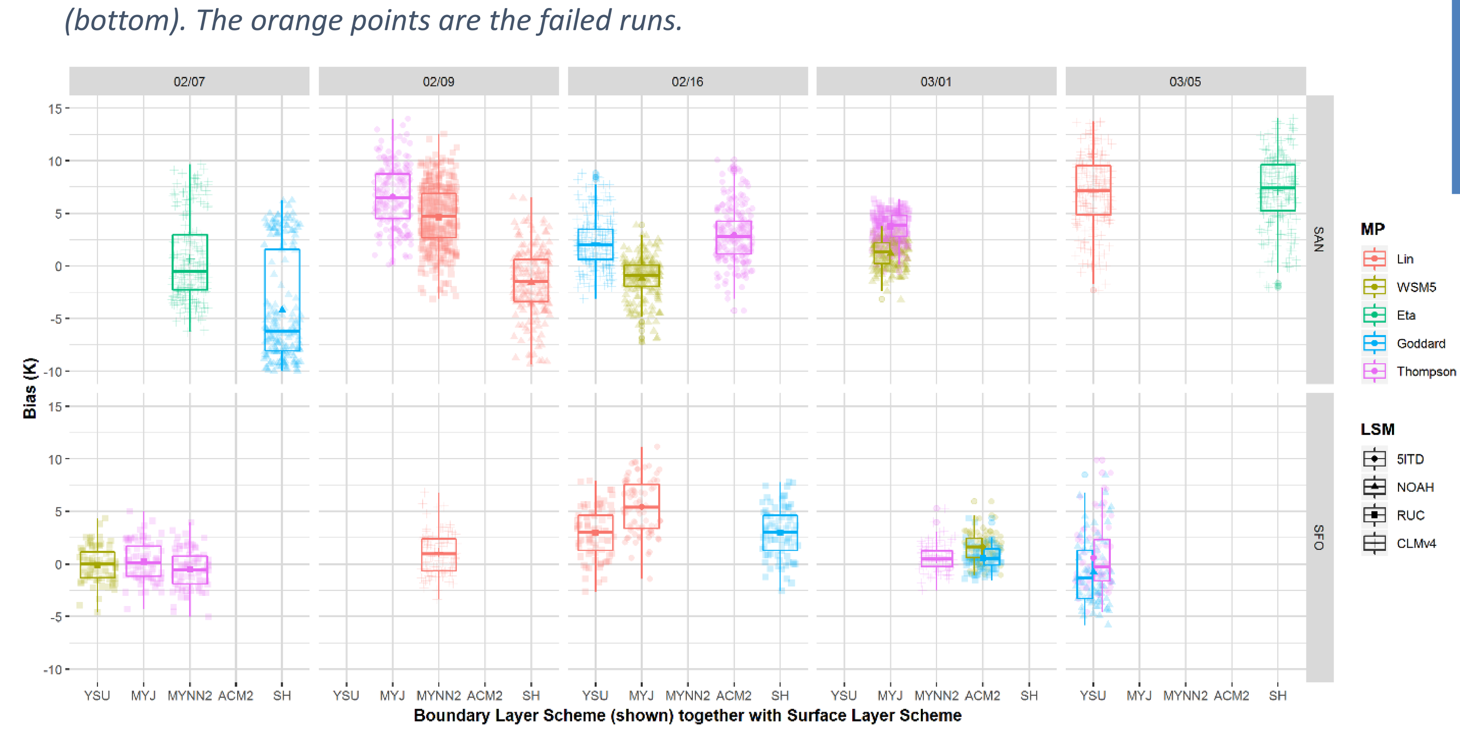
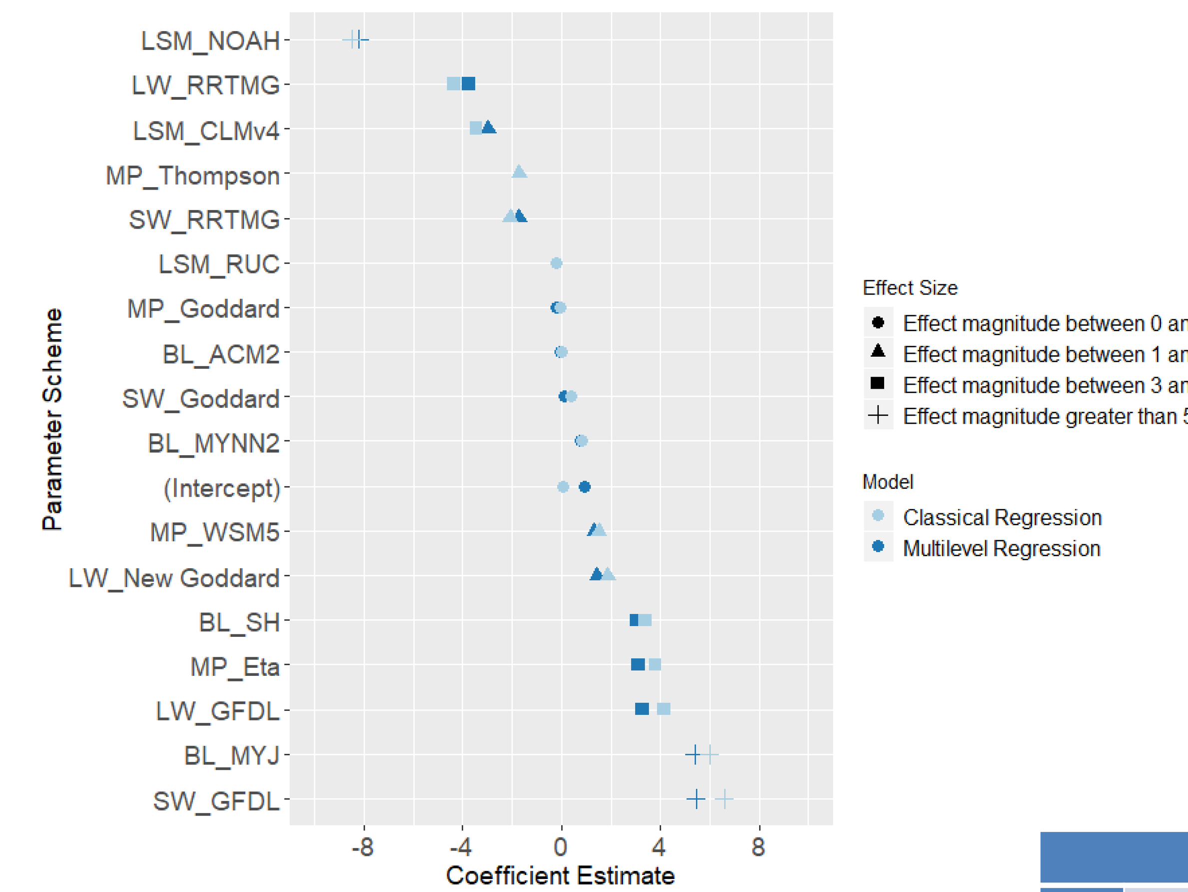


Figure 3: Dew point biases at 21Z by parameterization scheme organized by microphysics and land surface scheme. Mean values are indicated by a single glyph. Note: the 02/09 case for San Diego (SAN) using the MYNN2 Boundary Layer Schemes is runs 4 and 5 plotted together. The only feature distinguishing these runs is the choice of shortwave radiation scheme (Dudhia vs. New Goddard).

Results



Classical Linear Regression 21Z Model of Dew Point Bias (K)

		regression term	estimate	std.error	p.value
		(Intercept)	0.0973	0.7332	0.8945
	Domain	SFO	0.9140	0.5521	0.0980
	Case	02/09	1.4102	0.4748	0.0030
	Case	02/16	-4.0745	1.1998	0.0007
	Case	03/01	-2.3352	0.8661	0.0071
	Case	03/05	3.4801	0.5356	0.0000
Parameterization Scheme	PBL/SL	MYJ	6.6782	0.6919	0.0000
Parameterization Scheme	PBL/SL	MYNN2	1.8248	0.5837	0.0018
Parameterization Scheme	PBL/SL	ACM2	0.5538	0.4677	0.2365
Parameterization Scheme	PBL/SL	SH	3.8091	0.557	0.0000
Parameterization Scheme	Microphysics	WSM5	1.4266	0.5451	0.0089
Parameterization Scheme	Microphysics	Eta	2.9483	0.8422	0.0005
Parameterization Scheme	Microphysics	Goddard	0.0644	0.3913	0.8693
Parameterization Scheme	Microphysics	Thompson	-1.9915	0.3458	0.0000
Parameterization Scheme	Land Surface	NOAH	-8.7646	0.5907	0.0000
Parameterization Scheme	Land Surface	RUC	-0.2665	0.3715	0.4733
Parameterization Scheme	Land Surface	CLMv4	-3.402	0.6962	0.0000
Parameterization Scheme	Short-wave	Goddard	0.4272	0.3284	0.1935
Parameterization Scheme	Short-wave	RRTMG	-2.0962	0.5718	0.0003
Parameterization Scheme	Short-wave	GFDL	6.5121	1.0956	0.0000
Parameterization Scheme	Long-wave	RRTMG	-4.4738	0.6639	0.0000
Parameterization Scheme	Long-wave	New Goddard	1.4379	0.5461	0.0085
Parameterization Scheme	Long-wave	GFDL	3.7414	0.9288	0.0001

Multilevel Regression Model

		regression term	estimate	std.error
		(Intercept)	0.9850	1.3016
Groups	Domain	SAN	0.8851	0.1905
Groups	Domain	SFO	1.0848	0.1911
Groups	Case	02/07	0.6826	0.2409
Groups	Case	02/09	2.2686	0.2340
Groups	Case	02/16	-1.9293	0.2283
Groups	Case	03/01	-0.6237	0.2383
Groups	Case	03/05	4.5266	0.2515
Parameterization Scheme	PBL/SL	MYJ	5.9603	0.5551
Parameterization Scheme	PBL/SL	MYNN2	1.7837	0.5811
Parameterization Scheme	PBL/SL	ACM2	0.4893	0.4666
Parameterization Scheme	PBL/SL	SH	3.4258	0.5113
Parameterization Scheme	Microphysics	WSM5	1.2283	0.5339
Parameterization Scheme	Microphysics	Eta	2.0943	0.6491
Parameterization Scheme	Microphysics	Goddard	-0.0940	0.3838
Parameterization Scheme	Microphysics	Thompson	-1.9584	0.3435
Parameterization Scheme	Land Surface	NOAH	-8.3873	0.5516
Parameterization Scheme	Land Surface	RUC	-0.2085	0.3694
Parameterization Scheme	Land Surface	CLMv4	-2.8249	0.6123
Parameterization Scheme	Short-wave	Goddard	0.1383	0.2803
Parameterization Scheme	Short-wave	RRTMG	-1.6300	0.487
Parameterization Scheme	Short-wave	GFDL	5.1344	0.7264
Parameterization Scheme	Long-wave	RRTMG	-3.7351	0.4609
Parameterization Scheme	Long-wave	New Goddard	0.9102	0.4448
Parameterization Scheme	Long-wave	GFDL	2.7081	0.7093

References

1. AirDat, cited 2018: AirDat Real-Time TAMDAR Weather Data and Products. [Available online at http://www.airdat.com/.]
2. Dudhia, J., 2015: Overview of WRF Physics. 2015 Basic WRF Tutorial, National Center for Atmospheric Research.
3. Dumais, R. E., and B. P. Reen, 2013: Data assimilation techniques for rapidly relocatable weather research and forecasting modeling. Final Report ARL-TN-0546.
4. Dumais, R. E., Jr., B. P. Reen, J. A. Smith, D. I. Knapp, and H. Cai, 2015: Developing a WRF-based mixed variational and nudging data assimilation scheme for the US Army convection-scale nowcasting system. 95'th Annual AMS Meeting, 19th Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS) Paper 5.2.
5. Fisher, R. A., 1971: The Design of Experiments Eighth ed. Hafner Publishing Company, Inc. .
6. Gelman, A., and J. Hill, 2007: Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press.
7. Hu, X. M., J. W. Nielsen-Gammon, and F. Q. Zhang, 2010: Evaluation of Three Planetary Boundary Layer Schemes in the WRF Model. J. Appl. Meteorol. Climatol., 49, 1831-1844.
8. Mamrosh, R. D., 1997: The use of high- frequency ACARS soundings in forecasting convective storms. Weather and Forecasting Conference, American Meteorological Society.
9. National Oceanic and Atmospheric Administration (NOAA), cited 2018: Meteorological Assimilation Data Ingest System (MADIS). [Available online at http://madis.noaa.gov/.]
10. ——, cited 2018: National Operational Hydrologic Remote Sensing Center (NOHRSC), Snow Data Assimilation System (SNODAS). [Available online at https://www.nohrsc.noaa.gov/.]
11. ——, cited 2018: Real-Time Global Sea Surface Temperature (RTG_SST). [Available online at http://www.nco.ncep.noaa.gov/pmb/products/sst/.]
12. ——, cited 2018: Global Forecast System (GFS). [Available online at https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs.]
13. Skamarock, W. C., and Coauthors, 2008: A description of the advanced research WRF version 3. NCAR Technical Note NCAR/TN-475+STR.
14. Wu, C. F. J., and M. S. Hamada, 2009: Experiments Planning, Analysis, and Optimization 2nd ed. John Wiley & Sons, Inc., 716 pp.

APPROVED FOR PUBLIC RELEASE