Utah State University
BINGHAM RESEARCH CENTER

FINAL REPORT

IMPROVING WRF/CAMX MODEL PERFORMANCE USING SATELLITE DATA ASSIMILATION TECHNIQUE FOR THE UINTAH BASIN

Huy Tran
Trang Tran

DOCUMENT NUMBER: BRC_2021A
REVISION: ORIGINAL RELEASE
DATE: JANUARY 2021

(435) 722-1740 320 North Aggie Boulevard Vernal, Utah 84078 binghamresearch.usu.edu

Contents

Contents ii

List of Tables iii

List of Figures iv

1. Background and Introduction 1

2. Technical Approach 3

2.1. Regriding MODIS data to WRF model domain 4

2.2. Determining diurnal surface albedo from MODIS MCD43A1 dataset 5

2.3. WRF model configurations 6

2.4. CAMx model configurations 8

2.5. Implementation of MODIS data assimilation to WRF/CAMx model platform 9

2.6. Sensitivity simulations with Noah and PX LSMs 10

3. Results and Discussions 12

3.1. Impact of MODIS data assimilation on WRF model performance 12

3.2. Impact of MODIS data assimilation on CAMx model performance 17

3.3. Impacts of using sub-tiling in Noah land-use scheme on WRF model performance 18

3.4. Model performance sensitivity tests on Noah LSM versus Pleim-Xiu LSM with soil moisture nudging land-use scheme. 20

4. Conclusions and Future Works 22

5. References 23

List of Tables

Table 1. MODIS dataset applied in this study 3

Table 2. Parameters for 6SV model 6

Table 3. Grid definitions for WRF Preprocessor System 7

Table 4. WRF model grid configurations and topographical, land use and initial/boundary conditions 7

Table 5. Physics options used in the WRF 8

Table 6. Summary of CAMx Model Configurations 9

List of Figures

Figure 1. Diagram of processing MODIS data assimilation to WRF model..... 4

Figure 2. WRF one-way nested 12-4-1.33km domains (left) and zoom-in of 1.33km-domain with black circles indicating monitoring stations used for model performance evaluation (right). 8

Figure 3. Comparison of surface albedo obtained in simulations using the WRF default configuration (left) and using MODIS data assimilation (right)..... 12

Figure 4. Comparison of snow cover fraction (SNOWC) obtained in simulations using the WRF default configuration (left) and using MODIS data assimilation (right). 13

Figure 5. Comparison of snow cover fraction, snow water equivalent, and snow depth using the WRF default configuration (REF) and using MODIS data assimilation (MODIS). 14

Figure 6. Comparison of planetary boundary layer height (PBLH) and lapse rate using the WRF default configuration (REF) and using MODIS data assimilation (MODIS). 15

Figure 7. Comparison of overall performance of WRF using the WRF default configuration (REF) and using MODIS data assimilation (MODIS). 16

Figure 8. Comparison of photolysis rates simulated by CAMx using the default configuration (REF) and using MODIS data assimilation (MODIS). 17

Figure 9. Comparison of ozone at Ouray as simulated by CAMx using the default configuration (REF) and using MODIS data assimilation (MODIS). 18

Figure 10. February 2011 time series of Uintah Basin-wide average mixing ratio, temperature, wind speed and wind direction simulated by Noah-default (red), Noah-subtiling (blue) on 1.3km domain and observations (black)..... 19

Figure 11. February 2011 time series of Uintah Basin-wide average mixing ratio, temperature, wind speed and wind direction simulated by Noah-default (red), Pleim-Xiu with soil nudging (blue) on 1.3km domain and observations (black). 21

1. Background and Introduction

Ozone negatively impacts respiratory health, especially for those with lung diseases. During wintertime temperature inversion episodes, ozone in the Uinta Basin sometimes increases to levels that exceed the standard of 70 ppb set by the U.S. Environmental Protection Agency (EPA). Because of this, on 3 August 2018, portions of Uintah and Duchesne Counties below 6,250 feet in elevation were declared an ozone nonattainment area by EPA.

The Uinta Basin is one of only two places in the world that are known to experience wintertime ozone in excess of EPA standards (Wyoming's Upper Green River Basin is the other). Ozone forms in the atmosphere from reactions involving oxides of nitrogen (NO_x) and volatile organic compounds, and the majority of NO_x and organic compound emissions in the Uinta Basin are from oil and gas development. Inversion conditions trap these pollutants near ground level, increasing their concentrations and allowing them to generate ozone. The mix of pollutants during inversion episodes in the Uinta Basin is very different from those on Utah's urban Wasatch Front, leading to the formation of ozone rather than fine particulate matter ($PM_{2.5}$).

The number of ozone exceedance days and concentrations of ozone that occur each year are closely tied to meteorology. Years with persistent snow cover and high barometric pressure tend to have more days with strong winter inversions and high ozone. During inversion episodes, ozone concentrations tend to be higher at lower elevations, where inversion conditions are stronger and last longer. In the absence of snow cover and winter inversions, ozone concentrations in the Basin are similar to those in other rural, high-elevation locations around the western United States.

Since the first attempts were made to model winter ozone in the Uinta Basin, most modelers have found that the Weather Research and Forecasting (WRF) model, in its default configuration, is unable to capture the observed level of surface albedo (i.e., reflection of sunlight from the earth's surface). The maximum albedo able to be simulated by WRF is less than 0.7, whereas albedo above 0.8 has been observed at several locations in the Basin when snow cover is present. Underestimation of albedo, obviously, leads to underestimation of photolysis rates and, therefore, ozone formation. Because of this problem, most of the modeling studies for the Uinta Basin have implemented snow-related corrections to the WRF model, mostly manually, to increase the surface albedo outcome. Neemann, Crosman, Horel, and Avey (2015), for example, lowered the snow water equivalent threshold for certain vegetation types in a model grid cell so that the grid cell is allowed by the model to be fully covered by snow at lower snow depths. With such treatment, albedo higher than 0.8 was obtained over the entire Basin. The resultant modeled albedo, however, was unrealistically uniform. Most photochemical modelers subsequent to Neemann et al. have used this same approach (Matichuk et al., 2017; Tran, Tran, Mansfield, Lyman, & Crosman, 2018).

As we have performed model simulations for other winter episodes, we have encountered many difficulties in replicating high albedo values. Often, the albedo correction has been limited by a lack of observational data and lack of knowledge on surface conditions during the simulated episode. Overall, we have recognized that surface albedo corrections are: (1) performed on a case by case basis, and there is no standard method to perform the correction; (2) heavily reliant on knowledge and observations of the actual surface conditions during the simulated episode, and (3) Unrealistic in terms of the albedo distribution over the model domain.

These findings led us to the idea of utilizing satellite observational data to correct the model. Our objective in this work is to improve WRF performance with respect to surface albedo, snow cover, and vegetation fraction.

2. Technical Approach

Table 1 describes MODIS dataset that are utilized for this study. Figure 1 explains our method to assimilate MODIS data to WRF and CAMx model. All MODIS data was re-gridded from their original horizontal resolution (either 500m or 1 km) to the model resolution of 1.3 km, using the Earth System Modeling Framework. The Second Simulation of a Satellite Signal in the Solar Spectrum vector code (6SV) was utilized to construct lookup table for diffuse radiation fraction, which is then used in the Bidirectional Reflectance Distribution Function (BRDF) albedo determination to determine diurnal values of snow and spectrum shortwave albedo.

Table 1. MODIS dataset applied in this study

Dataset	***Descriptions***	***Horizontal Resolution***	***Temporal Resolution***
MCD43A1[1]	MODIS Terra + Aqua BRDF/Albedo Model Parameters Daily L3 Global - 500m V006	500 m	Daily
MCD19A1[2]	MODIS Terra + Aqua Land Surface BRF Daily L2G Global 500m, 1km and 10km SIN Grid V006	500 m	Daily
MCD19A2[3]	MODIS Terra + Aqua Land Aerosol Optical Thickness Daily L2G Global 1km SIN Grid V006	1 km	Daily
MOD10A1/MYD10A1[4]	MODIS Terra + Aqua Snow Cover Daily L3 Global 500m Grid. Only MOD10A1 is used.	500 m	Daily
MCD15A3H[5]	MODIS Terra + Aqua Leaf Area Index/FPAR 4-Day L4 Global 500 m (4 days composite)	500 m	Daily

Except the MODIS Leaf Area Index (LAI) which replaces the original and climatology seasonal LAI variable in WRF's input file, all other MODIS derived quantities are incorporate to WRF surface

[1] https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MCD43A1
[2] https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MCD19A1
[3] https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MCD19A2
[4] https://modis-snow-ice.gsfc.nasa.gov/?c=MOD10A1
[5] https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MCD15A3H

nudging module before entering WRF core model. Finally, we used the *wrfcamx* module to processes WRF outputs and prepare prepare meteorology inputs in binary format form CAMx.

Certain WRF and CAMx model source codes had to be modified to allow the MODIS data assimilation. One highlight in our data assimilation approach is that by assimilating MODIS data through WRF surface nudging module (*wrfsfdda*), we are able to utilize all standard options provide by this module to control the data assimilation such as turning the data assimilation on and off, how frequent WRF estimates is replaced with MODIS data, weighting factor, and ramping time, etc.

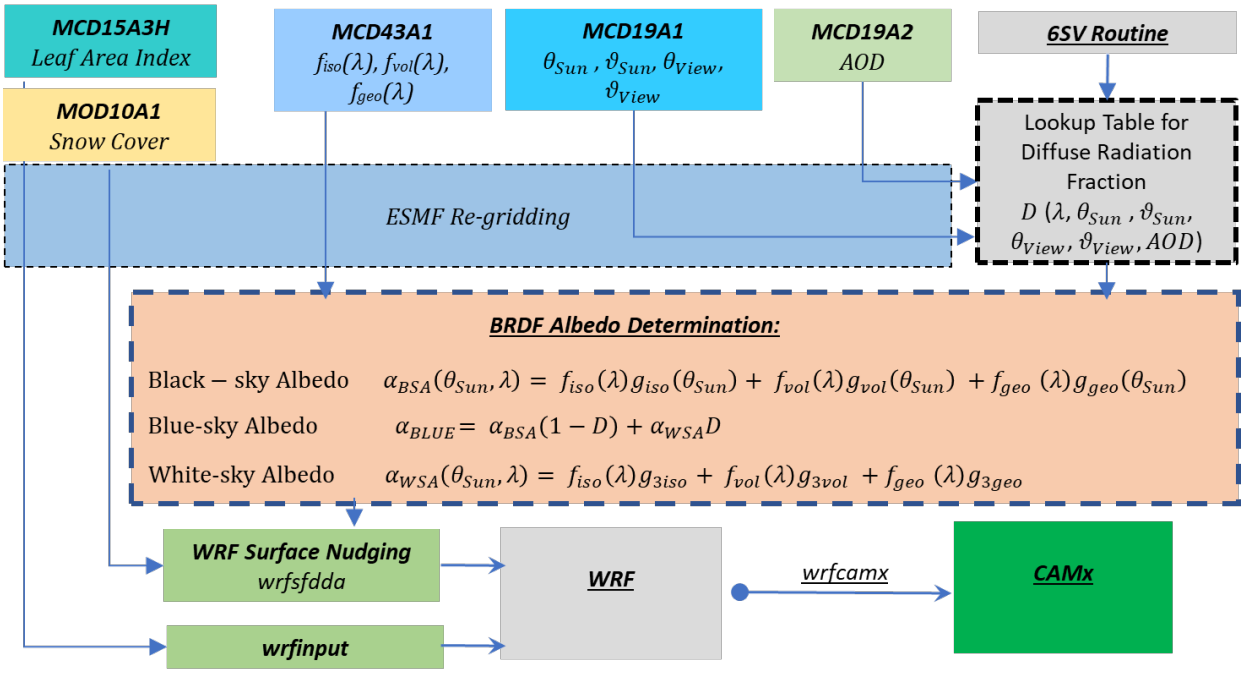

Figure 1. Diagram of processing MODIS data assimilation to WRF model

Details on the processes of processing MODIS data are dicussed in the below sections.

2.1. Regriding MODIS data to WRF model domain

We applied the Earth System Modeling Framework (ESMF[6]) to with NCAR Command Language (NCL[7]) front end to re-grid MODIS data from their original horizontal resolution (Table 1) to WRF's horizontal resolution of 1.3 km (see section 2.2). Several interpolation methods are

[6] https://earthsystemmodeling.org
[7] https://www.ncl.ucar.edu/Applications/ESMF.shtml

supported by ESMF including "*bilinear*", "*patch*", "*conserve*" and "*neareststod/nearestdtos*". The last method is nearest neighbor method which associate a point in source dataset to closet point in destination dataset, or vice versa. Our sensitive test suggests "*conserve*" method produce the best preservation of MODIS data in WRF domain. Since resolution of MODIS grid is higher than that of WRF, minimum interpolation is involved in the downscaling of MODIS data to WRF grid which explains for the better performance of "*conserve*" method over other interpolation methods.

2.2. Determining diurnal surface albedo from MODIS MCD43A1 dataset

MODIS MCD43A1 provides weighting parameters associated with the ***RossThickLiSparseReciprocal*** bidirectional reflectance distribution function (BRDF) that best describe the anisotropy of each pixel[8]. These parameters can be used to estimate black-sky (α_{BSA}) and white-sky (α_{WSA}) albedo for any solar zenith angle (θ) and selected bin of wavelength (λ) using the following equations:

$$\alpha_{BSA}(\theta_{Sun},\lambda) = f_{iso}(\lambda)g_{iso}(\theta_{Sun}) + f_{vol}(\lambda)g_{vol}(\theta_{Sun}) + f_{geo}(\lambda)g_{geo}(\theta_{Sun}) \quad (1)$$

$$\alpha_{WSA}(\lambda) = f_{iso}(\lambda)g_{3iso} + f_{vol}(\lambda)g_{3vol} + f_{geo}(\lambda)g_{3geo} \quad (2)$$

Blue sky albedo (α_{BLUE}), which refers to albedo calculated under real-world conditions with a combination of both diffuse and direct lighting based on atmospheric and view-geometry conditions, could be then determine by:

$$\alpha_{BLUE} = \alpha_{BSA}(1-D) + \alpha_{WSA}D \quad (3)$$

Parameters on the right hand side of equations (1) and (2) are provided by MCD43A1. The Second Simulation of a Satellite Signal in the Solar Spectrum – Vector (6SV[9]) model is used to calculate the split factor *D* between diffuse and direct lighting. Although 6SV simulates diffuse and direct effect under different atmospheric and surface conditions, we specified parameters in the model as shown in Table 2 with assumptions that these parameters best represent winter conditions in the Uintah Basin.

[8] https://www.umb.edu/spectralmass/terra_aqua_modis/v006/mcd43a1_brdif_albedo_model_parameters_product
[9] http://6s.ltdri.org/pages/manual.html

Table 2. Parameters for 6SV model

Parameters	***Values***	***Descriptions***
Atmospheric Profile	3	Midlatitude Winter
Aerosol Model	1	Continental Model
Sensor level	1000 km	Set to satellite altitude
Ground reflectance type	0	Homogenous surface
Directional effect	1	Directional effect is considered
Directional effect model	10	MODIS operational BDRF
Weigh factors for MODIS BDRF	1.0, 0.77, 0.2	Weights for Lambertian kernel, RossThick kernel, LiSparse kernel in MODIS BDRF
Atmospheric correction mode	-1	No atmospheric correction

Other inputs of 6SV model including view angles (zenith θ_{View} and azimuth ϑ_{View}) are taken from MCD19A1, and aerosol optical depth (AOD) is taken from MCD19A2. Although MCD19A2 is a daily dataset, its coverage is sporadic (i.e., not all pixels covering the model domain is available in each scan). Therefore, a monthly average AOD is derived from MCD19A2 to produce a complete coverage over the WRF model domain. We used 6SV model to create a lookup table for *D* values for wide range of wave length, solar angles, view angles, and aerosol optical depth (λ, θ_{solar}, ϑ_{solar}, θ_{view}, AOD).

2.3. WRF model configurations

For reference (REF) WRF simulation, we used the well-tested grid definitions, topography, land use, boundary/initial conditions and physics options as described in Tables 3, 4, 5 and Figure 2. Basically, these domain setups and model configurations followed 2017 Air Resource Management Strategy (ARMS)'s WRF developed by USU for Utah-BLM to address impacts of oil and gas emissions on air quality of the State of Utah and Uinta Basin. However, we turned off all nudging application that were used in 2017 ARMS to keep REF configurations as "default" for later comparison with MODIS assimilation.

The modeling episode was Feb 1st to 28th 2011 divided into four 5.5-day batches and one 8.5-day batch. The first 12 hours of each batch served as spin-up time and were discarded from the analysis. This period covered three distinct inversion events happening around Feb 4th-9th, Feb 12th -17th and Feb 21st -25th.

Table 3. Grid definitions for WRF Preprocessor System

Parameter name	Parameter value
Projection	Lambert conformal
Reference latitude	40 N
Reference longitude	-97 W
truelat1	33
truelat2 =	45
stand_lon =	-97
ref_x	190.5
ref_y	90.5

Table 4. WRF model grid configurations and topographical, land use and initial/boundary conditions.

	Domain 1	Domain 2	Domain 3 (*)
Grid Size (x,y)	201 x 191	253 x 253	298x322
Vertical levels	37	37	37
Vertical coordinates	Terrain-following Eta (non-hybrid)	Terrain-following Eta (non-hybrid)	Terrain-following Eta (non-hybrid)
Vertical grid spacing	12-16 m in boundary-layer	12-16 m in boundary-layer	12-16 m in boundary-layer
Horizontal resolution (km)	12	4	1.33
Model time step (s)	25	8.33	2.77
Topographic dataset	USGS GTOPO30	USGS GTOPO30	USGS GTOPO30
Land use data set	NLCD2011 modified 9s	NLCD2011 modified 9s	NLCD2011 modified 9s
Initial and boundary conditions	NAM-12km	Continuous updates nested from 12km domain	Continuous updates nested from 4km domain
Top and Bottom Boundary Conditions	- Top: Rayleigh dampening for the vertical velocity - Bottom: physical, non free-slip option	- Top: Rayleigh dampening for the vertical velocity - Bottom: physical, non free-slip option	- Top: Rayleigh dampening for the vertical velocity - Bottom: physical, non free-slip option
Veg parm table variables modified for winter simulations (*)	SNUP, MAXALB	SNUP, MAXALB	SNUP, MAXALB
Snow cover initialization	- SNODAS	- SNODAS	- SNODAS

Table 5. Physics options used in the WRF

WRF Treatment	Option Selected
Microphysics	Thompson
Longwave Radiation	RRTMG
Shortwave Radiation	RRTMG
Land Surface Model (LSM)	NOAH
Planetary Boundary Layer (PBL) scheme	MYJ
Cumulus parameterization	Kain-Fritsch in the 12 km domains. None in the 4 and 1.3 km domain.

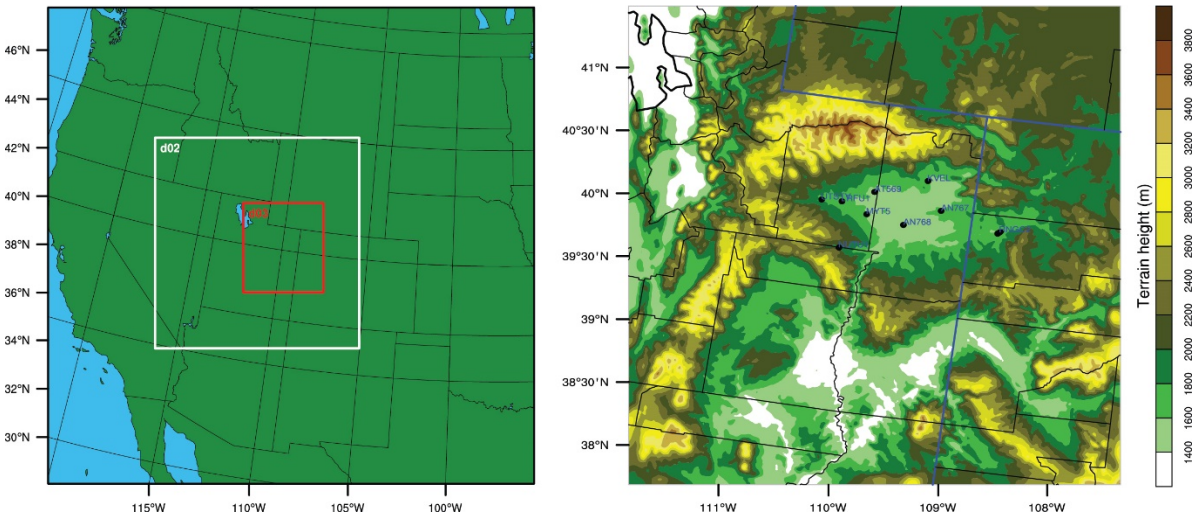

Figure 2. WRF one-way nested 12-4-1.33km domains (left) and zoom-in of 1.33km-domain with black circles indicating monitoring stations used for model performance evaluation (right).

2.4. CAMx model configurations

We used the Comprehensive Air Quality Model (CAMx) version 6.5[10] in the same configuration of the 2017 Air Resource Management Strategy (ARMS) modeling study (Mansfield, Tran, & Tran, 2020) in its finest domain (1.3 km), which is presented as ***d03*** in Figure 2. Table 5 provides details on the CAMx domain configurations. The Utah Department of Environmental Quality 's emission inventory for oil and gas production in the Uintah Basin in 2014 (UBEI2014) was

[10] http://www.camx.com/files/camxusersguide_v6-50.pdf

processed using Sparse Matrix Operator Kerner Emissions (SMOKE) version 4.5[11] and projected to modeling episode Feb 1st to 28th 2011 with scaling factors derived from oil and gas production rates in 2014 and 2011.

Table 6. Summary of CAMx Model Configurations

Science Options	Configuration
Model Code Version	CAMx V6.5
Horizontal Grid	1.33 km (298x322)
Vertical Grid	25 vertical layers
Initial and Boundary conditions	Processed from ARMS2017
Boundary Conditions	12 km BCs from WAQS 2011b
Land-use Data	Land-use fields from meteorological model
Photolysis Rate Preprocessor	TUV V4.8 (Clear-sky photolysis rates from TOMS data)
Gas-phase chemistry	CB6r4
Aerosol-phase	CF (coarse and fine mode aerosols)
Diffusion Scheme	
Horizontal-grid	Explicit horizontal diffusion
Vertical-grid	K-theory 1st-order closure
Deposition Scheme	
Dry deposition	ZHANG03 with modifications based on (Helmig, Ganzeveld, Butler, & Oltmans, 2007)
Wet deposition	CAMx specific formulation
Numerical Solvers	
Gas-phase chemistry	Euler Backward Iterative (EBI) solver
Horizontal advection	Piecewise Parabolic Method (PPM)
Vertical advection	Implicit scheme with vertical velocity update

2.5. Implementation of MODIS data assimilation to WRF/CAMx model platform

2.5.1. WRF model modification

We made modification to certain WRF's source codes to allow assimilation of MODIS's surface albedo (SW_ALB) and snow cover (SNOWC) to the model. Particularly, we implemented the assimilation to WRF through WRF's surface assimilation module (*wrfsfdda*) which allow WRF modeler to enable or disable MODIS data assimilation as needed and to control assimilation frequency (e.g., every 1, 2, 6 hours or longer). Being components of w*rfsfdda,* the assimilation

[11] https://www.cmascenter.org/smoke/documentation/4.5/manual_smokev45.pdf

of SW_ALB and SNOWC are also controlled by other standard options in WRF *fdda* module. New variables (see Appendix 1) are added to input file of *wrfsfdda* module (*wrfsfdda_d<domain>).* Note that unlike hourly SW_ALB which has diurnal cycle after being processed as described in section 2.2, the hourly SNOWC is aggregated from MODIS regridded daily snow cover data which means all hours within a same date have identical SNOWC value.

MODIS SNOWC and LAI are also written to WRF input files (*wrfinput_d<domain>*). MODIS LAI replace the default climatology LAI in *wrfinput* that often is incorrect and in coarse resolution. In WRF simulation, LAI is read in from *wrfinput* as initial condition and evolves with other simulated parameters such as snow depth and snow cover.

2.5.2. CAMx model modification

Although CAMx reads in most of meteorological inputs from WRF output files, CAMx determines its own ALBEDO and SNOWC as the functions of snow water equivalent, snow age and land use type. This means improvements made in WRF albedo and snow cover are not carried over to CAMx in standard model configuration. In order to force CAMx to use MODIS-assimilated data from WRF, modifications to CAMx code has to be made (see Appendix 2). These include modification made to *wrfcamx* module so that MODIS ALB and SNOWC are written to CAMx inputs.

In standard simulation and at the begin of the simulation, if there is lack of information of snow age, CAMx assumes albedo of near fresh snow condition, and therefore, CAMx-albedo may be higher than that estimated by WRF before decreases gradually as the snow ages. This implies that albedo calculated by CAMx model in standard configuration could be higher than the albedo value taken from WRF inputs.

2.6. Sensitivity simulations with Noah and PX LSMs

The Noah LSM have been using in our WRF simulations for Uintah Basin Ozone Study (UBOS) and the Air Resource Management Strategy (ARMS) modeling study (Mansfield et al., 2020). In standard configuration which is currently applied in our WRF simulations, Noah LSM assumes the dominant land cover type in each grid-cell as the land cover for that grid-cell and therefore ignores discard the subgrid-scale heterogeneity of land cover. Since WRF v3.6, sub-tiling option is made available so that modeler can specify number of land cover type to be considered by Noah LSM[12].

[12] https://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3.9/ARWUsersGuideV3.9.pdf

The PX LSM uses factional land cover type and thus is able to utilize the high-resolution land cover data such as the 500 m MODIS or the 30 m National Land Cover Database (NLCD) (Ran et al., 2015). More noticeably, PX LSM was developed for retrospective simulations and it allows soil nudging option where observational 2-m temperature and water vapor mixing ratio are utilized for deep soil moisture and temperature nudging[13].

Because of the distinct difference between Noah and PX LSMs, we will perform WRF sensitivity simulations using Noah LSM with sub-tiling option (*sf_surface_mosaic = 1*) and PX LSM with reanalysis nudging. We will compare WRF performances in these two configurations and compare them with WRF performance with standard Noah configuration (*sf_surface_mosaic = 0*)

Through these sensitivity simulations, we will examine whether WRF model's performance improves with different LSM configurations, and with high resolution land cover data.

[13] https://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3.9/ARWUsersGuideV3.9.pdf

3. Results and Discussions

3.1. Impact of MODIS data assimilation on WRF model performance

Figure 3 compares surface albedo simulated in the WRF default configuration (referred as **REF** hereafter) to WRF simulations that included MODIS data assimilation (referred as **MODIS** hereafter) as an average across 1 through 28 February 2011. In comparison to REF, MODIS not only gives higher albedo in the Uinta Basin, but the spatial distribution of albedo is more realistic, as one may expect when observed from space. Comparison of snow cover fraction in Figure 4 shows that, while snow cover in the Uinta Basin is somewhat lower in MODIS than in REF, its distribution is more realistic. In REF, snow cover is fairly consistent and at a high value at many locations in the model. Initialization of the model with the SNODAS data product (NOHRSC, 2004) could be one of the reasons for this snow cover distribution in REF.

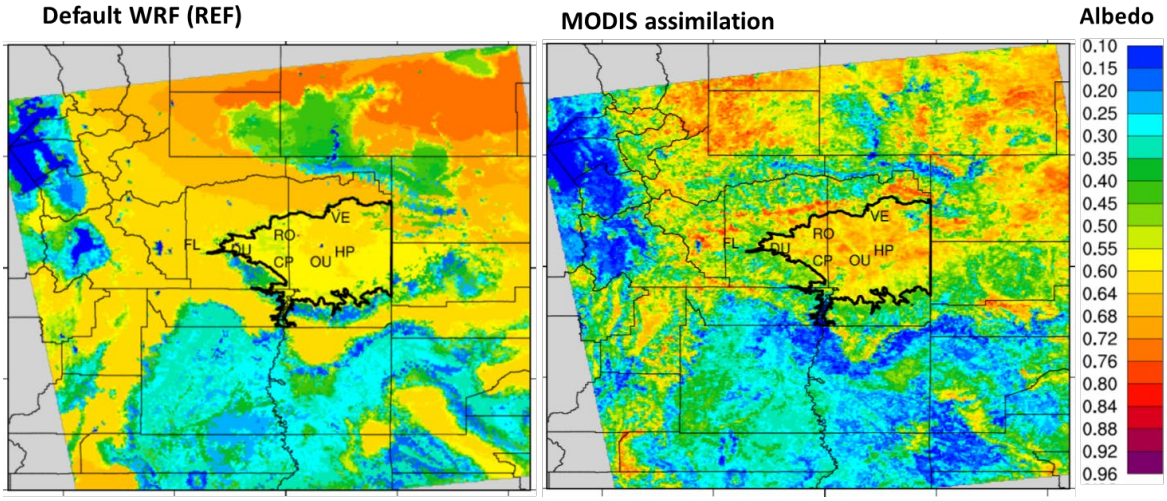

Figure 3. Comparison of surface albedo obtained in simulations using the WRF default configuration (left) and using MODIS data assimilation (right).

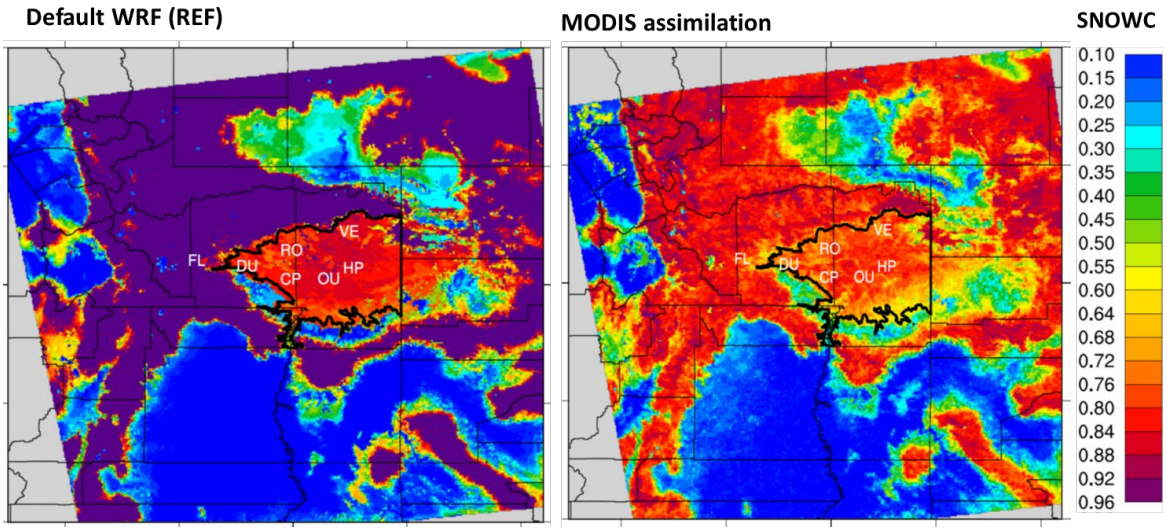

Figure 4. Comparison of snow cover fraction (SNOWC) obtained in simulations using the WRF default configuration (left) and using MODIS data assimilation (right).

Figure 5 compares the time evolution of WRF estimates of snow depth, snow water equivalent, and snow cover between REF and MODIS, as an average over 6 locations in the Uinta Basin. As seen in this figure, every time WRF was re-initialized with SNODAS data (indicated by green bars), snow depth, snow water equivalent, and snow cover sharply increased. This behavior was observed in both REF and MODIS, as they were both periodically re-initialized with SNODAS data. All the three parameters tended to decay faster in REF than in MODIS. Also, from 11 February onward, there were several storms that brought extra snow to the Basin, and WRF consistently estimated more snow in REF than in MODIS. The zigzag-like pattern of snow cover in MODIS in Figure 5 is the effect of data assimilation as snow cover fraction was directly taken from the MODIS dataset. Snow cover was typically higher in REF than in MODIS, which was likely caused by SNOWDAS data reinitialization. Our conclusion is that, for extended simulations and without reinitialization, snow-related characteristics were retained longer in MODIS than in REF.

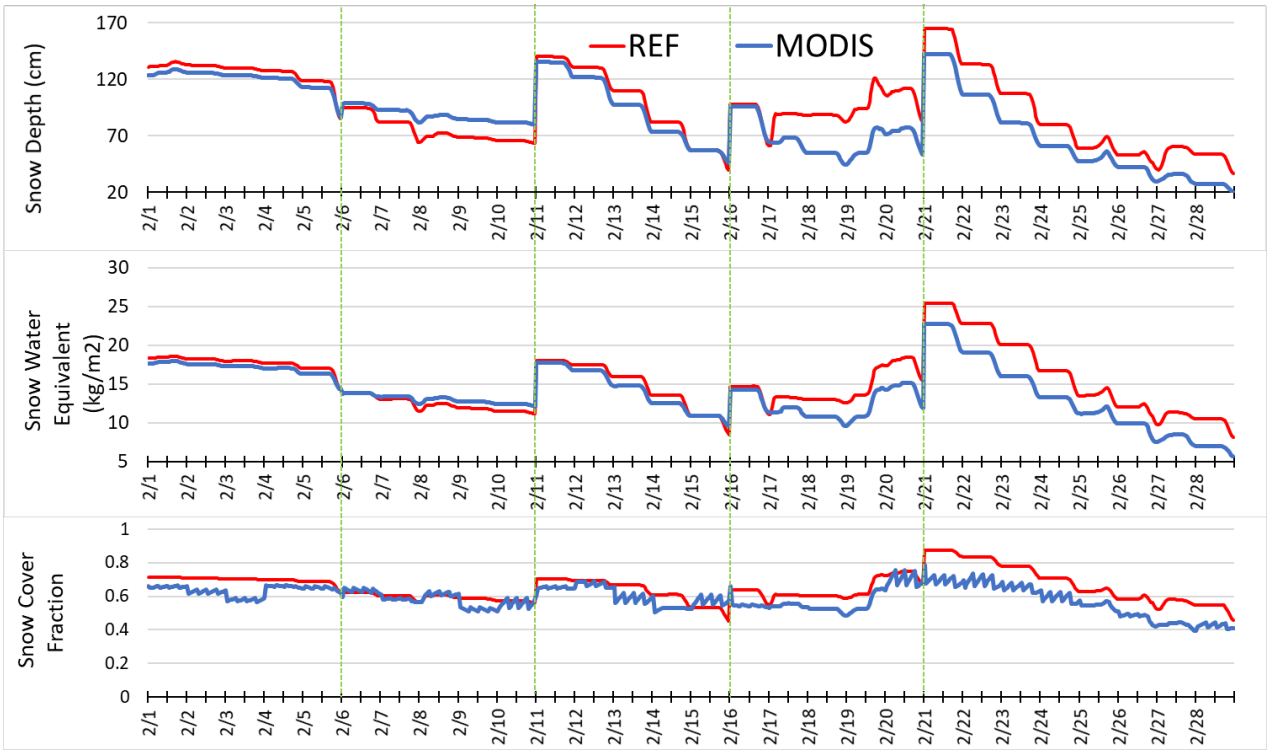

Figure 5. Comparison of snow cover fraction, snow water equivalent, and snow depth using the WRF default configuration (REF) and using MODIS data assimilation (MODIS). Green bars show periods where WRF reinitialized snow characteristics using the SNOWDAS dataset.

Planetary boundary layer height is the height of the mixed layer of the atmosphere that is in direct contact with the earth's surface. The lapse rate is a measure of inversion strength. We found no meaningful differences in the planetary boundary layer height and lapse rate between REF and MODIS at Ouray (Figure 6). Boundary layer height tended to be slightly shallower in MODIS compared to REF, but the lapse rate was slightly weaker in MODIS.

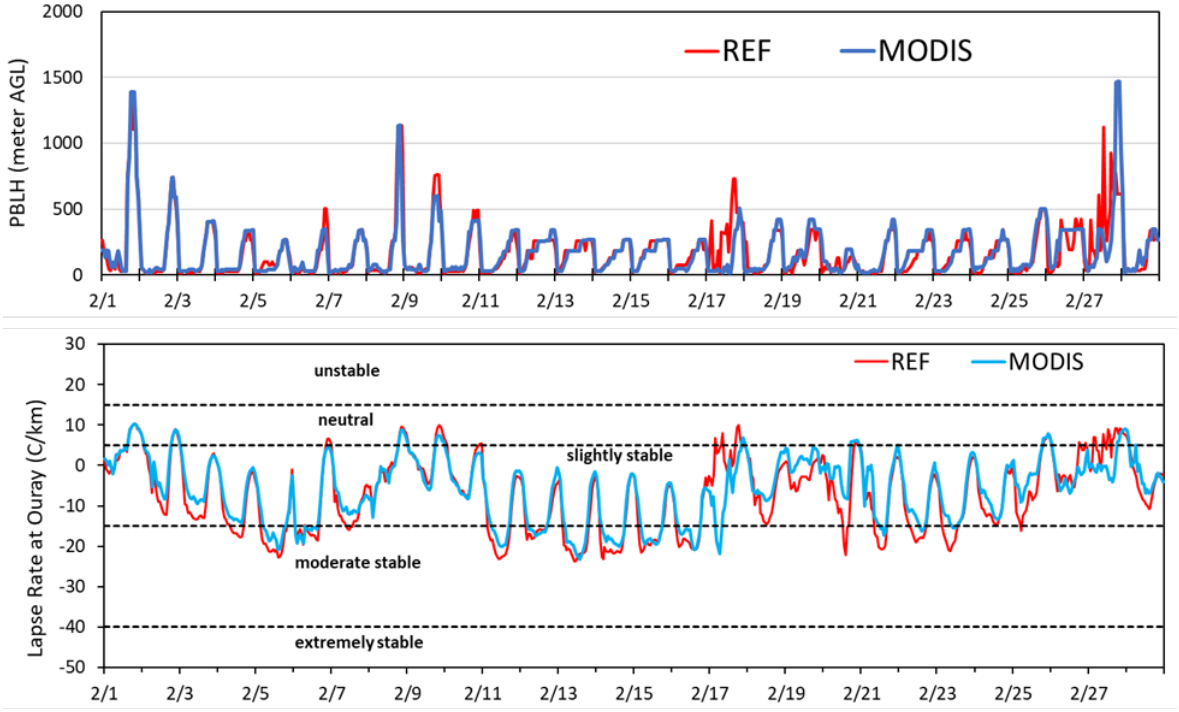

Figure 6. Comparison of planetary boundary layer height (PBLH) and lapse rate using the WRF default configuration (REF) and using MODIS data assimilation (MODIS).

The MODIS data assimilation technique, overall, does not result in better WRF performance, as illustrated in Figure 7. Skill scores for WRF output from the REF and MODIS simulation are presented in the figure, including mean absolute error (MAE), bias (BIAS), and index of agreement (IOA). Lower values of MAE and BIAS and higher IOA indicate better performance. Overall, the MODIS simulation tended to have lower skill scores than REF for some quantities. One exception was the simulation of wind speed. On some days when storm fronts passed through the Basin (indicated in the figure by green rectangles), the REF simulation estimated unrealistically high wind speed, whereas wind speeds were lower and closer to observed values in the MODIS simulation. This shows that assimilating MODIS albedo and snow cover data improved simulations of wind speed, though we don't have a clear explanation for this behavior.

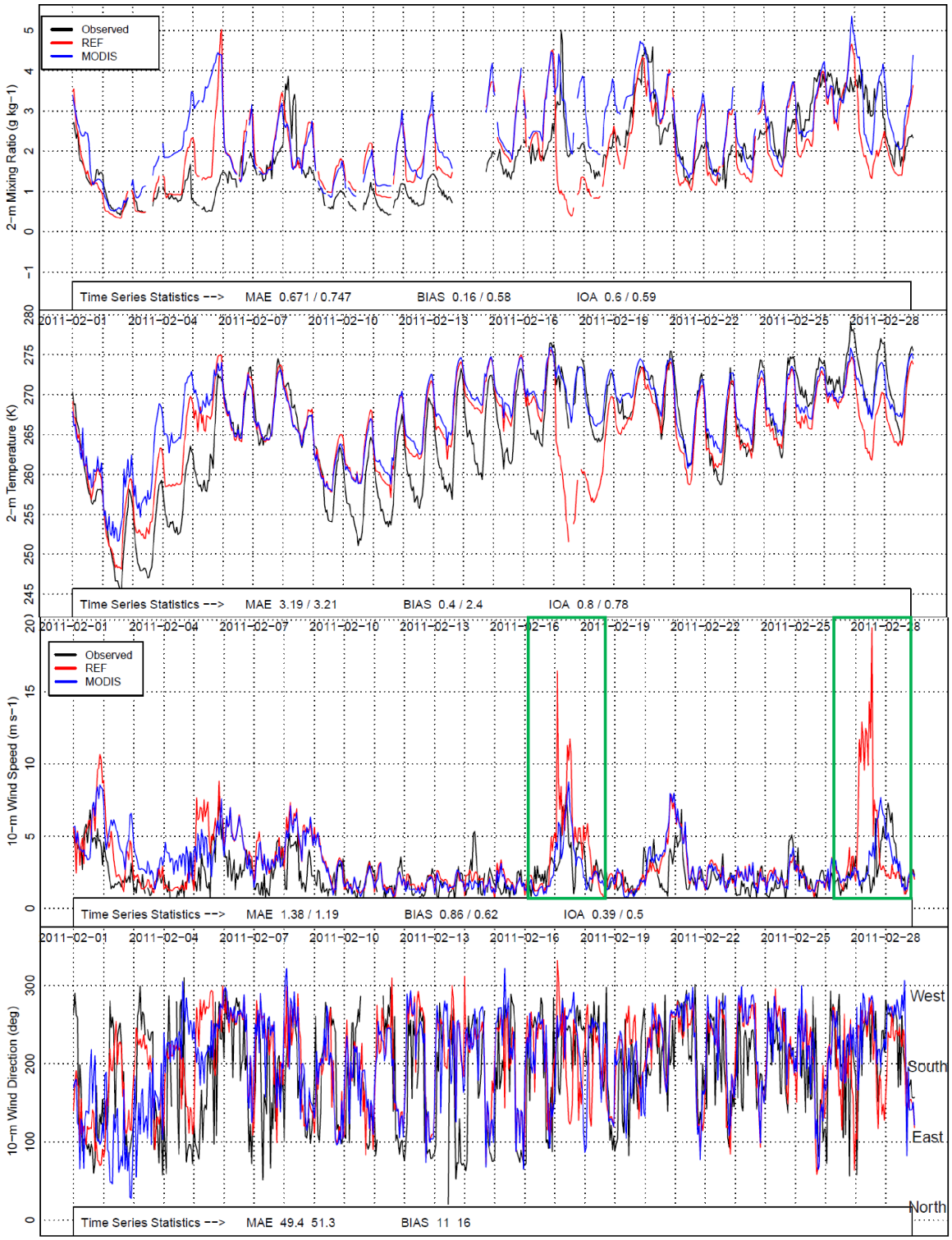

Figure 7. Comparison of overall performance of WRF using the WRF default configuration (REF) and using MODIS data assimilation (MODIS). Observed values are also shown. The panels show water vapor mixing ratio, temperature at 2 m above ground, wind speed at 10

m, and wind direction at 10 m (in descending order). Skill scores are shown at the bottom of each box-plot (REF/MODIS; see text for explanation of scores). The green boxes indicate days during which storm fronts passed through the Basin.

3.2. Impact of MODIS data assimilation on CAMx model performance

Figure 8 compares the photolysis rate of ozone at Ouray in REF and MODIS. From 1 to 10 February, no new snow was simulated in the model, and since snow tends to decay quicker in the REF simulations than in the MODIS simulations, the photolysis rate in the MODIS simulations eventually exceeds that in REF. However, on 11 February and afterward (indicated as the part of the figure to the right of the green vertical bar), as new snow was estimated in WRF, and because of the fresh-snow effect in CAMx, the photolysis rate in REF was consistently higher than in MODIS. Overall, we found that the higher albedo we obtained in WRF with the MODIS satellite data assimilation technique did not necessarily translate to a higher photolysis rate in the CAMx model. A positive effect of MODIS data assimilation on the CAMx photolysis rate could be expected if the simulation included an extended period of time where no new snow was simulated. Additionally, an on-line coupling meteorology-and-chemistry model such as WRF-Chem (see Section 5) may benefit more from MODIS data assimilation than a decoupled model platform such as WRF-CAMx.

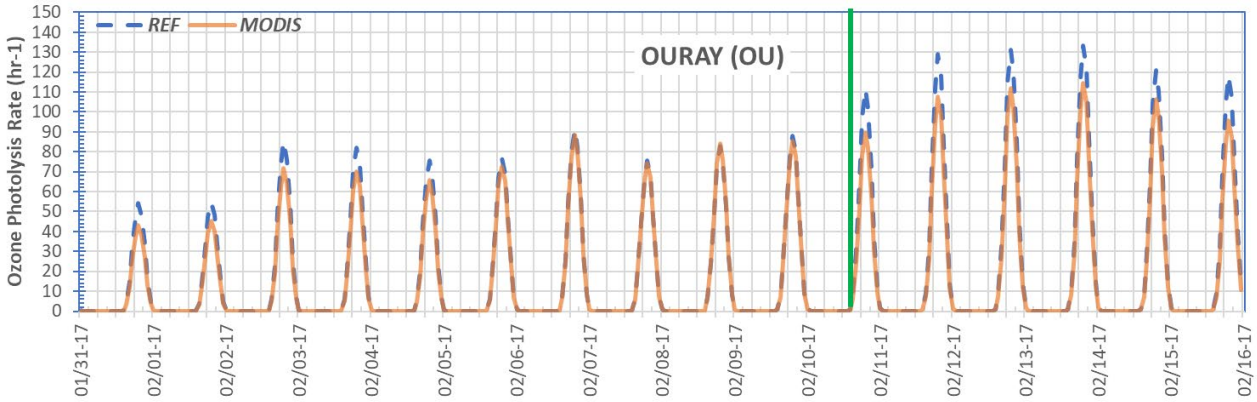

Figure 8. Comparison of photolysis rates simulated by CAMx using the default configuration (REF) and using MODIS data assimilation (MODIS). The green bar represents a new snow event.

As expected, we found no significant difference in ozone between the REF and MODIS simulations. Simulated ozone, as shown in Figure 9 for Ouray, behaved in agreement with the photolysis rate shown in Figure 8. Neither of the two scenarios was able to simulate ozone at levels that were actually observed during this episode, probably because of challenges with model emissions and chemistry, which also impact winter ozone.

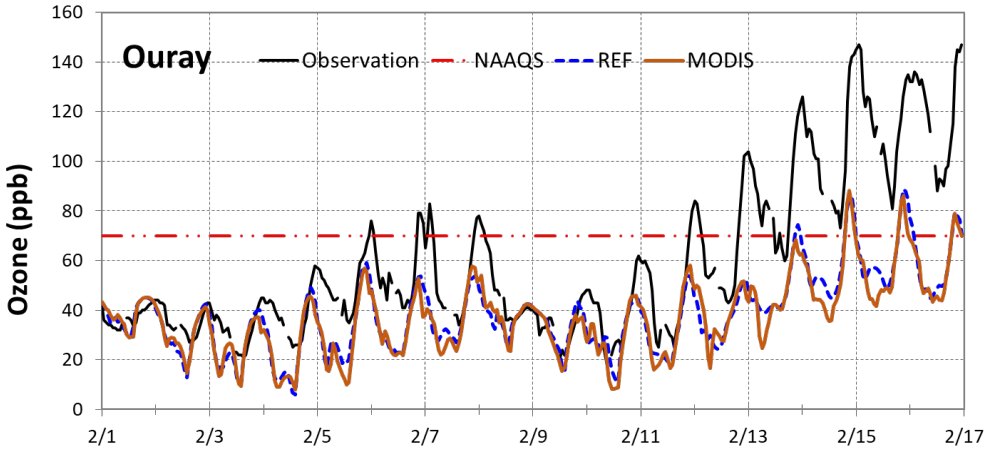

Figure 9. Comparison of ozone at Ouray as simulated by CAMx using the default configuration (REF) and using MODIS data assimilation (MODIS). **The U.S. EPA National Ambient Air Quality Standard (NAAQS) for ozone is also shown.**

3.3. Impacts of using sub-tiling in Noah land-use scheme on WRF model performance

Figure 10 presents Noah-default in comparison with Noah-subtiling model performance evaluation (MPE) results of surface temperature, humidity, wind speed and direction averaged over nine monitoring sites throughout Uinta Basin. Locations of these sites are shown in Figure 2.

For the basin during the studied period, the impact of WRF sub-tiling configuration accounting for subgrid-scale heterogeneity of land cover was negligible. As a remote region, Uinta Basin's land-use cover distributed relatively homogeneously with only three dominant NLCD40 categories (e.g., "Shrub", "Glassland" and "Barren Land"). Using mosaic land-use categories instead of dominant land-use in such region; therefore, did not pose significant impacts on model performance. Mallard and Spero (2019) sensitivity test on Noah-subtiling over entire US-CONUS domain shared the consistent result which indicated that southwest climate region including the State of Utah was not sensitive to subgrid-scale land-use cover accountability.

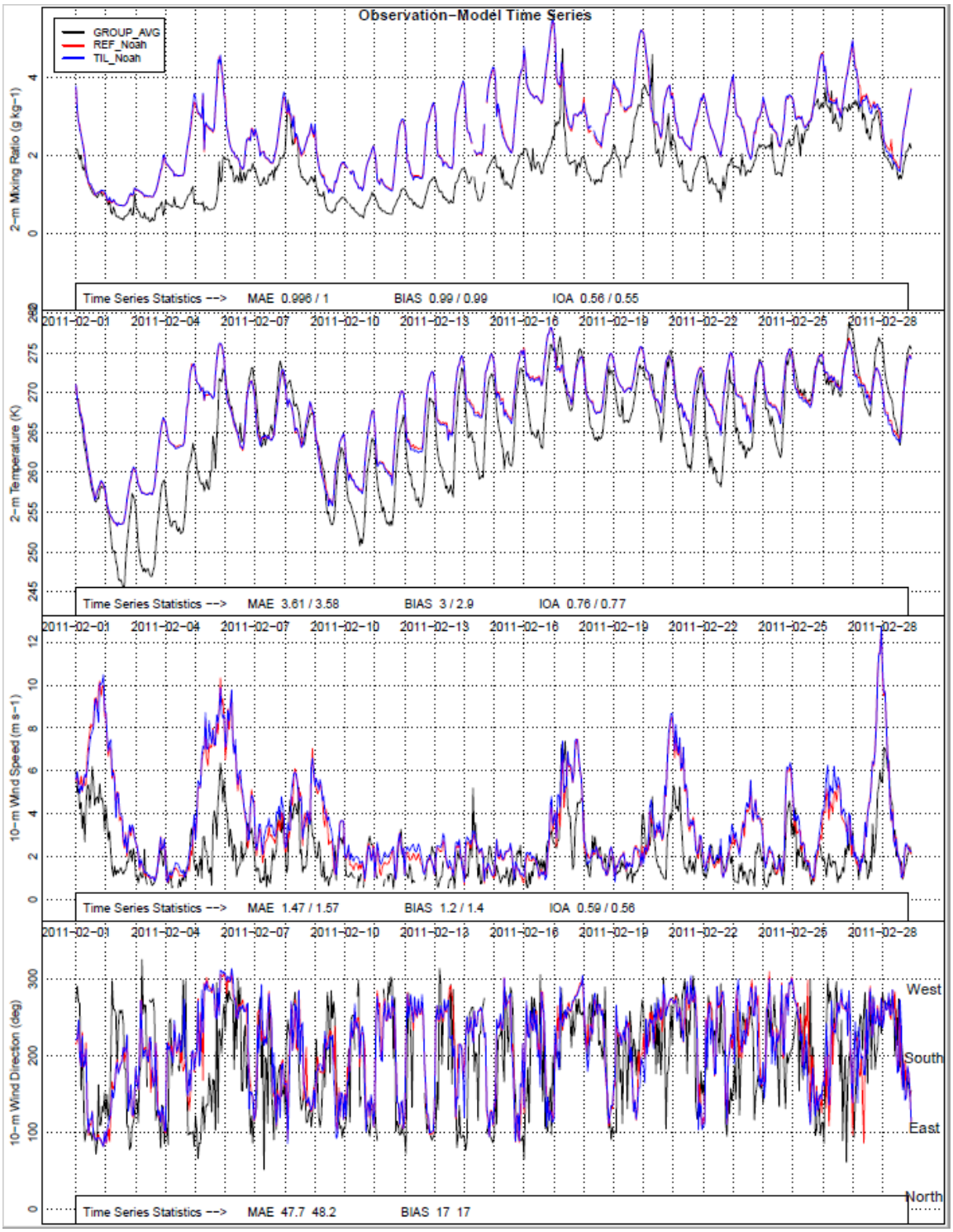

Figure 10. February 2011 time series of Uintah Basin-wide average mixing ratio, temperature, wind speed and wind direction simulated by Noah-default (red), Noah-subtiling (blue) on 1.3km domain and observations (black).

3.4. Model performance sensitivity tests on Noah LSM versus Pleim-Xiu LSM with soil moisture nudging land-use scheme.

Figure 11 presents Noah-default in comparison with Pleim-Xiu MPE results of surface temperature, humidity, wind speed and direction averaged over nine monitoring sites throughout Uinta Basin. Locations of these sites are shown in Figure 2.

During the studied period, Pleim-Xiu LSM consistently outperformed Noah LSM in simulating surface moisture and surface wind speed over Uinta Basin with significantly smaller Mean Absolute Error (MAE), bias and higher Index of Agreement (IOA). On the other hand, Pleim-Xiu LSM performed worse than Noah-LSM in term of surface temperature. The two LSMs performed similarly in predicting wind direction.

Tran et al. (2018) showed that the large model errors in surface temperature could be effectively corrected by surface observational nudging for temperature; whereas moisture errors were less sensitive to nudging impact leading to model over-predicting warm cloud and under-predicting radiation budget, which is important for ozone chemistry. Moreover, Pleim-Xiu LSM reflected the stability of inversion layer better than Noah LSM by predicting lower wind speed with higher agreement with observed data. Therefore, Pleim-Xiu LSM performance in combination with surface nudging could be very promising configuration for winter Uinta Basin weather predicting.

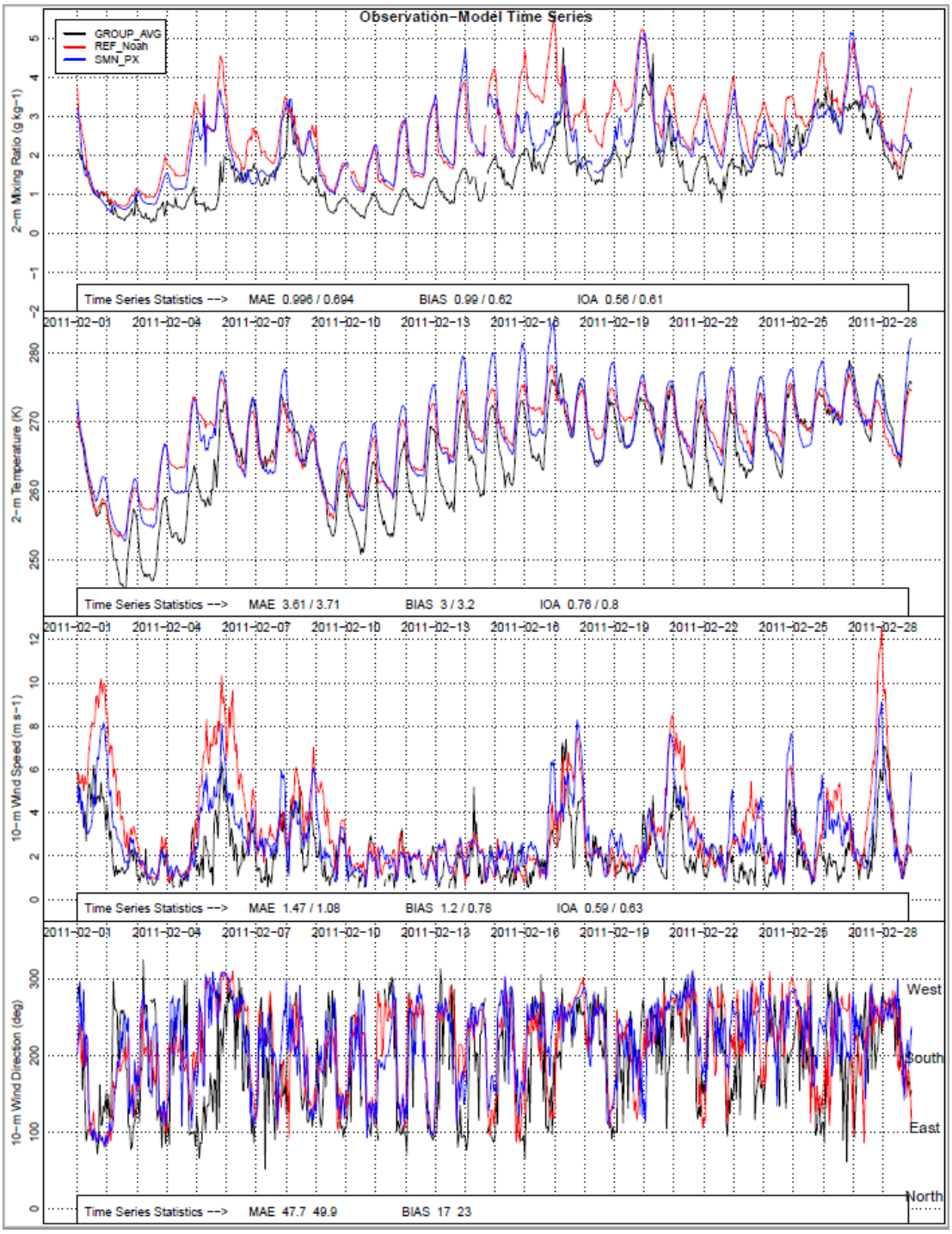

Figure 11. February 2011 time series of Uintah Basin-wide average mixing ratio, temperature, wind speed and wind direction simulated by Noah-default (red), Pleim-Xiu with soil nudging (blue) on 1.3km domain and observations (black).

4. Conclusions and Future Works

We have developed and streamlined the process to assimilate MODIS satellite data in to WRF and CAMx models and demonstrated its impacts on the models' performance. Instead of relying knowledge of albedo and snow cover basing on spare observations and conducting forcible correction to the model, our approach utilizes the high resolution and extensive coverage in both spatial and temporal scale of MODIS data to improve WRF simulations of these quantities. This approach, therefore, eliminates the guesswork from the modelers when dealing with WRF's discrepancies and thus eliminate the extra bias that would have been introduced to the model system.

Improvements to WRF simulation of surface albedo and snow cover is substantial. However, its impact on improving WRF's performance in other meteorological quantities are minimal. This suggests that a combination of MODIS data assimilation as presented in this study with surface and vertical nudging as discussed in Tran et al. (2018) would be the best approach to improve WRF performance.

In term of ozone performance in CAMx simulations, MODIS data assimilation did not result in significant improvement, and in some case even worse. This behavior dues to how albedo and snow cover are simulated in CAMx in standard configuration. However, we expect that positive impact of MODIS data assimilation on CAMx model performance will be more pronounce when examined over a longer simulation episode than what was examined in this study.

Our sensitivities simulations with using Noah and Pleim-Xiu LSM in WRF demonstrate another aspect where observational nudging approach could improve WRF model performance. Pleim-Xiu LSM consistently outperforms Noah LSM in simulating surface meteorology quantities. Our MODIS data assimilation technique has been tested with Noah LSM. Testing MODIS data assimilation technique with Pleim-Xiu LSM is the future work that we will perform.

This project is only our first step in using satellite data to improve current discrepancies in both meteorology and chemistry model. For example, using the same framework developed in this project we will investigate using satellite data to correct overestimation of cloud cover in WRF, and using chemistry observation from satellite system such as TROPOMI to improve photochemical model performance.

5. References

Helmig, D., Ganzeveld, L., Butler, T., & Oltmans, S. J. (2007). The role of ozone atmosphere-snow gas exchange on polar, boundary-layer tropospheric ozone - a review and sensitivity analysis. *Atmos. Chem. Phys., 7*(1), 15-30. doi:10.5194/acp-7-15-2007

Mallard, M. S., & Spero, T. L. (2019). Effects of Mosaic Land Use on Dynamically Downscaled WRF Simulations of the Contiguous United States. *Journal of Geophysical Research: Atmospheres, 124*(16), 9117-9140. doi:https://doi.org/10.1029/2018JD029755

Mansfield, M., Tran, H., & Tran, T. (2020). *2017 ARMS - Photochemical Grid Model Performance Evaluation for base year 2011*. Retrieved from Technical Report submitted to the Bureau of Land Management Utah Office, Utah State University - Bingham Research Center, Vernal UT 84078: https://binghamresearch.usu.edu/files/reports/blm_arms/BLM%20Utah%20ARMS17%20Base%20Year%20Performance%20Evaluation.pdf

Matichuk, R., Tonnesen, G., Luecken, D., Gilliam, R., Napelenok, S. L., Baker, K. R., . . . Lyman, S. N. (2017). Evaluation of the Community Multiscale Air Quality Model for Simulating Winter Ozone Formation in the Uinta Basin. *Journal of Geophysical Research: Atmospheres, 122*(24), 13545–13572.

Neemann, E. M., Crosman, E. T., Horel, J. D., & Avey, L. (2015). Simulations of a cold-air pool associated with elevated wintertime ozone in the Uintah Basin, Utah. *Atmospheric Chemistry and Physics, 15*(1), 135-151. doi:10.5194/acp-15-135-2015

NOHRSC. (2004). *Snow Data Assimilation System (SNODAS) Data Products at NSIDC*. Retrieved from: http://nsidc.org/data/docs/noaa/g02158_snodas_snow_cover_model/

Ran, L., Gilliam, R., Binkowski, F. S., Xiu, A., Pleim, J., & Band, L. (2015). Sensitivity of the Weather Research and Forecast/Community Multiscale Air Quality modeling system to MODIS LAI, FPAR, and albedo. *Journal of Geophysical Research: Atmospheres, 120*(16), 8491-8511. doi:doi:10.1002/2015JD023424

Tran, T., Tran, H., Mansfield, M., Lyman, S., & Crosman, E. (2018). Four dimensional data assimilation (FDDA) impacts on WRF performance in simulating inversion layer structure and distributions of CMAQ-simulated winter ozone concentrations in Uintah Basin. *Atmospheric Environment, 177*, 75-92. doi:https://doi.org/10.1016/j.atmosenv.2018.01.012

Appendix 1

Table A1. Summary of WRF model modifications

Module	Parameters	Descriptions
$WRF_HOME/Registry/Registry.EM_COMMON	SW_ALB_OLD, SW_ALB_NEW, SNOWC_OLD, SNOWC_NEW	New variables represent albedo (SW_ALB) and snow cover (SNOWC) added to *wrfsfdda_d<domain>* input files that will be read in by WRF during simulation. Frequency of data assimilation is controlled by variable *sgfdda_interval_m.* NEW and OLD represent values at the current and preceding hours, respectively.
$WRF_HOME /Registry/Registry.EM_COMMON $WRF_HOME /phys/module_physics_init.F	galb_sfc, gsnc_sfc	New variables added to WRF *namelist.input* to disable (values equal 0) or enable (values equal 1) albedo and snow cover assimilation
$WRF_HOME /phys/module_fdda_psufddagd.F $WRF_HOME /phys/ module_fddagd_driver.F	ALBEDO, SNOWC	Add ALBEDO and SNOWC as the two additional variables that are processed in fdda nudging process
$WRF_HOME /share/ output_wrf.F	nl_get_galb_sfc, nl_get_gsnc_sfc	Add two new subroutines to assimilate albedo and snow cover from MODIS data.
$WRF_HOME / dyn_em/ module_first_rk_step_part1.F	ALBEDO, SNOWC, sw_alb_old, sw_alb_new, snowc_old, snowc_new	Declare new variables and method to calculating them

$WRF_HOME is the top level directory of WRF model

Appendix 2

Table A2. Summary of CAMx model modifications

Module	Parameters	Descriptions
Module wrfcamx.f90	inalb, insnc, inlai	Add variables to read from wrf ouput files and process for CAMx binary input files.
$CAMX_HOME/Includes/camxfld.inc In $CAMX_HOME/CAMx: CAMx.f, getaldebo.f, startup.f, tstep_init.f	inalb, insnc, inlai, lrdmod	Add new variables represent albedo (*inalb*), snow cover (*insnc*) and leaf area index (*inlai*) that were processed by *wrfcamx* module .
$CAMX_HOME/Mod_src/ camxfld.f	inalb, insnc, inlai, lrdmod	Allocate memory for the new variables
In $CAMX_HOME/IO_bin: readinp.f, metinit.f		Add new code lines that read and store values of ALBEDO, SNOWCOVER and LAI from CAMx input binary files

$CAMX_HOME is the top level directory of CAMx model.