AQRP Monthly Technical Report

PROJECT TITLE	Incorporating Space-borne Observations to Improve Biogenic Emission Estimates in Texas	**PROJECT #**	14-017
PROJECT PARTICIPANTS	Arastoo Pour-Biazar; Richard McNider; Daniel Cohan, Rui Zhang	**DATE SUBMITTED**	5/8/2015
REPORTING PERIOD	**From:** April 1, 2015 **To:** April 30, 2015	**REPORT #**	12

A Financial Status Report (FSR) and Invoice will be submitted separately from each of the Project Participants reflecting charges for this Reporting Period. I understand that the FSR and Invoice are due to the AQRP by the 15th of the month following the reporting period shown above.

Detailed Accomplishments by Task

Progress Summary for WRF 2013 Simulations

The August and September 2013 WRF simulations for all domains (36-, 12-, and 4-km grid spacing) were completed. The simulations comprise control and satellite-cloud assimilation for each of the domains. Evaluation of 36-km domain was presented in the March report. Preliminary results for nests are consistent with the results from the coarse domain. WRF outputs from these simulations were shipped to Rice University to be used in producing a new set of biogenic emissions.

Progress Summary Evaluating UAH Insolation Satellite Retrievals During August 2006

Further evaluation of PAR/insolation products from satellite retrievals and the WRF simulations during August 2006 were made to compare the radiation performance, which is critical for biogenic emission model estimates. (REPORT ATTACED)

Progress Summary Developing new soil biome spatial map based on 12km CONUS 40-categroy 2006 NLCD-MODIS land use classification (NLCD40)

Process were made to develop a new soil biome spatial map based on 12km CONUS 40-category 2006 NLCD-MODIS land use classification (NLCD40) and climate zone definition, which will replace the GEOS-Chem biome map to better represent the up-to-date LU/LC change with finer details. (REPORT ATTACED)

Preliminary Analysis
Attached.

Data Collected
None for this period.

Identify Problems or Issues Encountered and Proposed Solutions or Adjustments
Due to the large size of outputs from UAH WRF simulation, transferring data over the internet was not feasible. We decided to copy the data on an external hard disk and send the data to Rice University via FedEx.

Goals and Anticipated Issues for the Succeeding Reporting Period
Finish the analysis of WRF simulations and continue emission estimate efforts. We are in the process of compiling the quarterly report.

Detailed Analysis of the Progress of the Task Order to Date
Attached.

Arastoo Pour Biazar

Submitted to AQRP by:

Principal Investigator: Arastoo Pour Biazar

Part I. Evaluating UAH insolation satellite retrievals during August 2006

Further evaluation of PAR/insolation products from satellite retrievals and the WRF simulations during August 2006 were made to compare the radiation performance, which is critical for biogenic emission model estimates. The four cases for different PAR/insolation estimates are: 1. WRF control case 'cntrl' with basic configurations; 2. WRF cloud assimilation case 'analytical' with the cloud assimilation from GOES observations; 3. PAR satellite retrievals from University of Maryland (UMD) with the resolution of 0.5 degrees (http://www.atmos.umd.edu/~srb/gcip/cgi-bin/historic.cgi?auth=no parameter=par) and 4. PAR satellite retrievals from UAH with the resolution of 4 km. PAR was computed from the WRF runs by scaling ground-reaching solar radiation (RGRND) by 50%. The details of WRF-MEGAN model configurations, simulation case arrangement, simulation time period selection as well as model performance evaluation are given in Table 1.

The four sets of PAR estimates were evaluated against observations from seven NOAA SURFRAD direct measurement sites (see Figure 1 for the site locations with purple dots) for August 2006. Evaluation statistics are given in Table 2, and corresponding scatterplots are given in Figure 2. Overall, both satellite retrieval products substantially outperformed the two WRF runs, including better correlation coefficients (R=0.96~0.97 versus R=0.93) and smaller simulation errors (NME=20.7%~20.1% versus NME=32.8%~35.5%). Both satellite retrievals also achieved much lower bias than the base WRF run, though the UMD retrieval tended to underestimate PAR (NMB=-12.4%) while the UAH retrieval tended to overestimate PAR (NMB=10.2%). Despite the different sign of the bias, the performance of the two satellite products by most other metrics was similar (Table 2). The outperformance of the satellite retrievals compared to the WRF runs is illustrated by the scatterplots in Figure 2, with the satellite cases (red and blue) clustering closer to the 1:1 ratio lines at all seven evaluation sites. The UAH PAR data achieved its best performance by most metrics at the FPK site and its worst performance at the PSU site, although its differences in performance across sites were not dramatic. More evaluation during other periods would be needed to test whether the spatial patterns of performance continue or are specific to the August 2006 period.

The insolation outputs from the two WRF runs (based on the WRF variable 'RGRND' for solar radiation reaching surface) and the UAH satellite retrieval were compared for August 2006 with 47 available broadband radiation monitoring stations over Texas (with the locations for blue dots given in Figure 1). The solar radiation observation data are provided by TCEQ. Table 3 summarizes the overall model performance for insolation for the three cases. Using satellite data substantially reduces the overprediction bias of the WRF control run, reducing NMB from 22.2% to 8.9% for the cloud-assimilated WRF run and 7.5% for the UAH retrieval. While cloud assimilation was able to reduce the overprediction bias in WRF, the UAH retrieval strongly outperformed both WRF runs in terms of error and correlation (Table 3).

Figure 3 provides the spatial maps of the correlation coefficient (R, upper panel) and normalized mean bias (NMB, lower panel) for the individual 47 sites among the three different cases. In terms of correlation coefficient R, the WRF 'cntrl' case performs worst (R~0.85) near the coastal sites especially around the greater Houston area. It performs better at inland sites near Dallas region(R~0.92). The cloud assimilation WRF run slightly improves model correlation at most of the evaluated sites. UAH satellite retrieval insolation products significantly improve the model correlation with the R values at inland sites approaching 1 and at the coastal sites mostly around 0.95. In terms of NMB, all three products performed better around Dallas and other inland sites than near Houston, and the satellite-based cases far outperformed the base WRF run (Figure 3).

Also during this period, we corrected the utility code in the UNC Spatial Allocator (www.epa.gov/AMD/Tools/spatialAllocator.html) to ensure that the 4km resolution UAH satellite products mapped properly to the model grid. The new coding solved the dislocation problem by using the old UAH satellite regridding utility and corrected conflicting assumptions about Earth's radius and shape. Figure 4 demonstrates that the new Spatial Allocator code properly maps the UAH 4km GOES satellite PAR retrievals to the three TCEQ domain settings given in Table 1. Note that there is a portion of Washington State out of the satellite retrieval domain and hence no values over corresponding region.

UAH already finished and sent the two sets of WRF model outputs for September 2013 (as 'Task 4' promised in the project) to Rice U in order to complete the 'Task 5' emission estimates for 2013 simulations by MEGAN and 'Task 6' performing air quality simulation using the different BVOC emission estimates on the three TCEQ domain settings (Table 1). Note that we will use CMAQ instead of CAMx as the host model to run the simulations. The series of CMAQ outputs will be evaluated against Discover-AQ data to investigate the sources of uncertainty reported in the literature with respect to BVOCs.

PART II. Developing new soil biome spatial map based on 12km CONUS 40-categroy 2006 NLCD-MODIS land use classification (NLCD40)

Current offline Berkeley-Dalhousie Soil NOx Parameterization (BDSNP) module developed by Rice use the soil biome map directly re-gridded from global atmospheric chemistry model GEOS-Chem, which is too coarse for regional model implementation. Biome map is related with the land use/ land cover (LU/LC) classification and K**ö**ppen-Geiger climate zone definition (Kottek et al., 2006) and will determine the base soil NOx emission strength. Here, we present the process to develop a new soil biome spatial map based on 12km CONUS 40-category 2006 NLCD-MODIS land use classification (NLCD40) and climate zone definition, which will replace the GEOS-Chem biome map to better represent the up-to-date LU/LC change with finer details.

For the first step, a mapping table to transfer the 40 categories of NLCD40 at each modeling grid to the 24 biome types with for which soil NO emission factors are available from Steinkamp and Lawrence (2011) was constructed. Table 4 details the connecting algorithm between the NLCD40 and 24 biome type for soil NO emission estimates. For the categories including in both sources with the identical names, the mapping are direct, such as 'evergreen needleleaf forest', 'deciduous needleaf forest', 'mixed forest', 'savannas' and 'grassland'. Further separation will be done for the biome categories with different emission factors at different climate zone. For the categories in NLCD40 with detail definitions than corresponding biome category such as water and urban lands, they will be consolidated into one category in biome by addition. For example, 'soil and ice' in soil biome category is equal to the addition of 'permanent snow and ice' and 'perennial ice-snow' in NLCD40 MODIS category; 'urban and build-up lands' in soil biome category is equal to the addition of 'developed open space', 'developed low intensity', 'developed medium intensity' and 'developed high intensity'. For the categories appearing only in NLCD40, the mapping algorithm is determined refer to the CMAQ mapping scheme, which is documented in each the CSQY_DATA_* under the MECHS/ directory at the CMAQ source code release. One of the examples is to map 'lichens' and 'moss' in NLCD40 to the category 'grassland' in soil biome.

For the second step, a model resolution compatible Köppen-Geiger climate zone classification is also needed to create in order allocates the different emission factor for the same biome type (e.g. 'grassland') at different locations. For the 12km CONUS domain, we use Spatial Allocator to generate the 5 category climate zone map in Figure 5 (A: equatorial, B: arid, C: warm temperature, D: snow, E: polar) based on the county level text file climate zone documentation as the surrogate (http://koeppen-geiger.vu-wien.ac.at/data/KoeppenGeiger.UScounty.txt). The 12km climate zone product (Figure 5, bottom) matches well with original files (Figure 5, top) and with most of southern US classified as the 'warm temperature' region. Only a slight portion of Rocky mountain summit is classified as 'polar' climate and the south corner of Florida State are classified as 'equatorial' climate.

Figure 6 provides the comparison between the new soil biome spatial map based on the finer resolution LU/LC definition used in current CMAQ simulation and the old soil biome spatial map based on coarse GOES-Chem LU/LC setting. As the reference, the independent 30m resolution 2011 NLCD spatial map based on the Landsat satellite (http://www.mrlc.gov/nlcd2011.php) were also provided here. In order to be comparable with the NLCD 2011 classification system, similar color legend was used to visualize the biome soil type. It can be found that the new CMAQ 12km soil biome has much more detailed texture and closer geolocation correspondence than the old soil biome derived from GOES-Chem. For example, the new biome map has more identified 'cropland' at the central US states (e.g. Oklahoma, Kansas) while the old biome map are all 'grassland' at the same places. Also the new biome map identifies a lot of ‘wetland’ near the southern coastal line area, which match with the NLCD2011 classification for ’90 woody wetlands’ and ’95 emergent herbaceous wetlands’.

We will use the new set of soil biome map to run the soil BDSNP model and evaluate its impact on soil NO emission estimates.

References:

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel, 2006: World Map of the Köppen-Geiger climate classification updated. Meteorol. Z., 15, 259-263. DOI: 10.1127/0941-2948/2006/0130.

Steinkamp, Jörg and Mark Gary Lawrence. "Improvement and evaluation of simulated global biogenic soil NO emissions in an AC-GCM." Atmospheric Chemistry and Physics 11.12 (2011): 6063-6082.

Table 1. WRF-MEGAN model configurations in this study

WRF			
Version:	ARW V3.6.1	**Shortwave radiation:**	RRTMG scheme
Horizontal resolution:	D1 (CONUS, 36km); D2 (SW US, 12km)	**Surface layer physic:**	Pleim-Xiu surface model
	D3 (E Texas, 4km)	**PBL scheme:**	ACM2
Vertical resolution:	42 layer (first layer height ~ 37 m)	**Microphysics:**	Morrison double-moment scheme
Boundary Condition:	NARR 32km	**Cumulus Parameterization:**	Kain-Fritsch scheme
Initial condition:	NECP-ADP	**Assimilation:**	Analysis nudging @ D1 Option run w/ cloud assimilation from GOES
Longwave radiation:	RRTMG scheme		
MEGAN			
Version:	V2.10	**Emission factor:**	Global emission factor (ver. 2011)
Horizontal resolution:	Same as WRF	**Leaf area index:**	30 sec, MODIS 8 day average
Plant function type:	16 CLM PFT types, 30 sec	**Gas-phase mechanism:**	CB5
Simulation Case Arrangement			
1. PAR_cntrl:	Base WRF simulation to provide insolation for MEGAN		
2. PAR_analytical:	Base WRF + cloud assimilation from GOES to provide insolation for MEGAN		
3. PAR_UMD:	Direct use PAR retrievals from UMD, other met inputs same as case 'PAR_analytical'		
4. PAR_UAH:	Direct use PAR retrievals from UAH, other met inputs same as case 'PAR_analytical'		
Simulation Time Period			
	Sep 1-30, 2013		
Model Performance Evaluation			
1. NOAA SURFRAD (Surface Radiation) Network	(http://www.esrl.noaa.gov/gmd/grad/surfrad/)		
2. TCEQ broadband radiation monitoring Network	(http://www.tceq.state.tx.us/agency/data/air_met_data.html)		

Table 2. Summary of statistics of PAR simulation/retrievals for different cases at 7 SURFRAD network sites

CASE	SITE	OBS_AVE (W/m2)	SIM_AVE (W/m2)	IA	R	RMSE (W/m2)	MB (W/m2)	MAGE (%)	NMB (%)	NME (%)
PAR_cntrl	BON	92.6	119.8	0.94	0.93	68.1	27.1	38.5	29.2	41.5
	DRA	140.8	162.1	0.99	0.99	41.6	21.4	26.0	15.2	18.5
	FPK	109.8	130.2	0.96	0.96	56.2	24.2	32.8	22.0	29.9
	GWN	111.0	132.7	0.94	0.92	71.2	21.8	40.5	19.6	36.5
	PSU	92.1	133.9	0.93	0.94	77.0	37.2	43.0	40.4	46.7
	SXF	101.6	126.1	0.95	0.93	65.5	24.4	33.4	24.0	32.8
	TBL	105.9	113.9	0.92	0.86	77.7	7.9	45.4	7.5	42.9
	average	107.7	131.2	0.95	0.93	65.3	23.4	37.1	22.6	35.5
PAR_analytical	BON	92.6	107.5	0.95	0.94	58.3	14.9	34.4	16.0	37.1
	DRA	140.8	162.0	0.99	0.99	41.3	21.3	25.8	15.2	18.4
	FPK	109.8	121.9	0.97	0.96	49.8	16.3	29.7	14.8	27.1
	GWN	111.0	124.8	0.96	0.94	61.9	13.9	36.4	12.5	32.8
	PSU	92.1	118.1	0.94	0.93	65.4	21.8	37.0	23.6	40.2
	SXF	101.6	119.2	0.95	0.93	60.2	17.5	32.0	17.2	31.4
	TBL	105.9	93.7	0.89	0.81	80.7	-12.4	45.0	-11.7	42.5
	average	107.7	121.0	0.95	0.93	59.6	13.3	34.3	12.5	32.8
PAR_UMD	BON	92.6	87.2	0.99	0.98	25.8	-5.4	15.3	-5.9	16.5
	DRA	140.8	108.2	0.97	0.97	51.7	-32.2	32.5	-22.9	23.1
	FPK	109.8	87.1	0.97	0.97	39.0	-19.4	24.8	-17.7	22.6
	GWN	111.0	103.1	0.99	0.98	30.0	-7.9	16.5	-7.1	14.9
	PSU	92.1	95.6	0.98	0.97	28.9	-0.1	15.9	-0.1	17.2
	SXF	101.6	94.3	0.98	0.96	36.5	-7.3	20.4	-7.1	20.1
	TBL	105.9	78.5	0.94	0.95	57.2	-27.4	32.1	-25.9	30.3
	average	107.7	93.4	0.97	0.97	38.5	-14.2	22.5	-12.4	20.7
PAR_UAH	BON	92.6	110.6	0.97	0.96	44.1	19.4	25.2	21.3	27.6
	DRA	140.8	130.4	0.98	0.96	44.0	-10.0	18.8	-7.1	13.4
	FPK	109.8	112.8	0.98	0.96	39.9	4.2	19.4	3.7	17.2
	GWN	111.0	127.3	0.98	0.97	44.2	16.6	23.7	15.0	21.4
	PSU	92.1	117.1	0.97	0.97	44.2	21.1	25.7	23.0	28.1
	SXF	101.6	114.9	0.98	0.96	39.1	11.2	19.5	10.8	18.8
	TBL	105.9	111.4	0.97	0.94	45.4	5.1	22.0	4.8	20.7
	average	107.7	117.8	0.97	0.96	43.0	9.7	22.0	10.2	21.0

Note: IA-index of agreement, R-correlation coefficient, RMSE-root mean square error, MB-mean bias, MAGE-mean aggregate gross error, NMB-normalized mean bias, and NME-normalized mean error

Table 3. Summary of statistics of insolation simulation/retrievals for different cases at 47 TCEQ network sites

	OBS_AVE (W/m2)	SIM_AVE (W/m2)	IA	R	RMSE (W/2)	MB	MAGE	NMB (%)	NME (%)
WRF cntrl	248.6	299.8	0.95	0.91	142.3	53.9	74.7	22.2	30.7
WRF analytical	248.6	266.8	0.95	0.91	143.9	20.3	74.9	8.9	30.7
UAH satellite	248.6	263.6	0.96	0.96	123.2	17.3	71.8	7.5	29.5

Note: IA-index of agreement, R-correlation coefficient, RMSE-root mean square error, MB-mean bias, MAGE-mean aggregate gross error, NMB-normalized mean bias, and NME-normalized mean error

Table 4. Mapping table used to create the soil biome map based on NLCD40 MODIS land use/land cover categories

ID	NLCD40 MODIS CATEGORY (40)	ID	SOIL BIOME CATEGORY (24)
1	Evergreen Needleleaf Forest	19	Evergr. Needel. Foresst
2	Evergreen Broadleaf Forest	16 and 21	Evergr. Broadl. Forest
3	Deciduous Needleleaf Forest	18	Dec. Needel. Forest
4	Deciduous Broadleaf Forest	17 and 20	Dec. Broadl. Forest
5	Mixed Forests	15	Mixed Forest
6	Closed Shrublands	7	Closed shurb
7	Open Shrublands	8 and 9	Open shrubland
8	Woody Savannas	14	Woody savannah
9	Savannas	11 and 12	Savannah
10	Grasslands	10 and 13	Grassland
11	Permanent Wetlands	2	Permanent Wetland
12	Croplands	22	Cropland
13	Urban and Built Up	23	Urban and build-up lands
14	Cropland-Natural Vegetation Mosaic	24	Cropland/nat. veg. mosaic
15	Permanent Snow and Ice	3	Snow and ice
16	Barren or Sparsely Vegetated	6	Barren
17	IGBP Water	1	Water
18	Unclassified	1	Water
19	Fill value	1	Water
20	Open Water	1	Water
21	Perennial Ice-Snow	3	Snow and ice
22	Developed Open Space	23	Urban and build-up lands
23	Developed Low Intensity	23	Urban and build-up lands
24	Developed Medium Intensity	23	Urban and build-up lands
25	Developed High Intensity	23	Urban and build-up lands
26	Barren Land (Rock-Sand-Clay)	24	Cropland/nat. veg. mosaic
27	Unconsolidated Shore	24	Cropland/nat. veg. mosaic
28	Deciduous Forest	16 and 21	Evergr. Broadl. Forest
29	Evergreen Forest	19	Evergr. Needel. Foresst
30	Mixed Forest	15	Mixed Forest
31	Dwarf Scrub	8 and 9	Open shrubland
32	Shrub-Scrub	8 and 9	Open shrubland
33	Grassland-Herbaceous	10 and 13	Grassland
34	Sedge-Herbaceous	14	Woody savannah
35	Lichens	10 and 13	Grassland
36	Moss	10 and 13	Grassland
37	Pasture-Hay	24	Cropland/nat. veg. mosaic
38	Cultivated Crops	22	Cropland
39	Woody Wetlands	2	Permanent Wetland
40	Emergent Herbaceous Wetlands	2	Permanent Wetland

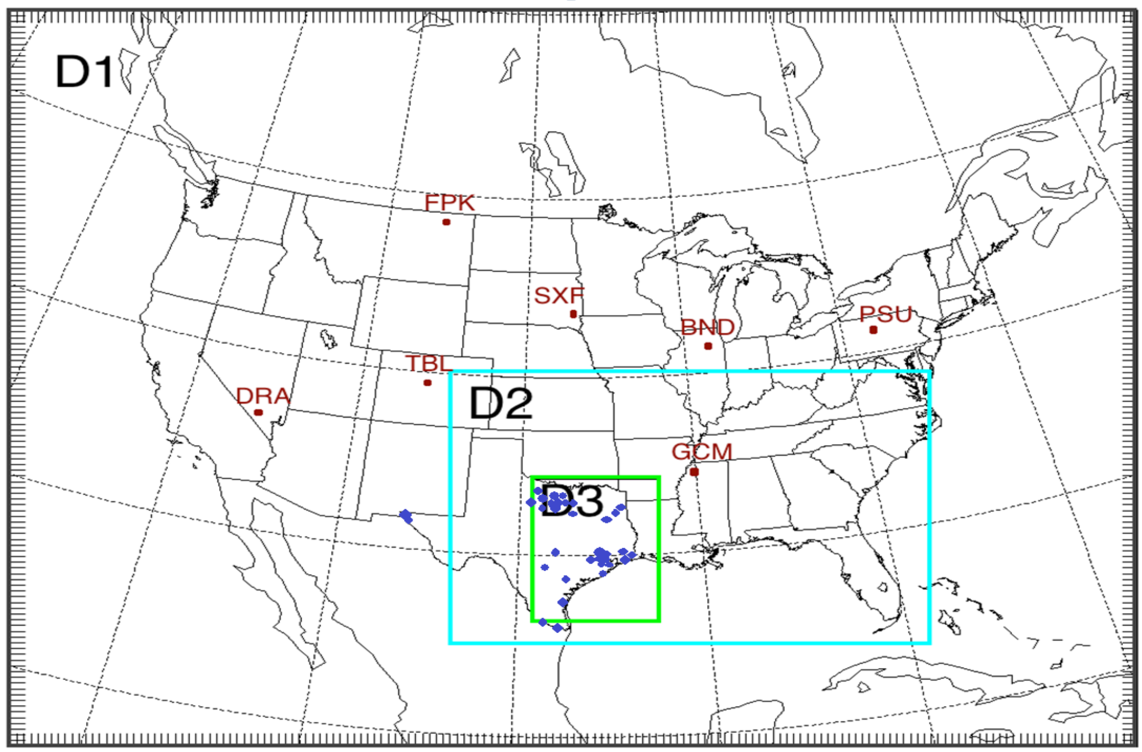

Figure 1. Three nesting domains for WRF-MCIP simulation in this study and locations of the insolation/PAR evaluation sites at SURFRAD network (red) and TCEQ broadband radiation network (blue)

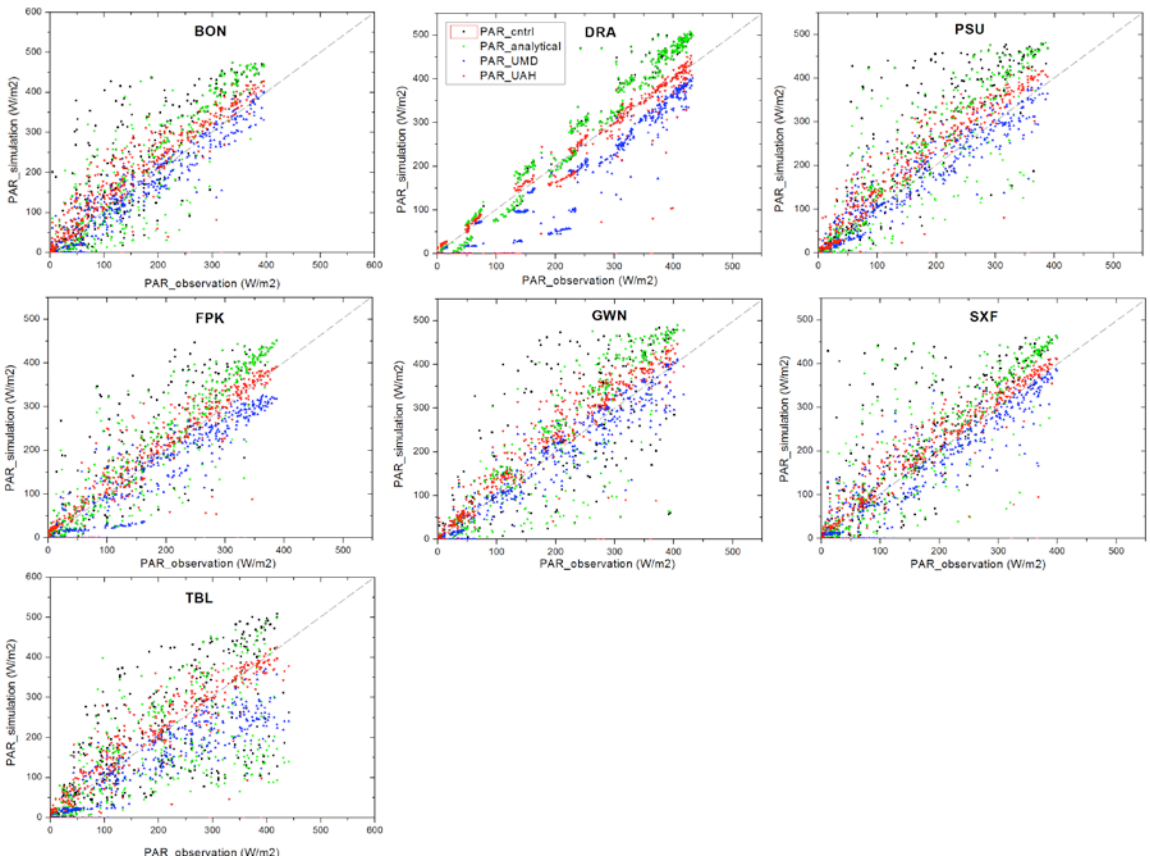

Figure 2. Scatter plots between four different hourly simulated/retrieved PAR and observed PAR during August 2006 at 7 SURFAD sites

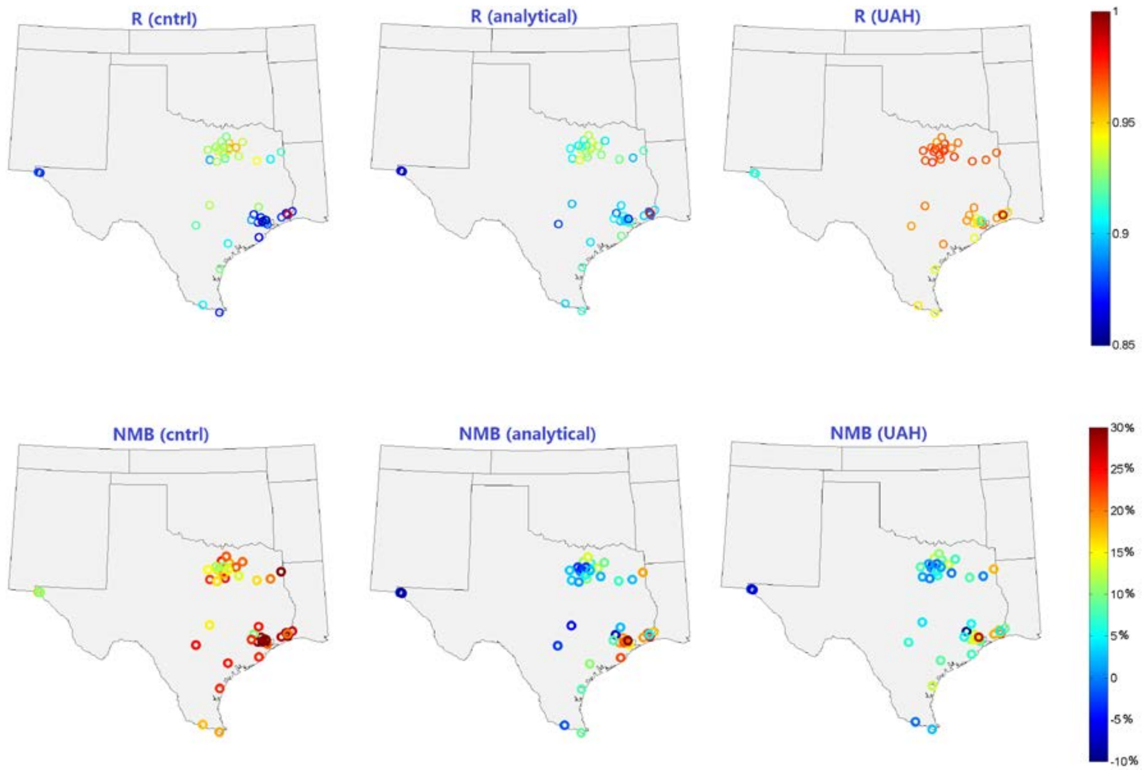

Figure 3. Performance of WRF incoming solar radiation (RGRND) from simulation case 'cntrl' (left) and 'analytical' (right) as well as UAH insolation retrievals (right) at TCEQ sites. The upper panel shows the correlation coefficient (R) and the lower panel shows the normalize mean bias (NMB)

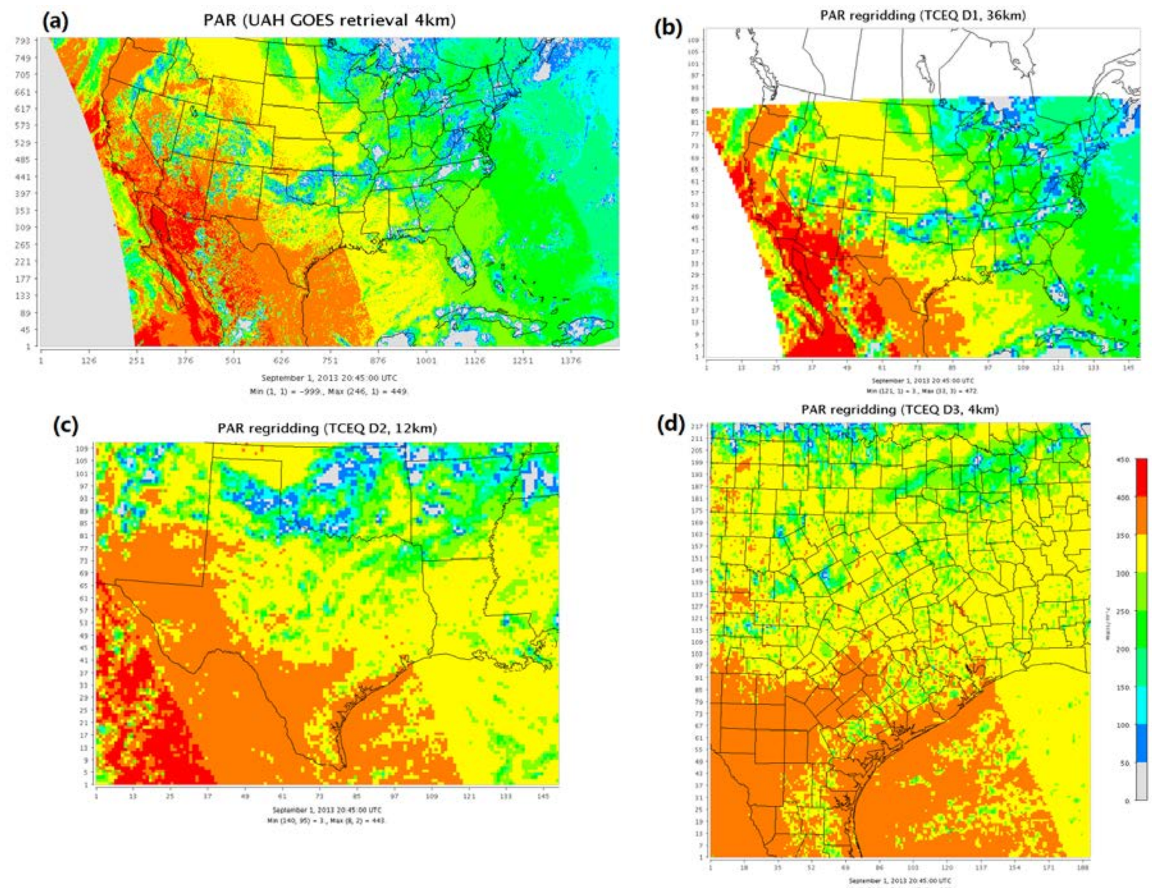

Figure 4. Domain regridding check from the (a) UAH 4km PAR from GOES satellite retrieval to (b) TCEQ D1 36 km CONUS domain; (c) TCEQ D2 12km Texas domain and (d) TCEQ 4km Southern Texas domain on September 1, 2013 20:45:00 UTC

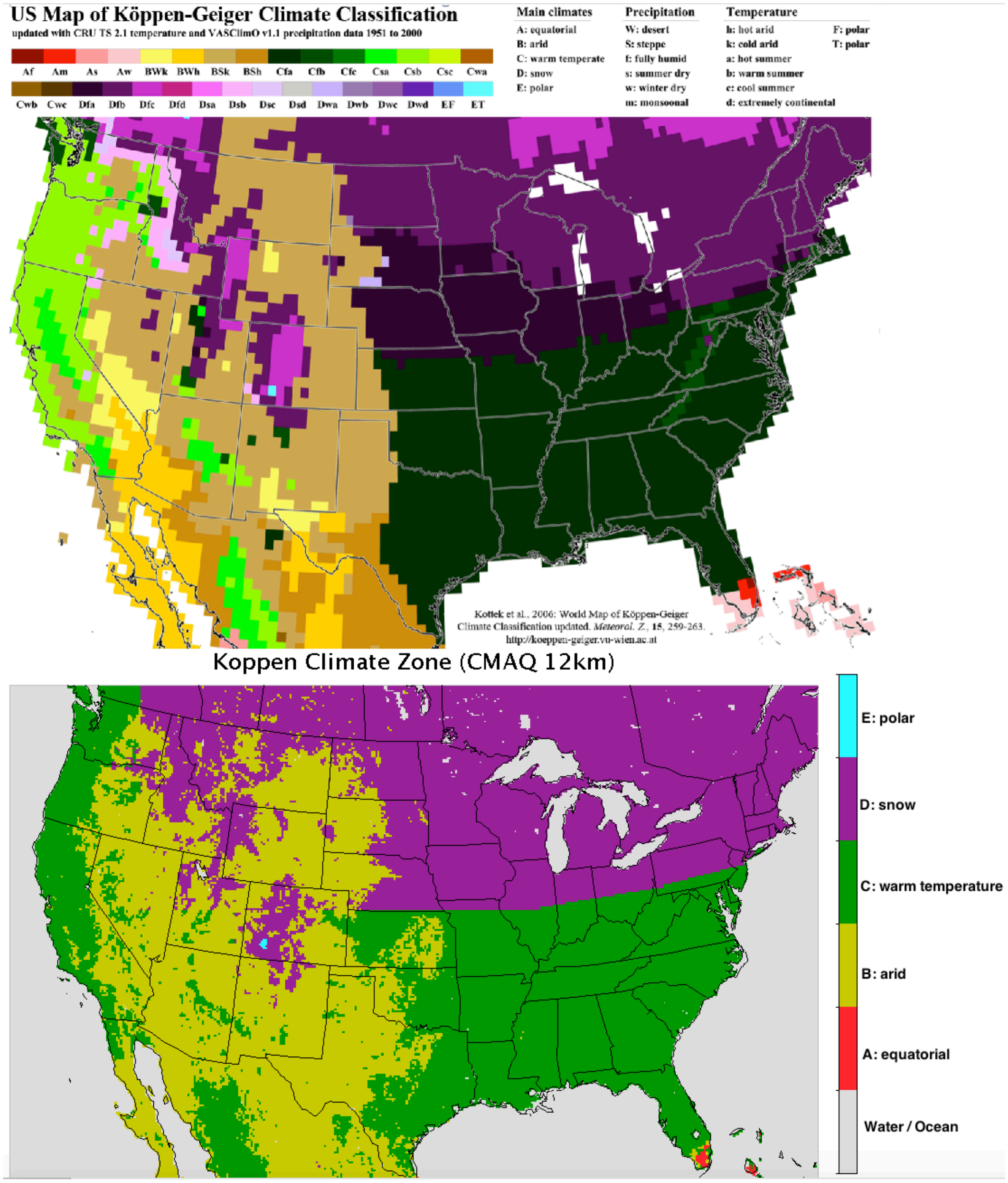

Figure 5. Spatial map of five climate zone over CMAQ CONUS 12km domain (bottom) based on Köppen-Geiger climate classification (top)

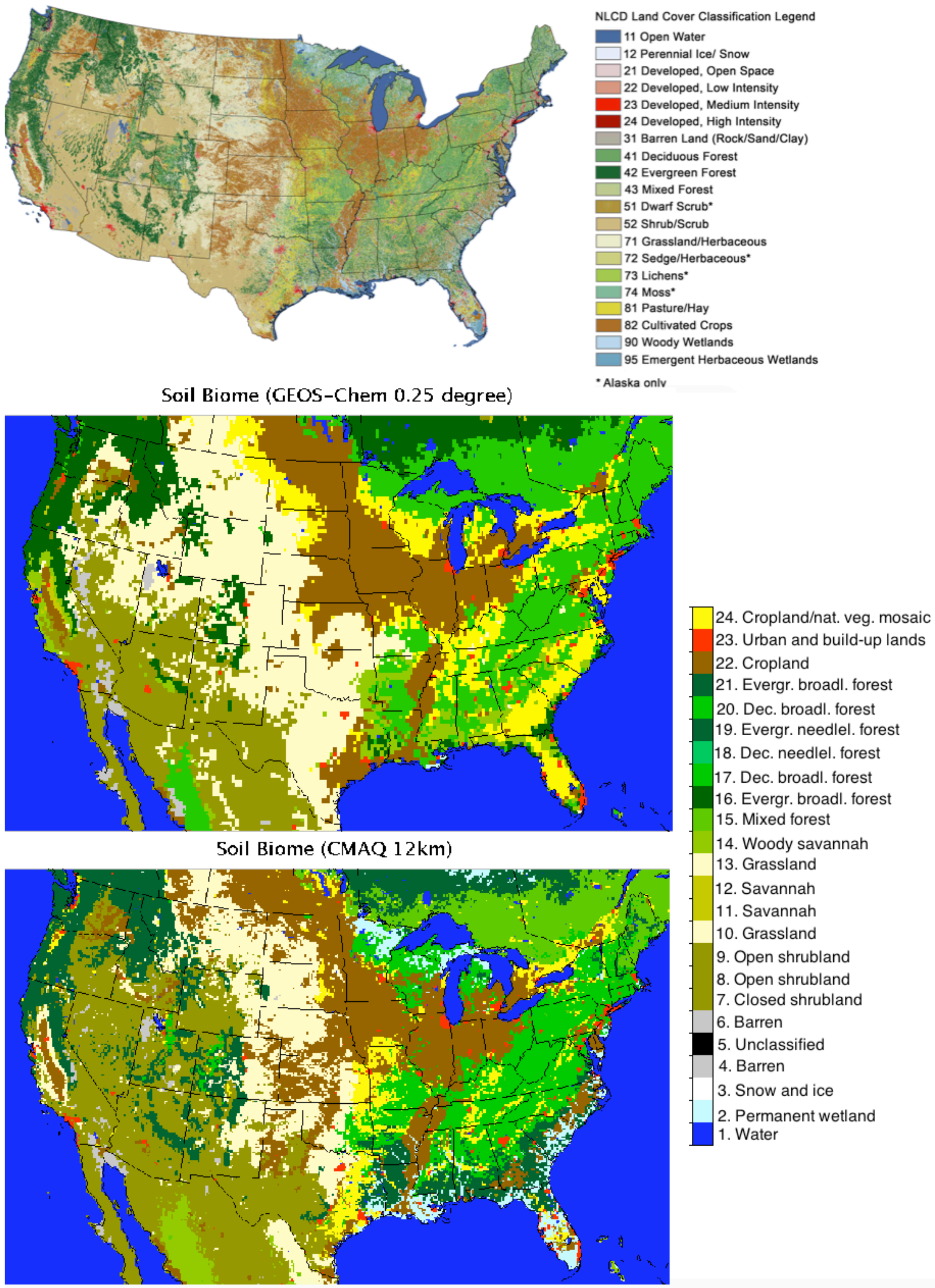

Figure 6. Comparison the spatial pattern of 2011 National Land Cover Database (30m resolution, top) with the soil biome type developed either from GEOS-Chem (0.25 degree resolution, middle) or from MODIS NLCD40 classification in CMAQ (12km resolution, bottom)