The effect of model parameters on the spatial structure of forecast fields

Caren Marzban[1,2], Robert Tardif[3], Scott Sandgathe[1],
Corinne Jones[2], Xiaochuan Du[2], Ning Li[2],
Natalia Hryniw[3], Nicholas C. Lederer[4],
James D. Doyle[5], Yi Jin[6]

[1] Applied Physics Lab, University of Washington, Seattle, WA 98195
[2] Dept of Statistics, University of Washington, Seattle, WA 98195
[3] Department of Atmospheric Sciences, University of Washington, Seattle, WA 98195
[4] The Boeing Company, Applied Mathematics, Seattle, WA 98124-2207
[5] Naval Research Laboratory, Monterey, CA 93943-5502

Abstract

In spatial verification techniques, it is often important to take into account the spatial structure of forecasts and observations. As such, it is desirable to be able to affect the spatial structure of forecasts. In this talk, it is shown that some model parameters have coherent and significant effects which can, in principle, be employed to revise the spatial structure of the forecasts. A variety of forecast quantities are examined from two NWP models - COAMPS and WRF.

Introduction

It is well-known that model parameters affect forecasts.

The question we ask here is not How?

But rather How should one go about answering the How?

I.e., this talk is not about results (must be treated cautiously),

But rather about methodology.

Bird's-Eye View

WRF/ARW/SKEBS	COAMPS
Version 3.7.0	4.2.2
25-km domain over CONUS	81-km domain over CONUS
003-120hr forecasts	24hr forecasts
prcp, t2m, t500hPa wind speed at 250hPa, and at 850hPa	convective and stable precip surface temp, water vapor
9 days (Dec '14 - Mar '15) 10 days apart	40 days (Feb - July '09) 3 days apart
2-8 model params	11 model params
Fractional Factorial Designs (LSD and 2^{k-p})	Latin Hypercube Sampling

Only a sample of the proposed devices are shown here.

Model Parameters

COAMPS:

ID	Name	Description
1	cumulus1	Temp increment at the LCL for KF trigger
2	cumulus2	Cloud radius function
3	cumulus3	Fraction of precip fed back to grid scale
4	PBL1	Mixing length
5	PBL2	Surface flux perturbations
6	Cumulus4	Vertical velocity for trigger
7	Cumulus5	2nd method to perturb temp at the LCL for trigger
8	Micro1	Autoconversion factor
9	Micro2	Autoconversion factor
10	Micro3	Slope intercept parameter for rain
11	Micro4	Slope intercept parameter for snow

KF = Kain-Fritsch

PBL = Planetary Boundary Layer

LCL = Lifted Condensation Level

WRF:

Stochastic Kinetic Energy Backscatter Schemes (SKEBS)

ID	Name	Description
1	rexponent_t	Spectral slope for potential temperature perturbations
2	rexponent_psi	Spectral slope for streamfunction perturbations
3	kminforc	Min forcing wave# in longitude for streamfunction pert
4	tot_backscat_psi	Total backscattered dissipation rate for streamfunction
5	ztau_t	Decorrelation time for potential temperature perturbations
6	ztau_psi	Decorrelation time for streamfunction perturbations
7	tot_backscat_t	Total backscattered dissipation rate for potential temp
8	lminforc	Min forcing wave# in latitude for streamfunction pert

Method

Statistical model for WRF sensitivity:

$$y_{i,j,k\cdots m} = \mu + \mathrm{Day}_i + \mathrm{X1}_j + \mathrm{X2}_k + \cdots + \epsilon_{i,j,k\cdots m},$$

$y_{ijk\cdots lm}$ = response (e.g., observed precip at a grid point)
on the i^{th} Day,
for the j^{th}, k^{th}, $\cdots$, values of the WRF params X1, X2, $\cdots$
for the m^{th} replication of the experiment.
$X1_j$ = True effect (mean response - μ) of WRF param X1
Etc.

Fixed effects model: all factors fixed (non-random), except $\epsilon \sim N(0, \sigma^2_\epsilon)$.
t-test of $H_0 : \mathrm{X1} = \mu$, etc., or
F-test of $H_0 : \mathrm{Xi} = \mu \;\; \forall i$

Random effects model: all factors zero-mean random (except μ), with

$$\sigma^2_{\mathrm{Response}} = \sigma^2_{\mathrm{Day}} + \sigma^2_{\mathrm{X1}} + \sigma^2_{\mathrm{X2}} + \cdots + \sigma^2_\epsilon \; .$$

F-test of $H_0 : \sigma^2_i = 0$
F-test of $H_0 : \sigma^2_i = 0 \;\; \forall i$
Intraclass correlation $\rho = \frac{\sigma^2_i}{\sigma^2_{\mathrm{Response}}}$

Statistical model for COAMPS sensitivity:

Multivariate Multiple Regression (MMR)

# of responses = $9 = 3 \times 3$.
Pillai's trace test for if a model param has effect on *any* of the 9 responses.

MMR accounts for spatial correlations and multiple hypothesis testing.

Method: Continued ...

For each day, for each model param $\rightarrow$ map of p-values (or ρ values).

More useful to "combine" maps across days into a single map, for each param.

How?

At each grid point, the daily p-values are subjected to a test of uniformity (chi-squared and/or Kolmogorov-Smirnov).

No α (significance level)!

Even though MMR gives corrected p-values, it's not necessary because we don't compare p-values with α.

Sampling Designs

We have k factors, L levels each.

Full factorial design requires L^k runs.
Fractional Factorial Designs require fewer runs.
The price: aliasing (i.e., cannot be estimated separately).
Magic: We know special runs for estimating the main effects.

Graeco-Latin Square Design (GLSD) requires L^2 runs.

Example of a 5-factor GLSD, with $L = 5$ levels each:

111	222	333	444	555
534	145	251	312	423
452	**513**	124	235	341
325	431	542	153	214
243	354	415	521	132

2^{k-p} Designs (i.e., $L = 2$) require 2^{k-p} runs.

$2^{8-4} = 16$ runs involving 8 factors
$2^{14-10} = 16$ runs involving 14 factors
$2^{15-11} = 16$ runs involving 15 factors (instead of $2^{15} = 32,768$ runs)

The 16 runs for a 2^{8-4} design, involving 8 factor:

X1	X2	X3	X4	X5	X6	X7	X8
-1	-1	-1	-1	-1	-1	-1	-1
1	-1	-1	-1	-1	1	1	1
-1	1	-1	-1	1	-1	1	1
1	1	-1	-1	1	1	-1	-1
-1	-1	1	-1	1	1	1	-1
1	-1	1	-1	1	-1	-1	1
-1	1	1	-1	-1	1	-1	1
1	1	1	-1	-1	-1	1	-1
-1	-1	-1	1	1	1	-1	1
1	-1	-1	1	1	-1	1	-1
-1	1	-1	1	-1	1	1	-1
1	1	-1	1	-1	-1	-1	1
-1	-1	1	1	-1	-1	1	1
1	-1	1	1	-1	1	-1	-1
-1	1	1	1	1	-1	-1	-1
1	1	1	1	1	1	1	1

Latin Hypercube Sampling (LHS) requires specification of sample size.

Example, for COAMPS, we specify 99 parameter values for the 11 parameters.
LHS is "space filling."

Results

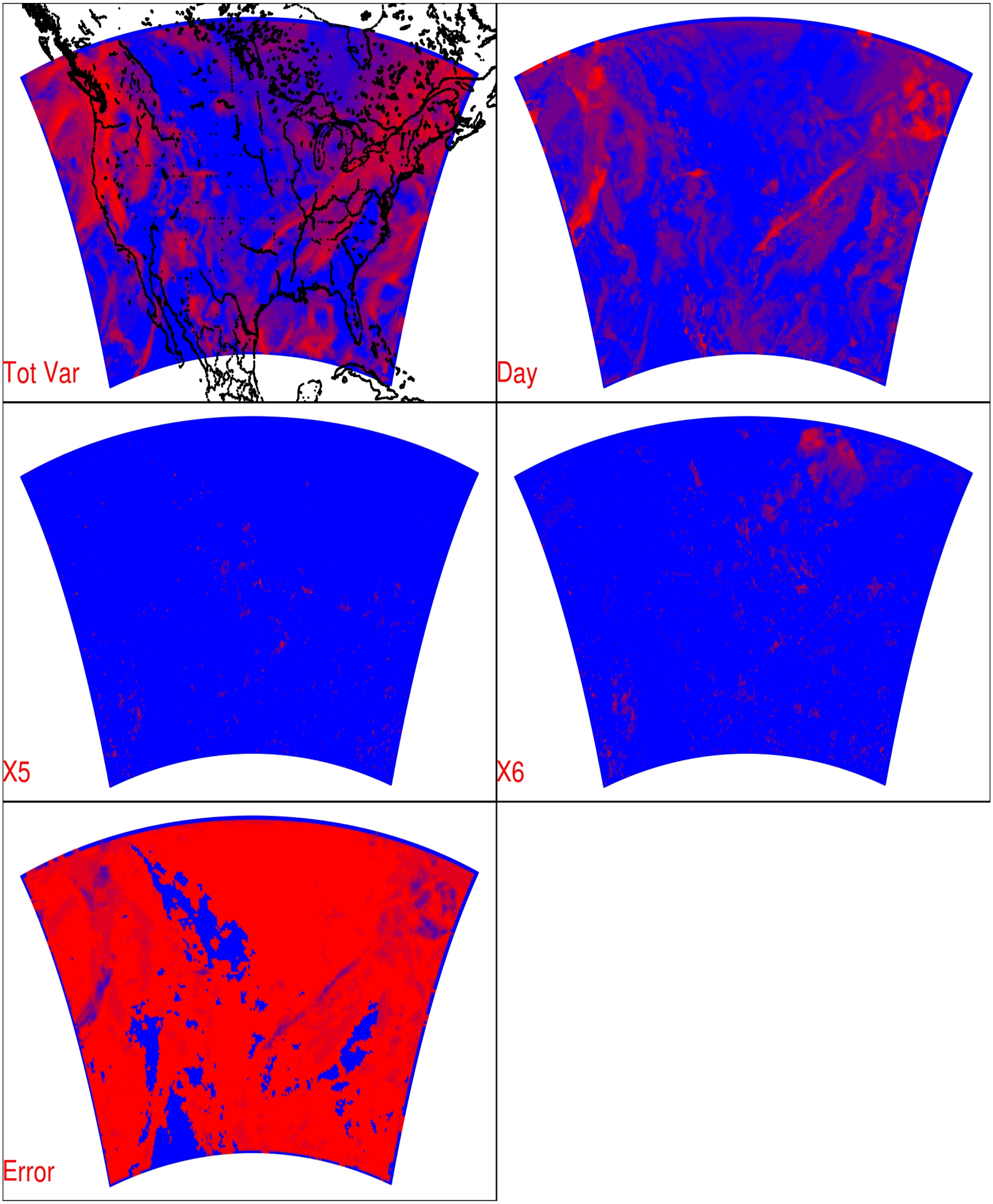

Fig1a. WRF sensitivity of precip, ρ values from random effects model, with LSD.

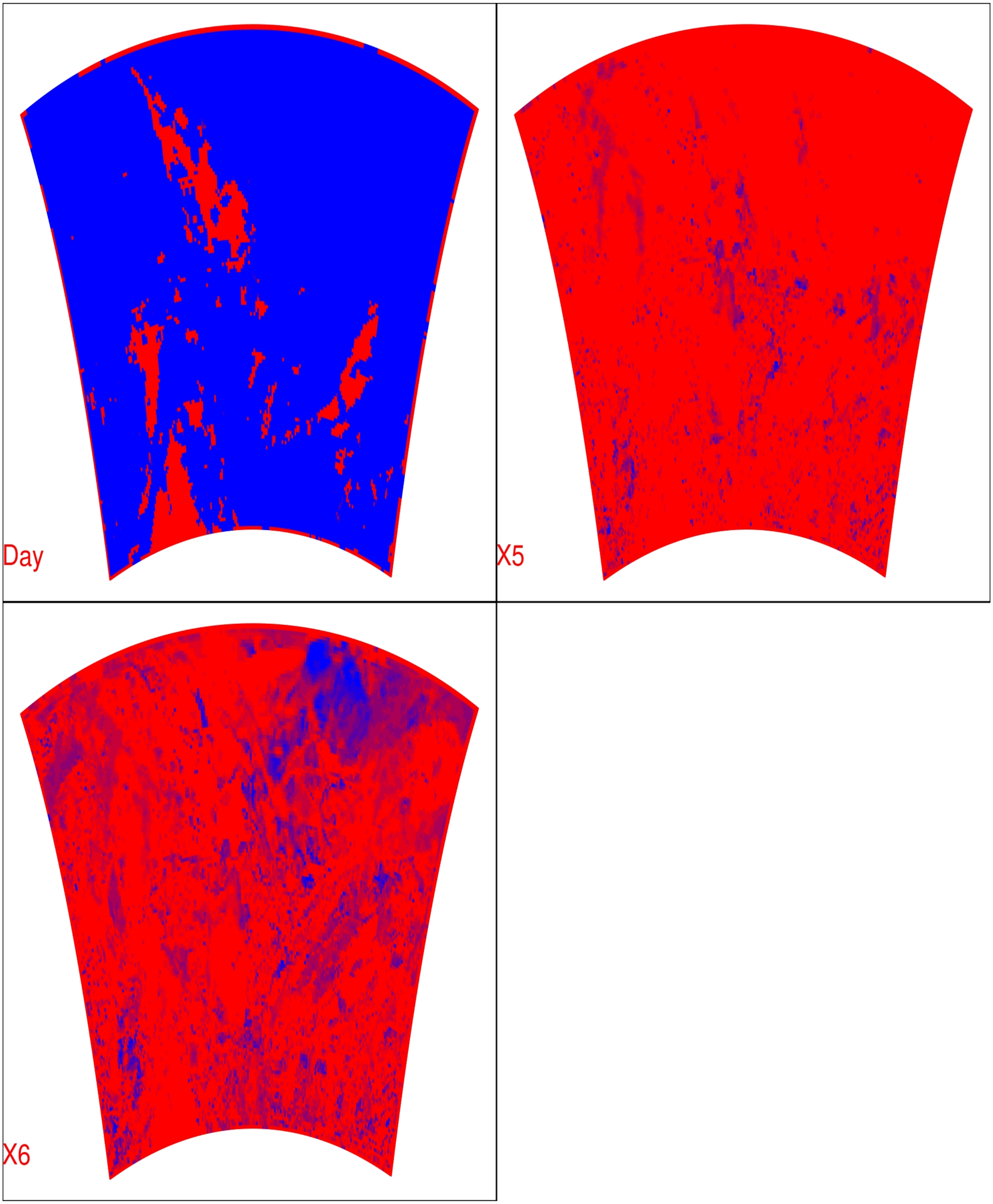

Fig1b. WRF sensitivity of precip, p-values from fixed effects model, with LSD.

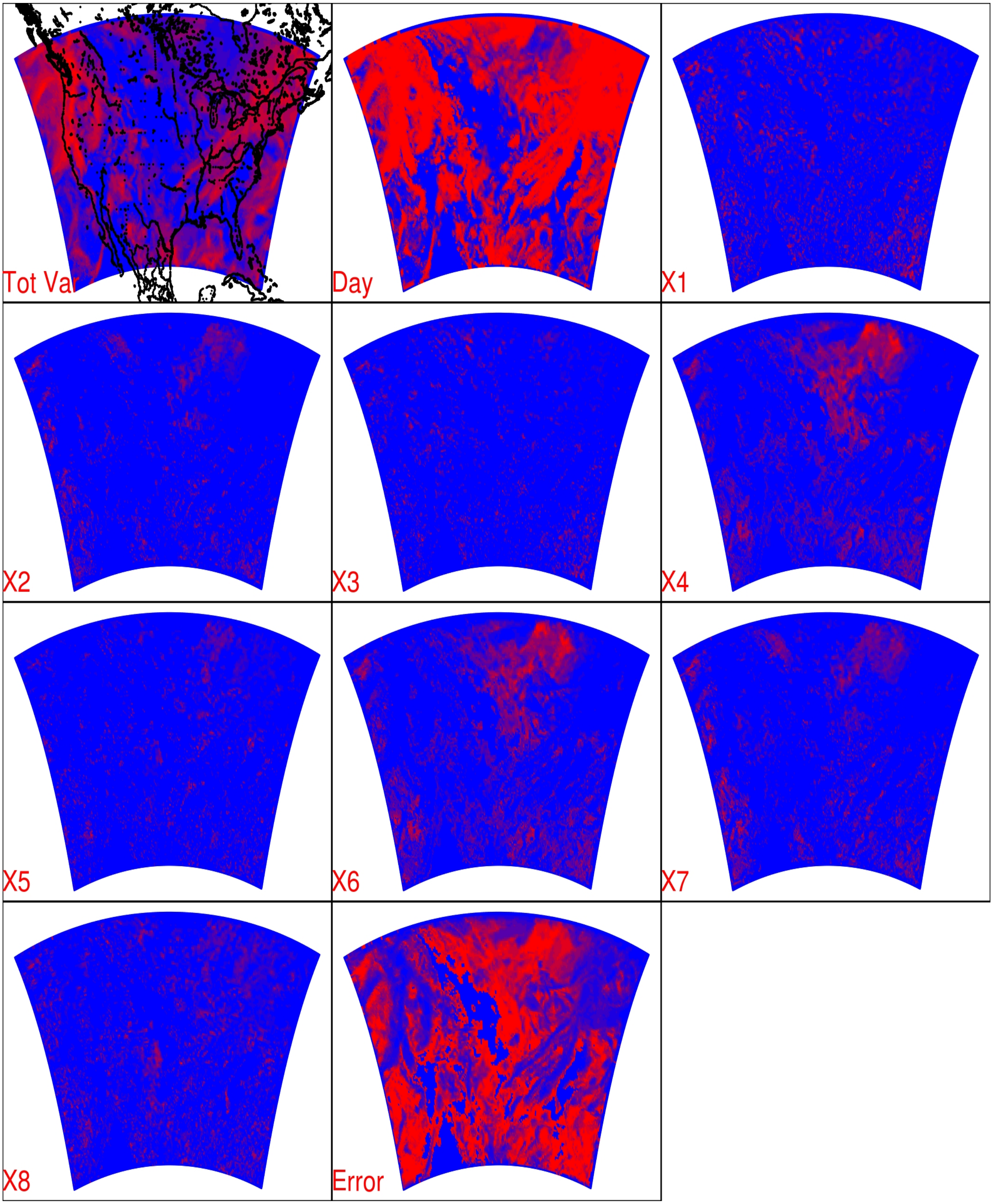

Fig2a. WRF sensitivity of precip, ρ values from 2^{k-p}.

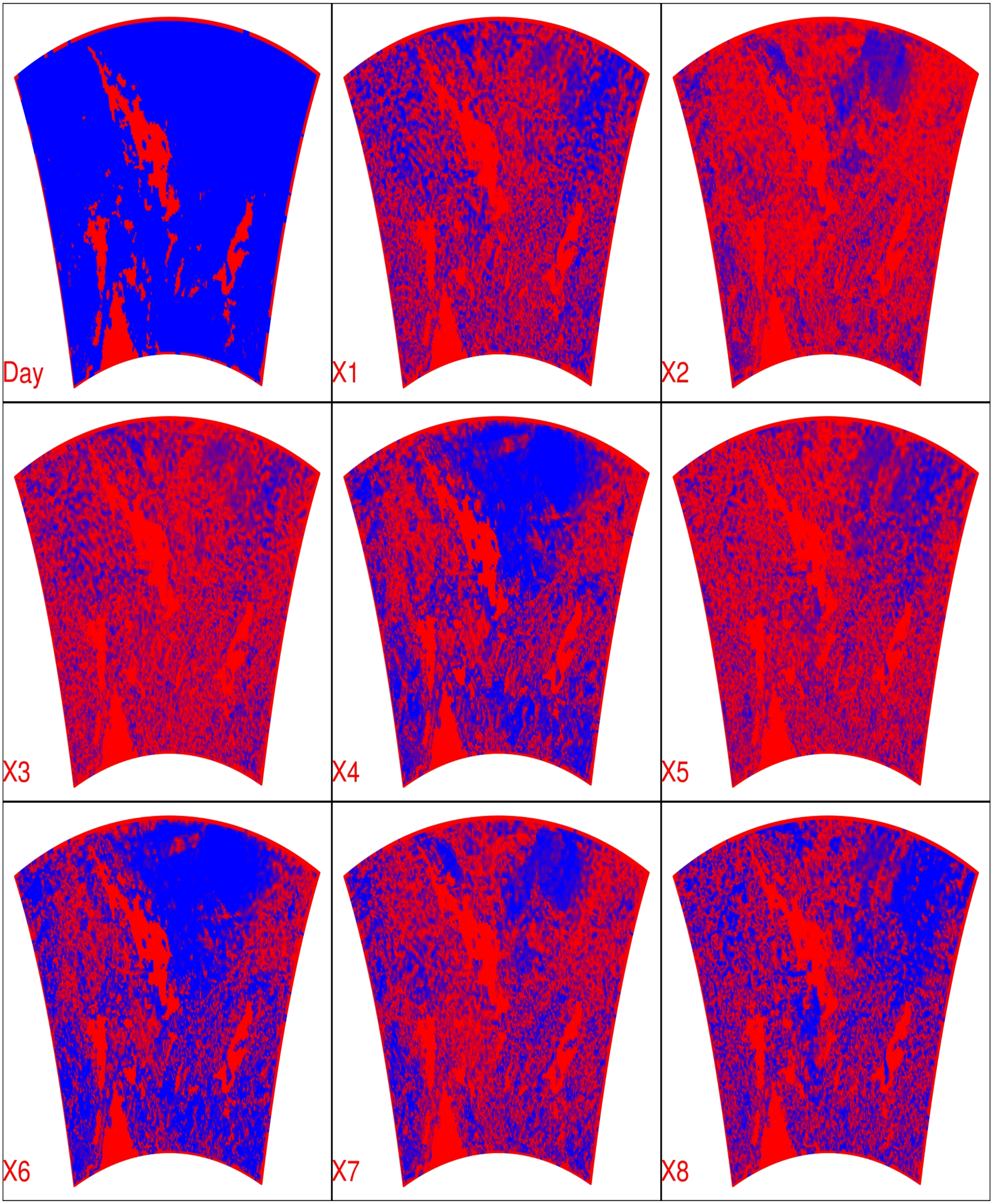

Fig2b. WRF sensitivity of precip, p-values from 2^{k-p}.

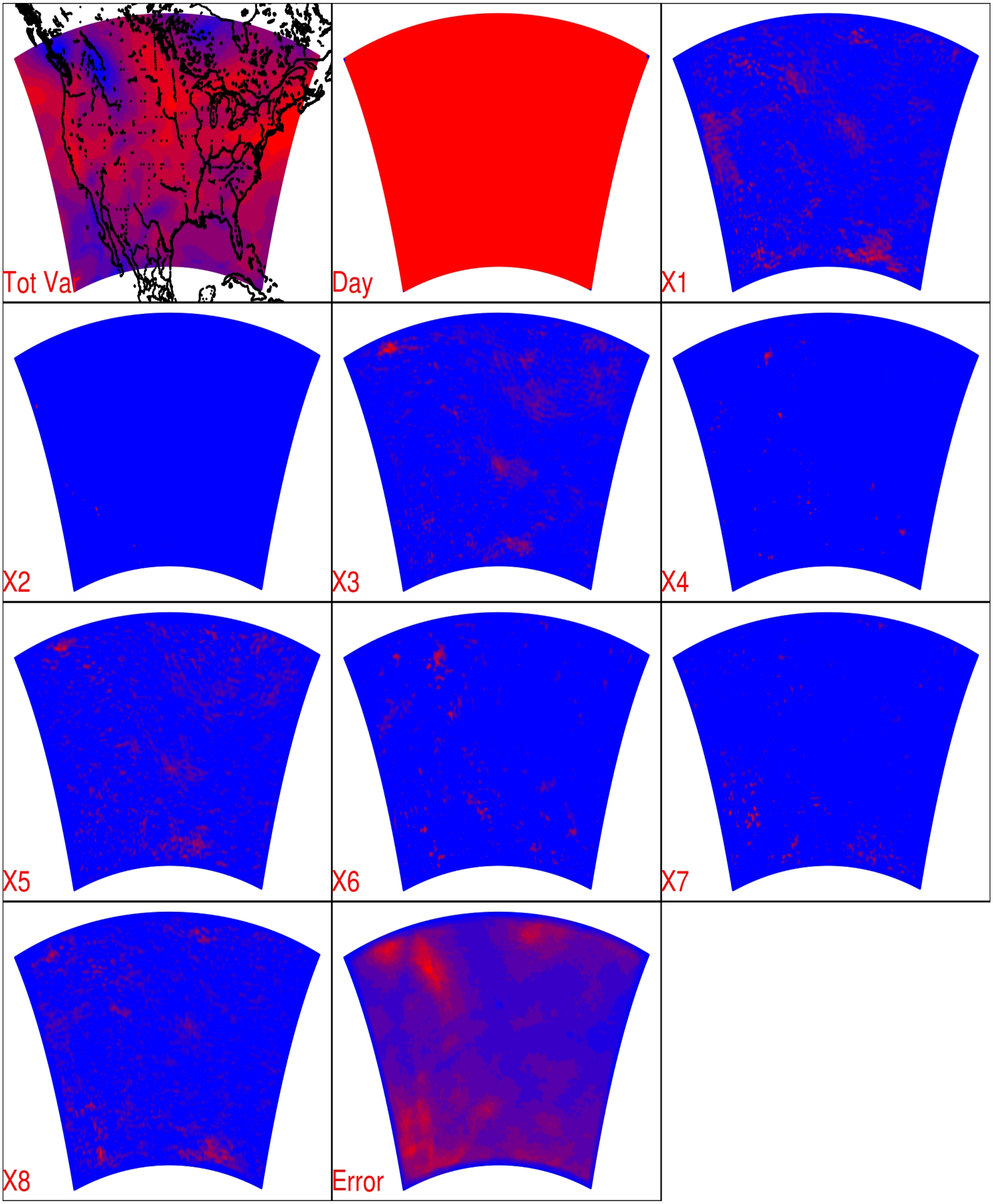

Fig3a. WRF wind speed at 250m, ρ values from 2^{k-p}.

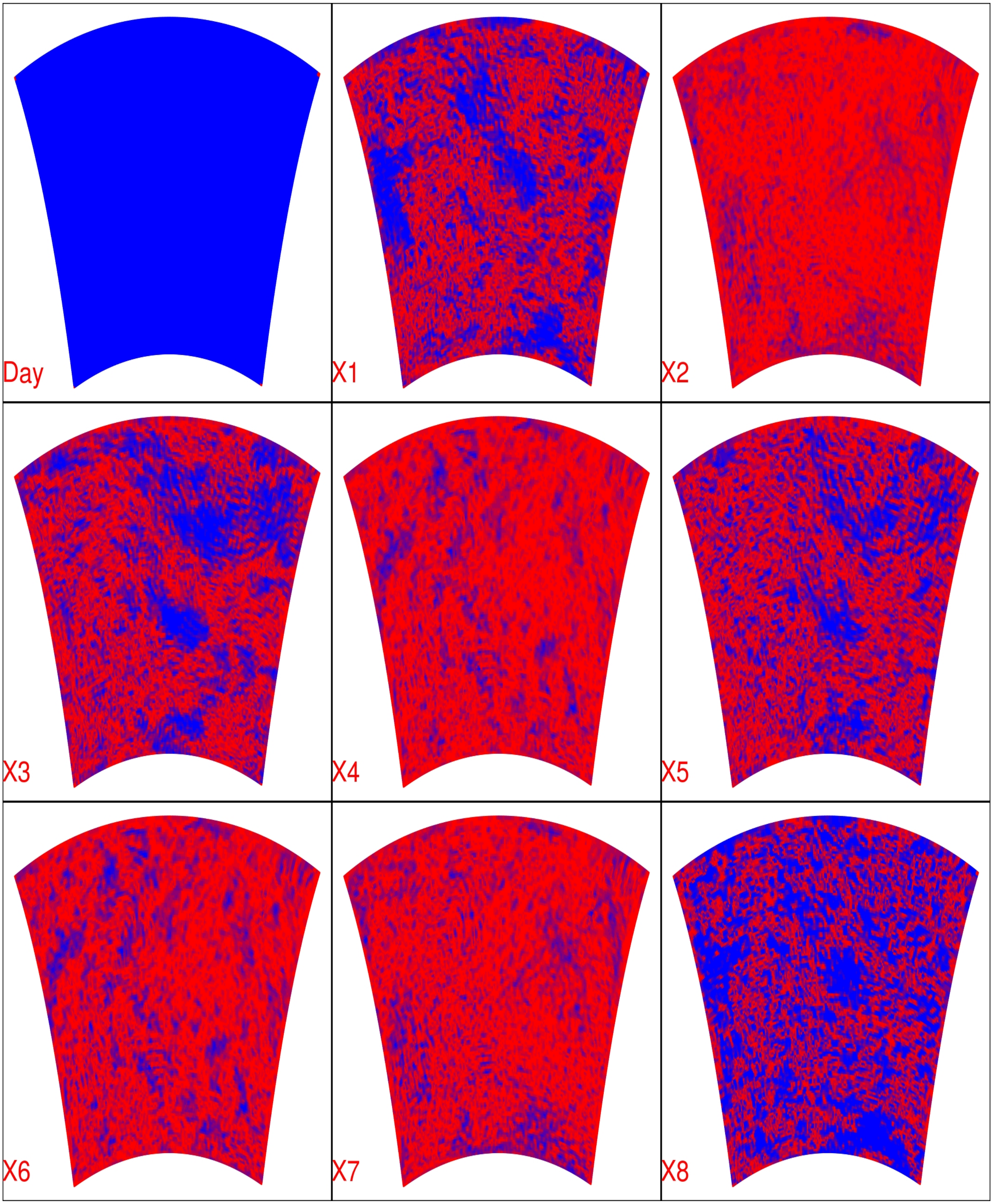

Fig3b. WRF wind speed at 250m, p-values from 2^{k-p}.

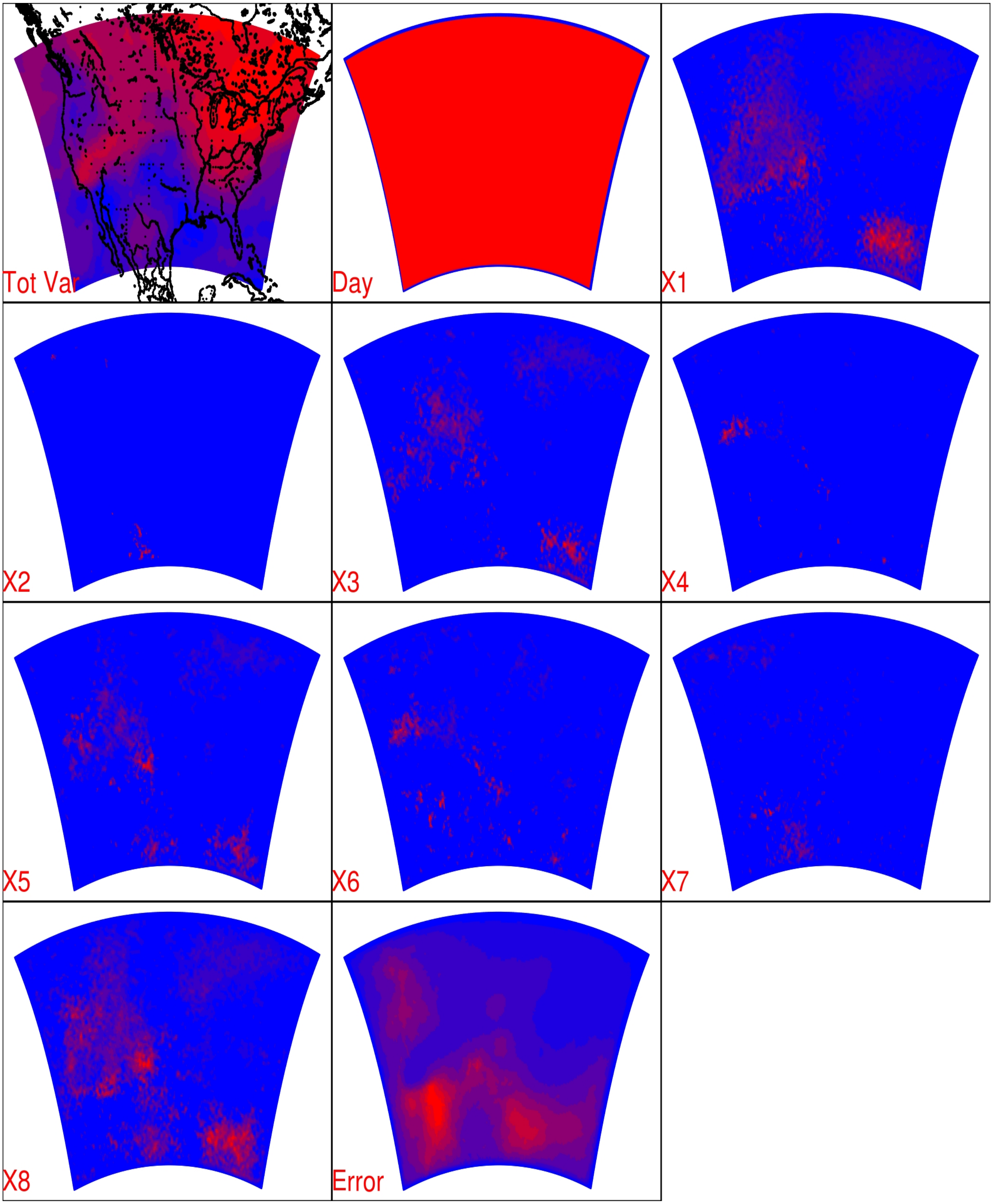

Fig4a. WRF temperature at 500m, ρ values from 2^{k-p}.

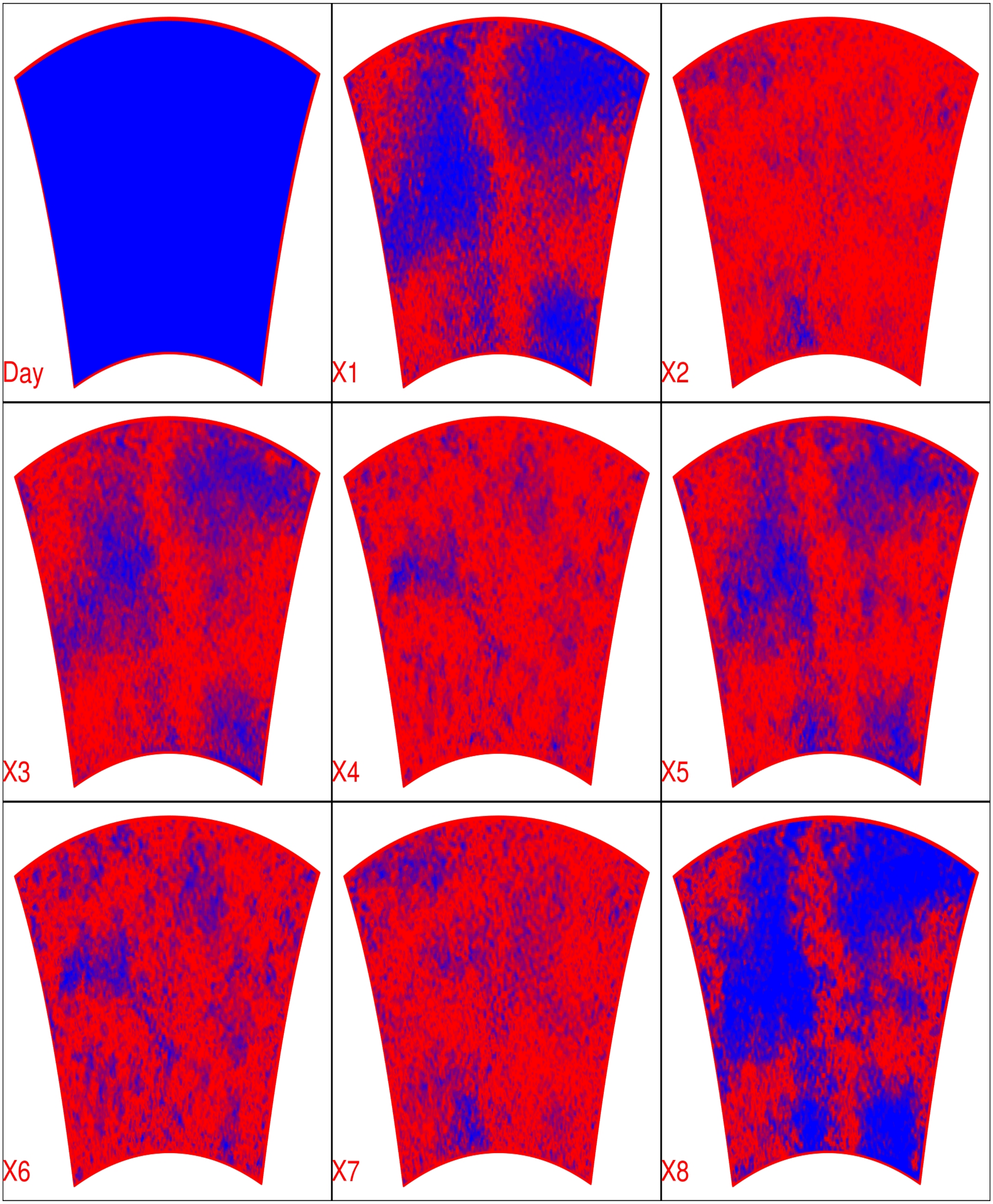

Fig4b. WRF temperature at 500m, p-values from 2^{k-p}.

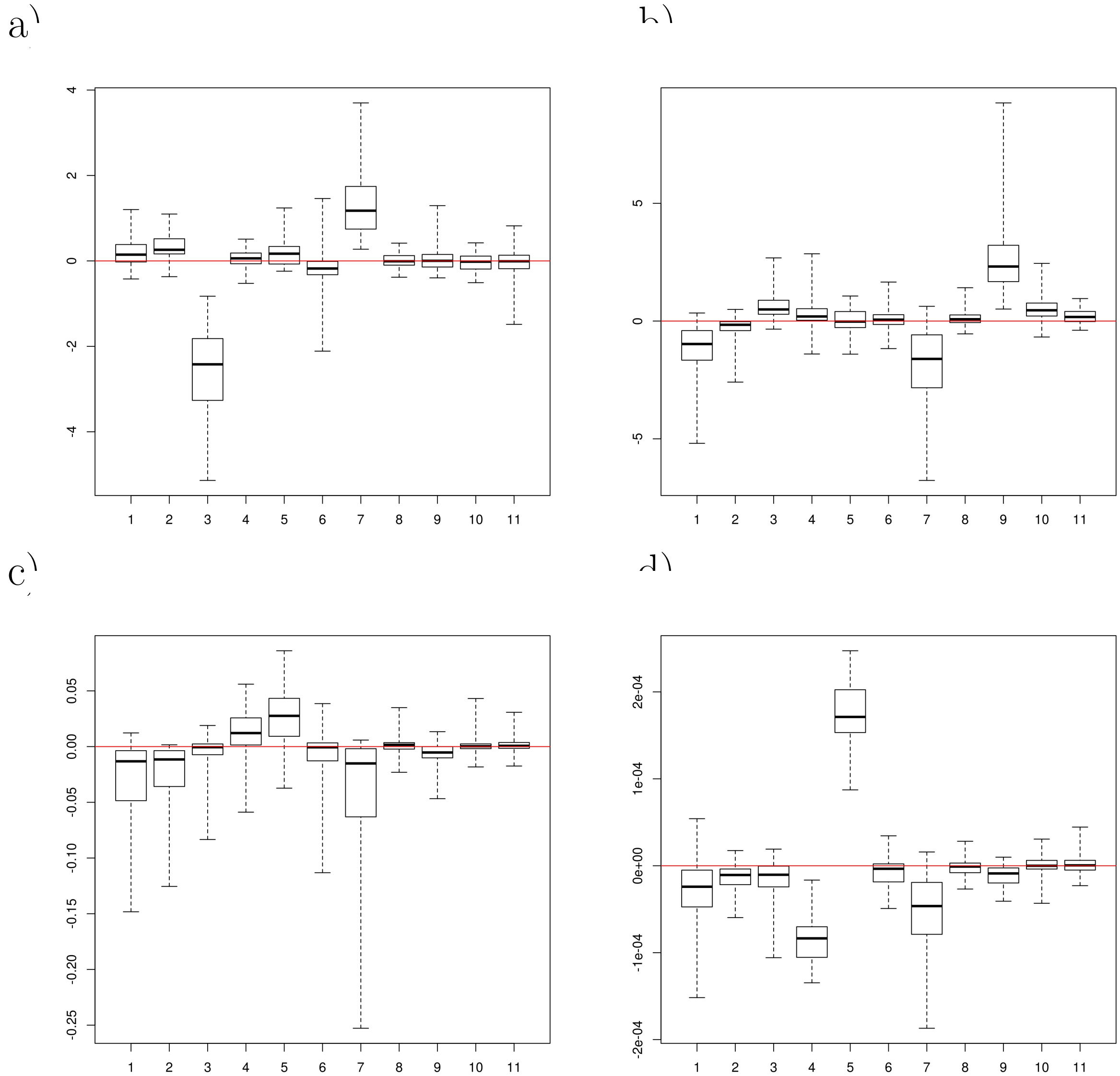

Fig 5. COAMPS sensitivity (on y-axis) of the domain mean of a) convective precipitation, b) stable precipitation, c) surface air temperature, and d) water vapor, with respect to the 11 model parameters (on x-axis). The boxplots display the variability of the sensitivity across the 40 days examined here.

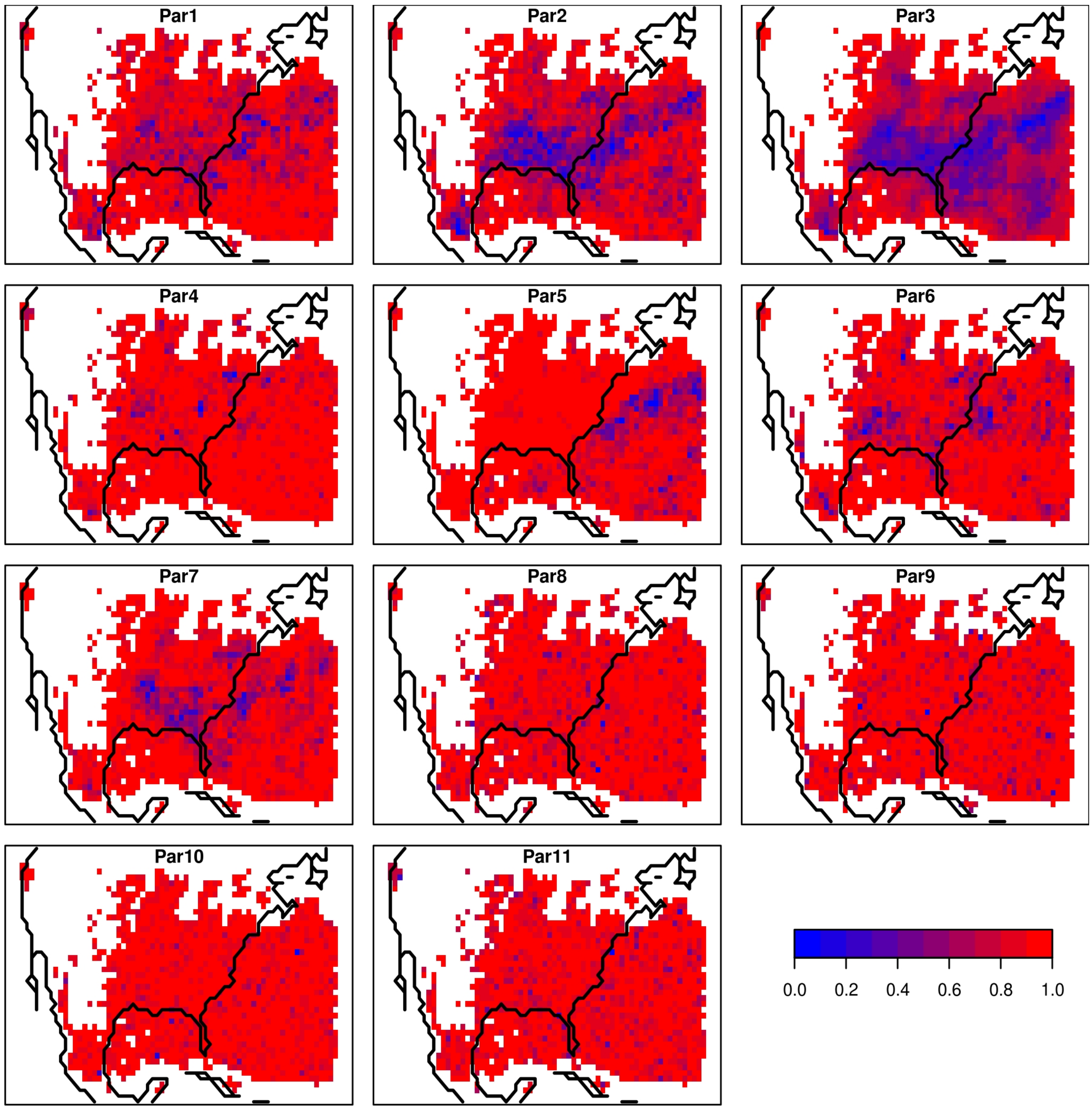

Fig6. COAMPS p-values for convective precip.

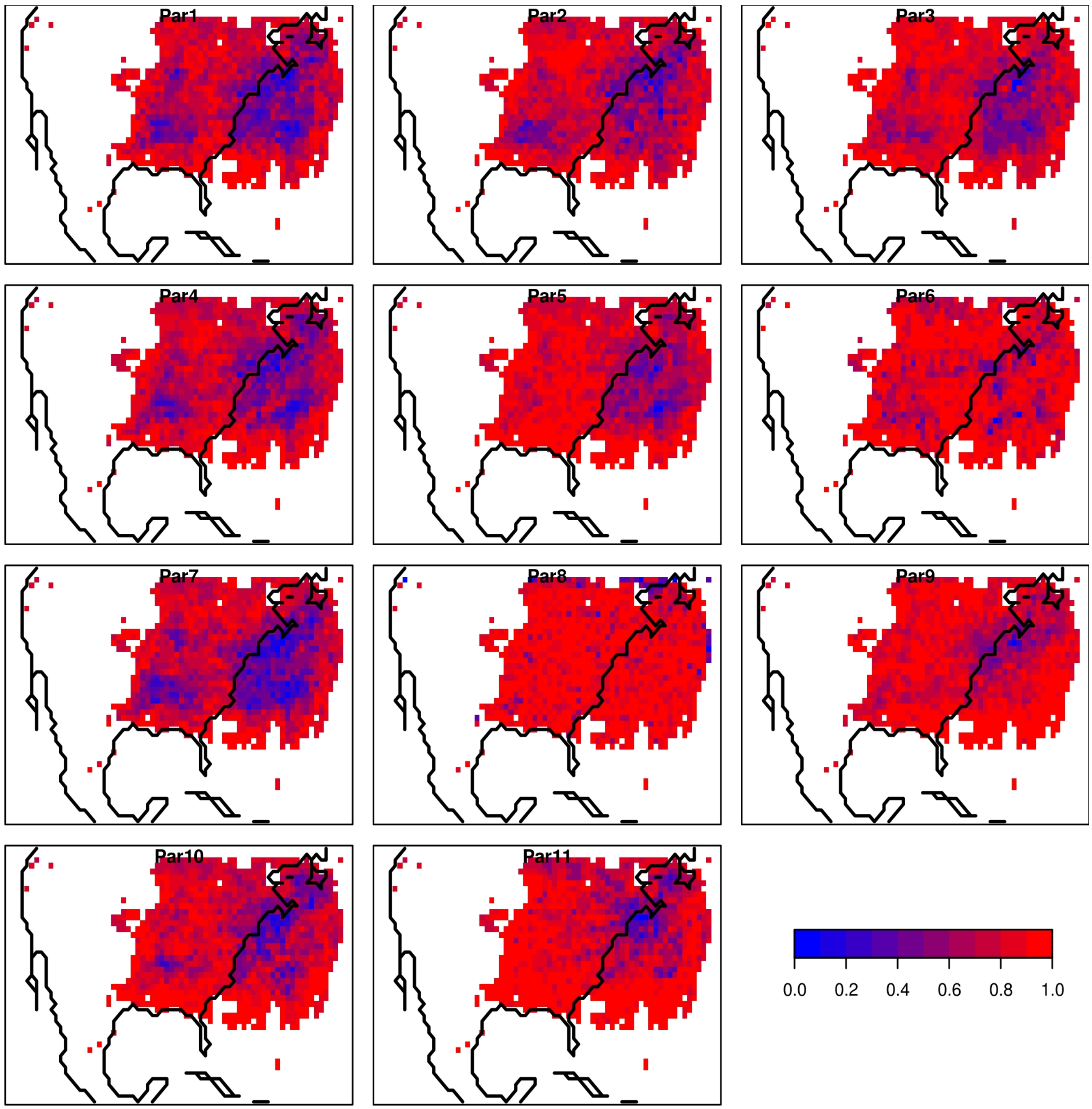

Fig7. COAMPS p-values for stable precip.

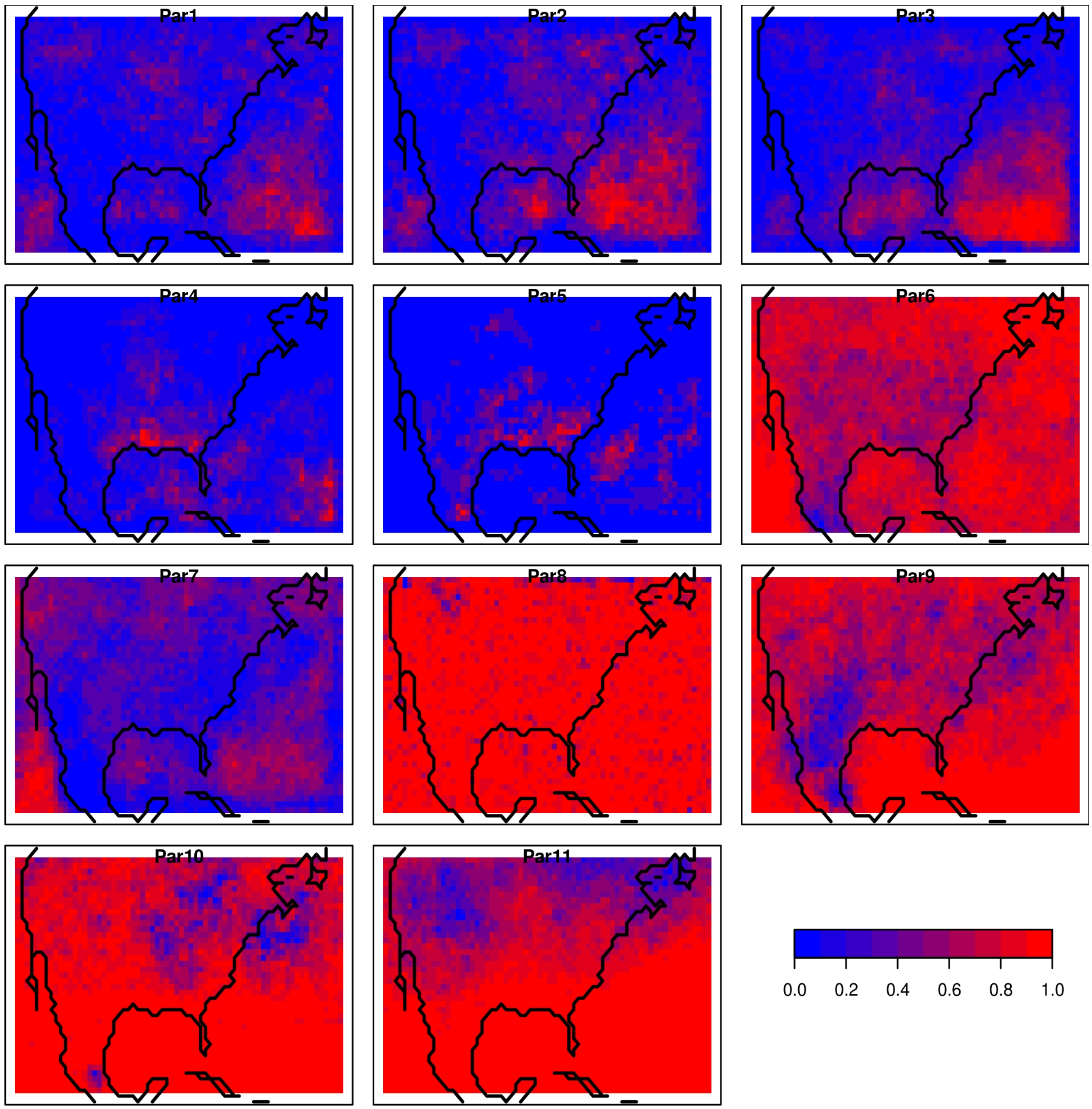

Fig8. COAMPS p-values for air temperature.

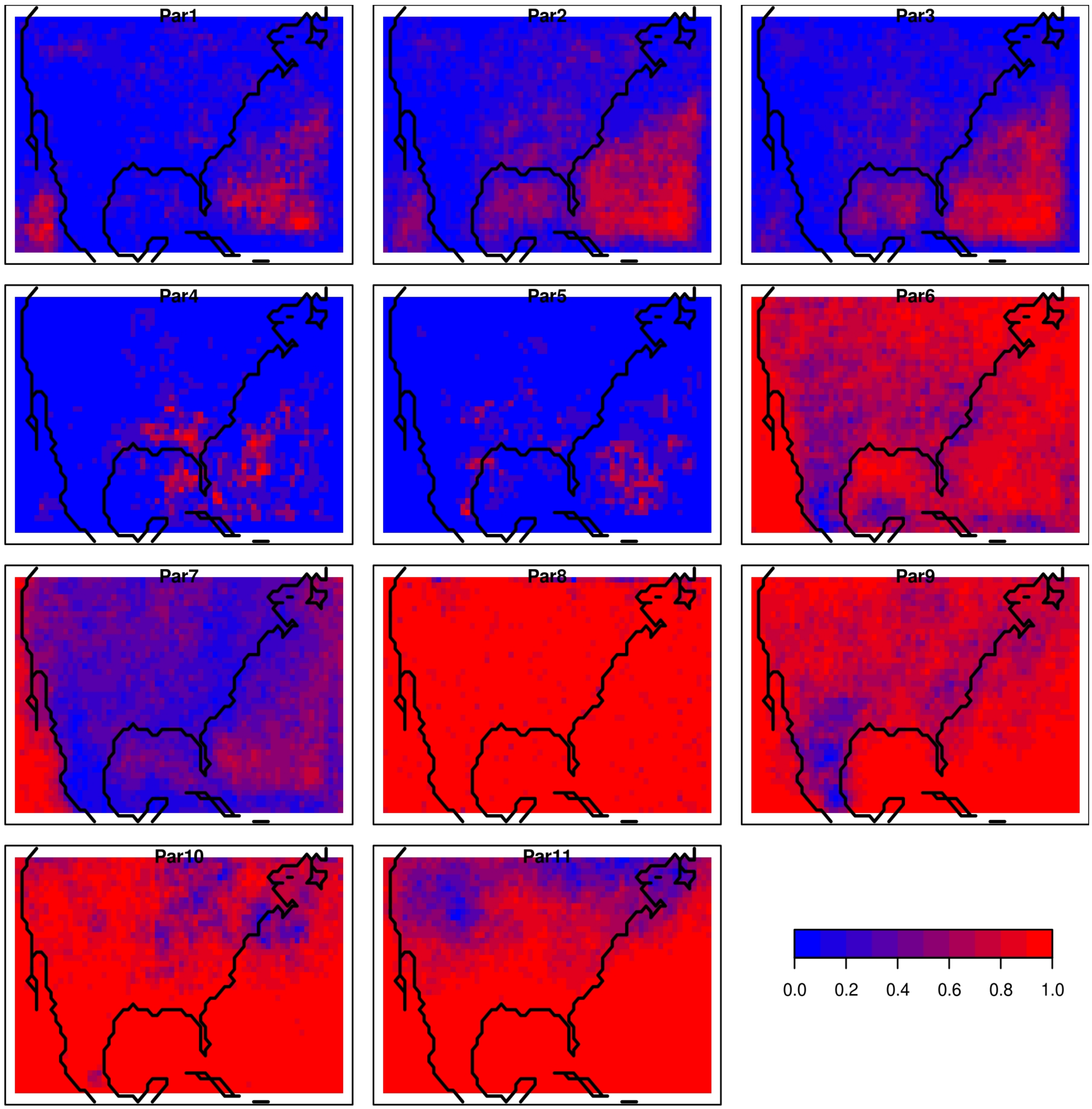

Fig9. COAMPS p-values for water vapor.

Conclusion

Sensitivity analysis on NWP forecasts with respect to model params is hard because

- NWP models are generally too computationally intensive to run many times, hence sampling methods play an important role.

- The data have temporal and spatial structure, hence “careful” statistical modeling is warranted.

- There are ambiguities in terms of what is meant by sensitivity, hence fixed-effects vs. random-effects models and more,

- The final devices are visual, e.g., boxplots & maps of p-values (a feature), which introduce some subjectivity (a drawback).

- But all of these issues are addressable.

We are currently varying similar model parameters in COAMPS and in WRF.